



(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)



1-0049206

(51)^{2020.01} G10L 19/16; H04S 3/00; G10L 19/008

(13) B

(21) 1-2020-05296

(22) 02/07/2019

(86) PCT/US2019/040271 02/07/2019

(87) WO2020/010064 A1 09/01/2020

(30) 62/693,246 02/07/2018 US

(45) 25/07/2025 448

(43) 25/05/2021 398A

(73) 1. DOLBY LABORATORIES LICENSING CORPORATION (US)

1275 Market Street San Francisco, California 94103 (US)

2. DOLBY INTERNATIONAL AB (NL)

Apollo Building, 3E Herikerbergweg 1-35 1101 CN Amsterdam Zuidoost (NL)

(72) BRUHN, Stefan (DE); TORRES, Juan Felix (US).

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

(54) PHƯƠNG PHÁP VÀ THIẾT BỊ MÃ HÓA ĐỂ TẠO RA DÒNG BIT, PHƯƠNG
PHÁP VÀ THIẾT BỊ GIẢI MÃ ĐỂ THU ĐƯỢC DỮ LIỆU LIÊN QUAN ĐẾN TÍN
HIỆU ÂM THANH NỔI SÓNG ĐỘNG AMBISONIC TỪ DÒNG BIT

(21) 1-2020-05296

(57) Sáng chế đề cập đến phương pháp (500) tạo dòng bit (101), trong đó dòng bit (101) bao gồm chuỗi siêu khung (400) cho chuỗi các khung của tín hiệu âm thanh nổi sống động (111). Phương pháp (500) bao gồm các bước, theo cách lặp lại cho chuỗi siêu khung (400), chèn (501) dữ liệu âm thanh mã hóa (206) cho một hoặc nhiều khung của một hoặc nhiều tín hiệu kênh giảm mức (203) thu được từ tín hiệu âm thanh nổi sống động (111), vào các trường dữ liệu (411, 421, 412, 422) của siêu khung (400); và chèn (502) siêu dữ liệu (202, 205) để tái tạo một hoặc nhiều khung của tín hiệu âm thanh nổi sống động (111) từ dữ liệu âm thanh mã hóa (206), vào trường siêu dữ liệu (403) của siêu khung (400). Sáng chế còn đề cập đến thiết bị mã hóa được tạo cấu hình để tạo ra dòng bit, và phương pháp và thiết bị giải mã để thu được dữ liệu liên quan đến tín hiệu âm thanh nổi sống động từ dòng bit.

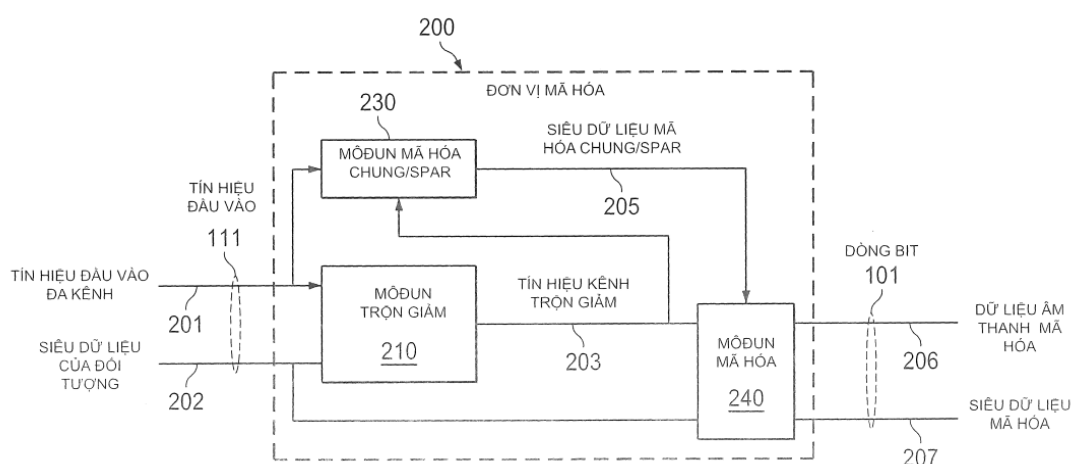


FIG. 2

Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến các tín hiệu âm thanh nổi sống động có thể bao gồm các tín hiệu biểu diễn trường âm thanh, nhất là các tín hiệu ambisonics. Cụ thể, sáng chế đề cập đến việc tạo ra và giải mã dòng bit chứa tín hiệu âm thanh nổi sống động.

Tình trạng kỹ thuật của sáng chế

Âm thanh hoặc trường âm thanh trong môi trường nghe của người nghe ở vị trí nghe có thể được mô tả nhờ sử dụng tín hiệu ambisonics. Tín hiệu ambisonics có thể được xem là tín hiệu âm thanh đa kênh, với mỗi kênh tương ứng với một đặc tuyến hướng tính cụ thể của trường âm thanh ở vị trí nghe của người nghe. Tín hiệu ambisonics có thể được mô tả nhờ sử dụng hệ tọa độ Đề-cát ba chiều (three-dimensional - 3D), với gốc của hệ tọa độ tương ứng với vị trí nghe, trục x hướng về phía trước, trục y hướng về bên trái và trục z hướng lên trên.

Bằng cách làm tăng số lượng tín hiệu hoặc kênh âm thanh và bằng cách làm tăng số lượng đặc tuyến hướng tính tương ứng (và các chức năng quét tương ứng), độ chính xác mô tả trường âm thanh có thể được tăng lên. Ví dụ, tín hiệu ambisonics bậc một bao gồm 4 kênh hoặc dạng sóng, đó là kênh W biểu thị thành phần đẳng hướng của trường âm thanh, kênh X mô tả trường âm thanh có đặc tuyến hướng tính lưỡng cực tương ứng với trục x, kênh Y mô tả trường âm thanh có đặc tuyến hướng tính lưỡng cực tương ứng với trục y, và kênh Z mô tả trường âm thanh có đặc tuyến hướng tính lưỡng cực tương ứng với trục z. Tín hiệu ambisonics bậc hai bao gồm 9 kênh bao gồm 4 kênh của tín hiệu ambisonics bậc một (còn được gọi là định dạng B) cộng với 5 kênh bổ sung cho các đặc tuyến hướng tính khác nhau. Nhìn chung, tín hiệu ambisonics bậc L bao gồm $(L+1)^2$ kênh bao gồm L^2 kênh của các tín hiệu ambisonics bậc $(L-1)$ cộng với $[(L+1)^2 - L^2]$ kênh bổ sung cho các đặc tuyến hướng tính bổ sung (khi sử dụng định dạng ambisonics 3D). Các tín hiệu ambisonics bậc L đối với $L > 1$ có thể được gọi là các tín hiệu ambisonics bậc cao hơn (higher order ambisonics - HOA).

Tín hiệu HOA có thể được sử dụng để mô tả trường âm thanh 3D độc lập với việc bố trí loa, mà được sử dụng để kết xuất tín hiệu HOA. Các bố trí loa ví dụ bao gồm các tai nghe hoặc một hoặc nhiều bố trí loa phóng thanh hoặc môi trường kết xuất thực tế ảo. Do vậy, có thể có lợi nếu cung cấp tín hiệu HOA đến bộ kết xuất âm thanh, để cho phép bộ kết xuất âm thanh thích ứng một cách linh hoạt với các bố trí loa khác nhau.

Các tín hiệu biểu diễn trường âm thanh (Soundfield representation - SR), như các tín hiệu ambisonics, có thể được bổ sung các đối tượng âm thanh và/hoặc các tín hiệu đa kênh, để tạo ra tín hiệu âm thanh nổi sống động (immersive audio - IA). Sáng chế nhằm xử lý các vấn đề kỹ thuật về truyền và/hoặc lưu trữ tín hiệu IA, với chất lượng cảm nhận cao theo cách có hiệu quả về băng thông. Cụ thể, sáng chế giải quyết vấn đề kỹ thuật của việc tạo dòng bit hiệu quả mà biểu thị tín hiệu IA. Các vấn đề kỹ thuật này được giải quyết bởi các điểm yêu cầu bảo hộ độc lập. Các ví dụ ưu tiên được mô tả trong các điểm yêu cầu bảo hộ phụ thuộc.

Bản chất kỹ thuật của sáng chế

Theo một khía cạnh, sáng chế mô tả phương pháp tạo dòng bit, trong đó dòng bit bao gồm chuỗi siêu khung cho chuỗi các khung của tín hiệu âm thanh nổi sống động. Phương pháp bao gồm, theo cách lặp lại cho chuỗi siêu khung, chèn dữ liệu âm thanh mã hóa cho một hoặc nhiều khung của một hoặc nhiều tín hiệu kênh giảm mức thu được từ tín hiệu âm thanh nổi sống động, vào các trường dữ liệu của siêu khung. Hơn thế nữa, phương pháp bao gồm chèn siêu dữ liệu, nhất là siêu dữ liệu mã hóa, để tái tạo một hoặc nhiều khung của tín hiệu âm thanh nổi sống động từ dữ liệu âm thanh mã hóa, vào trường siêu dữ liệu của siêu khung.

Theo một khía cạnh khác, sáng chế mô tả phương pháp thu được dữ liệu liên quan đến tín hiệu âm thanh nổi sống động từ dòng bit, trong đó dòng bit bao gồm chuỗi siêu khung cho chuỗi các khung của tín hiệu âm thanh nổi sống động. Phương pháp bao gồm, theo cách lặp lại cho chuỗi siêu khung, trích xuất dữ liệu âm thanh mã hóa cho một hoặc nhiều khung của một hoặc nhiều tín hiệu kênh giảm mức thu được từ tín hiệu âm thanh nổi sống động, từ các trường dữ liệu của siêu khung. Hơn thế nữa, phương pháp bao gồm bước trích xuất siêu dữ liệu để tái tạo một hoặc nhiều khung của tín hiệu âm thanh nổi sống động từ dữ liệu âm thanh mã hóa, từ trường siêu dữ liệu của siêu khung.

Theo một khía cạnh khác, sáng chế mô tả chương trình phần mềm. Chương trình phần mềm này có thể được làm thích ứng để thực thi trên bộ xử lý và để tiến hành các bước của phương pháp được nêu trong sáng chế khi tiến hành trên bộ xử lý.

Theo một khía cạnh khác, sáng chế mô tả phương tiện lưu trữ. Phương tiện lưu trữ này có thể bao gồm chương trình phần mềm được làm thích ứng để thực thi trên bộ xử lý và để tiến hành các bước trong phương pháp được nêu trong sáng chế khi tiến hành trên bộ xử lý.

Theo một khía cạnh khác, sáng chế mô tả sản phẩm chương trình máy tính. Chương trình máy tính có thể bao gồm các lệnh có thể thực thi để thực hiện các bước trong phương pháp được nêu trong sáng chế khi chạy trên máy tính.

Theo một khía cạnh khác, sáng chế mô tả siêu khung của dòng bit, trong đó dòng bit bao gồm chuỗi siêu khung cho chuỗi các khung của tín hiệu âm thanh nổi sống động. Siêu khung bao gồm các trường dữ liệu cho dữ liệu âm thanh mã hóa cho một hoặc nhiều (nhất là cho nhiều) khung của một hoặc nhiều tín hiệu kênh giảm mức, thu được từ tín hiệu âm thanh nổi sống động. Hơn thế nữa, siêu khung bao gồm (một) trường siêu dữ liệu cho siêu dữ liệu được làm thích ứng để tái tạo một hoặc nhiều (nhất là nhiều) khung của tín hiệu âm thanh nổi sống động từ dữ liệu âm thanh mã hóa.

Theo một khía cạnh khác, sáng chế mô tả thiết bị mã hóa được tạo cấu hình để tạo ra dòng bit. Dòng bit bao gồm chuỗi siêu khung cho chuỗi các khung của tín hiệu âm thanh nổi sống động. Thiết bị mã hóa được tạo cấu hình để, theo cách lặp lại cho chuỗi siêu khung, chèn dữ liệu âm thanh mã hóa cho một hoặc nhiều (nhất là nhiều) khung của một hoặc nhiều tín hiệu kênh giảm mức thu được từ tín hiệu âm thanh nổi sống động, vào các trường dữ liệu của siêu khung; và chèn siêu dữ liệu để tái tạo một hoặc nhiều (nhất là nhiều) khung của tín hiệu âm thanh nổi sống động từ dữ liệu âm thanh mã hóa vào trường siêu dữ liệu của siêu khung.

Theo một khía cạnh khác, thiết bị giải mã được tạo cấu hình để thu được dữ liệu liên quan đến tín hiệu âm thanh nổi sống động từ dòng bit, trong đó dòng bit bao gồm chuỗi siêu khung cho chuỗi các khung của tín hiệu âm thanh nổi sống động. Thiết bị giải mã được tạo cấu hình để, theo cách lặp lại cho chuỗi siêu khung, trích xuất dữ liệu âm thanh mã hóa cho một hoặc nhiều (nhất là nhiều) khung của một hoặc nhiều tín hiệu kênh

giảm mức thu được từ tín hiệu âm thanh nổi sống động, từ các trường dữ liệu siêu khung; và trích xuất siêu dữ liệu để tái tạo một hoặc nhiều (nhất là nhiều) khung của tín hiệu âm thanh nổi sống động từ dữ liệu âm thanh mã hóa từ trường siêu dữ liệu của siêu khung.

Nên chú ý rằng các phương pháp, thiết bị và hệ thống bao gồm các phương án được ưu tiên của nó như được nêu trong sáng chế có thể được sử dụng riêng lẻ hoặc kết hợp với phương pháp, thiết bị và hệ thống khác được bộc lộ trong bản mô tả. Hơn thế nữa, tất cả các khía cạnh của phương pháp, thiết bị và hệ thống được nêu trong đơn sáng chế có thể được kết hợp tùy ý. Cụ thể, các dấu hiệu trong yêu cầu bảo hộ có thể được kết hợp với nhau theo cách bất kỳ.

Mô tả vắn tắt các hình vẽ

Sáng chế được giải thích sau đây theo cách minh họa có dẫn chiếu đến các hình vẽ kèm theo, trong đó:

Fig.1 thể hiện hệ thống mã hóa làm ví dụ;

Fig.2 thể hiện đơn vị mã hóa làm ví dụ để mã hóa tín hiệu âm thanh nổi sống động;

Fig.3 thể hiện một đơn vị giải mã khác để giải mã tín hiệu âm thanh nổi sống động;

Fig.4 thể hiện cấu trúc siêu khung làm ví dụ cho tín hiệu âm thanh nổi sống động, nhất là cho dữ liệu mã hóa biểu thị tín hiệu âm thanh nổi sống động;

Fig.5 thể hiện lưu đồ của phương pháp làm ví dụ tạo dòng bit chứa chuỗi siêu khung biểu thị tín hiệu âm thanh nổi sống động; và

Fig.6 thể hiện lưu đồ của phương pháp làm ví dụ để trích xuất thông tin từ dòng bit chứa chuỗi siêu khung biểu thị tín hiệu nổi sống động.

Mô tả chi tiết sáng chế

Như nêu trên, sáng chế đề cập đến việc mã hóa hiệu quả các tín hiệu âm thanh nổi sống động như các tín hiệu HOA, các tín hiệu âm thanh đa kênh và/hoặc đối tượng, trong

đó nhất là các tín hiệu HOA ở đây thường được gọi là các tín hiệu biểu diễn trường âm thanh (SR). Hơn thế nữa, sáng chế đề cập đến việc lưu trữ hoặc truyền tín hiệu âm thanh nổi sống động (IA) trên mạng truyền trong dòng bit.

Như đã nêu ở phần giới thiệu, tín hiệu SR có thể bao gồm số lượng kênh hoặc dạng sóng tương đối cao, trong đó các kênh khác nhau đề cập đến các chức năng quét khác nhau và/hoặc đến các đặc tuyến hướng tính khác nhau. Ví dụ, tín hiệu ambisonics bậc một 3D bậc thứ L (FOA) hoặc tín hiệu HOA bao gồm $(L+1)^2$ kênh. Tín hiệu ambisonics bậc một (First Order Ambisonics - FOA) là tín hiệu Ambisonics bậc $L=1$ chứa 4 kênh. Tín hiệu SR có thể được biểu diễn ở các định dạng khác nhau.

Trường âm thanh có thể được xem là bao gồm một hoặc nhiều sự kiện âm thanh bắt nguồn từ các hướng tùy ý quanh vị trí nghe. Kết quả là các vị trí của

một hoặc nhiều sự kiện âm thanh có thể được xác định trên bề mặt của hình cầu (với vị trí nghe hoặc tham chiếu là ở tâm của hình cầu).

Định dạng trường âm thanh như FOA hoặc ambisonics bậc cao hơn (Higher Order Ambisonics - HOA) được xác định theo cách cho phép trường âm thanh được kết xuất qua các bố trí loa tùy ý (tức là các hệ thống kết xuất tùy ý). Tuy nhiên, các hệ thống kết xuất (như hệ thống Dolby Atmos) thường bị ràng buộc theo hướng là các góc ngẩng có thể của loa được cố định ở số lượng mặt phẳng xác định (ví dụ, mặt phẳng ngang tai (nằm ngang), mặt phẳng trần hoặc trên và/hoặc mặt phẳng sàn hoặc thấp hơn). Do vậy, khái niệm trường âm thanh cầu lý tưởng có thể được thay đổi thành trường âm thanh gồm có các đối tượng âm thanh nằm ở các vòng khác nhau ở các độ cao khác nhau trên bề mặt hình cầu (tương tự với các vòng lồng nhau tạo thành hình tổ ong).

Như được thể hiện trên Fig.1, hệ thống mã hóa âm thanh 100 bao gồm đơn vị mã hóa 110 và đơn vị giải mã 120. Đơn vị mã hóa 110 có thể được tạo cấu hình để tạo ra dòng bit 101 để truyền đến đơn vị giải mã 120 dựa trên tín hiệu đầu vào 111, trong đó tín hiệu đầu vào 111 có thể bao gồm hoặc có thể là tín hiệu âm thanh nổi sống động (được sử dụng ví dụ, cho các ứng dụng thực tế ảo (Virtual Reality - VR)). Tín hiệu âm thanh nổi sống động 111 có thể bao gồm tín hiệu SR, tín hiệu đa kênh và/hoặc nhiều đối tượng (mỗi đối tượng gồm có tín hiệu đối tượng và siêu dữ liệu đối tượng). Đơn vị giải mã 120

có thể được tạo cấu hình để cung cấp tín hiệu đầu ra 121 dựa trên dòng bit 101, trong đó tín hiệu đầu ra 121 có thể bao gồm hoặc có thể là tín hiệu âm thanh nổi sống động tái tạo.

Fig.2 minh họa đơn vị mã hóa làm ví dụ 110, 200. Đơn vị mã hóa 200 có thể được tạo cấu hình để mã hóa tín hiệu đầu vào 111, trong đó tín hiệu đầu vào 111 có thể là tín hiệu âm thanh nổi sống động (IA) 111. Tín hiệu IA 111 có thể bao gồm tín hiệu đầu vào đa kênh 201. Tín hiệu đầu vào đa kênh 201 có thể bao gồm tín hiệu SR và một hoặc nhiều tín hiệu đối tượng. Hơn thế nữa, siêu dữ liệu đối tượng 202 cho nhiều tín hiệu đối tượng có thể được cung cấp dưới dạng một phần của tín hiệu IA 111. Tín hiệu IA 111 có thể được tạo ra bởi máy tiêu thụ nội dung, trong đó máy tiêu thụ nội dung có thể được tạo cấu hình để thu được các đối tượng và/hoặc tín hiệu SR từ nội dung IA (phức tạp) như nội dung VR mà có thể bao gồm tín hiệu SR, một hoặc nhiều tín hiệu đa kênh và/hoặc một hoặc nhiều đối tượng.

Đơn vị mã hóa 200 bao gồm môđun giảm mức 210 được tạo cấu hình để giảm mức tín hiệu đầu vào đa kênh 201 thành nhiều tín hiệu kênh giảm mức 203. Nhiều tín hiệu kênh giảm mức 203 có thể tương ứng với tín hiệu SR, nhất là với tín hiệu ambisonics bậc một (FOA). Việc giảm mức có thể được thực hiện trong miền băng con hoặc miền QMF (ví dụ, nhờ sử dụng 10 băng con hoặc nhiều hơn).

Đơn vị mã hóa 200 còn bao gồm môđun mã hóa chung 230 (nhất là môđun SPAP), được tạo cấu hình để xác định siêu dữ liệu mã hóa chung 205 (nhất là siêu dữ liệu SPAR (Spatial Audio Resolution Reconstruction - tái tạo độ phân giải âm thanh không gian)) được tạo cấu hình để tái tạo tín hiệu đầu vào đa kênh 201 từ nhiều tín hiệu kênh giảm mức 203. Môđun mã hóa chung 230 có thể được tạo cấu hình để xác định siêu dữ liệu mã hóa chung 205 trong miền băng con. Theo một ví dụ, công cụ tái tạo âm thanh không gian (spatial audio reconstruction - SPAR) là công cụ mã hóa để cải thiện sự mã hóa số lượng kênh và đối tượng âm thanh tương đối lớn. Để tăng hiệu quả mã hóa, công cụ này hỗ trợ sự tái tạo kênh và đối tượng âm thanh từ số lượng kênh âm thanh đầu vào chung thấp hơn và thông tin phụ chi phí thấp.

Để xác định siêu dữ liệu mã hóa chung hoặc SPAR 205, nhiều tín hiệu kênh giảm mức 203 có thể được biến đổi thành miền băng con và/hoặc có thể được xử lý trong miền băng con. Hơn thế nữa, tín hiệu đầu vào đa kênh 201 có thể được biến đổi thành miền

bằng con. Sau đó, siêu dữ liệu mã hóa chung hoặc SPAR 205 có thể được xác định trên cơ sở mỗi băng con, nhất là sao cho bằng cách tăng mức tín hiệu băng con của nhiều tín hiệu kênh giảm mức 203 nhờ sử dụng siêu dữ liệu mã hóa chung hoặc SPAR 205, thu được dạng gần đúng của tín hiệu băng con của tín hiệu đầu vào đa kênh 201. Siêu dữ liệu mã hóa chung hoặc SPAR 205 cho các băng con khác nhau có thể được chèn vào dòng bit 101 để truyền đến đơn vị giải mã 120 tương ứng.

Ngoài ra, đơn vị mã hóa 200 có thể bao gồm môđun mã hóa 240 được tạo cấu hình để thực hiện việc mã hóa dạng sóng nhiều tín hiệu kênh giảm mức 203, nhờ đó tạo ra dữ liệu âm thanh mã hóa 206. Mỗi tín hiệu trong số các tín hiệu kênh giảm mức 203 có thể được mã hóa nhờ sử dụng bộ mã hóa dạng sóng đơn (ví dụ, mã hóa 3GPP EVS), nhờ đó cho phép việc mã hóa hiệu quả. Các ví dụ khác để mã hóa nhiều tín hiệu kênh giảm mức 203 là MPEG AAC, MPEG HE-AAC và các bộ mã hóa giải mã âm thanh MPEG khác, các bộ mã hóa giải mã 3GPP, Dolby Digital / Dolby Digital Plus (AC-3, eAC-3), Opus, LC-3 và các bộ mã hóa-giải mã tương tự khác. Theo ví dụ khác, các công cụ mã hóa có trong bộ mã hóa-giải mã AC-4 có thể được tạo cấu hình để thực hiện các hoạt động của đơn vị mã hóa 200.

Hơn thế nữa, môđun mã hóa 240 có thể được tạo cấu hình để thực hiện việc mã hóa entropy siêu dữ liệu mã hóa chung (tức là siêu dữ liệu SPAR) 205 và mã hóa siêu dữ liệu đối tượng 202, nhờ đó tạo ra siêu dữ liệu mã hóa 207. Dữ liệu âm thanh mã hóa 206 và siêu dữ liệu mã hóa 207 có thể được chèn vào dòng bit 101. Dòng bit 101 có thể biểu thị cấu trúc siêu khung mà được mô tả theo sáng chế. Phương pháp 500 được mô tả theo sáng chế có thể được thực hiện bởi môđun mã hóa 240.

Fig.3 thể hiện đơn vị giải mã làm ví dụ 120, 350. Đơn vị giải mã 120, 350 có thể bao gồm bộ thu để nhận dòng bit 101 mà có thể bao gồm dữ liệu âm thanh mã hóa 206 và siêu dữ liệu mã hóa 207. Đơn vị giải mã 120, 350 có thể bao gồm bộ xử lý và/hoặc bộ giải ghép kênh để giải ghép kênh dữ liệu âm thanh mã hóa 206 và siêu dữ liệu mã hóa 207 từ dòng bit 101. Đơn vị giải mã 350 bao gồm môđun giải mã 360 được tạo cấu hình để thu được nhiều tín hiệu kênh tái tạo 314 từ dữ liệu âm thanh mã hóa 206. Môđun giải mã 360 có thể còn được tạo cấu hình để thu được siêu dữ liệu mã hóa chung hoặc SPAR 205 và/hoặc siêu dữ liệu đối tượng 202 từ siêu dữ liệu mã hóa 207. Phương pháp 600 được mô tả theo sáng chế có thể được thực hiện bởi môđun giải mã 360.

Ngoài ra, đơn vị giải mã 350 bao gồm môđun tái tạo 370 được tạo cấu hình để thu được tín hiệu đa kênh tái tạo 311 từ siêu dữ liệu mã hóa chung hoặc SPAR 205 và từ nhiều tín hiệu kênh tái tạo 314. Siêu dữ liệu mã hóa chung hoặc siêu dữ liệu SPAR 205 có thể truyền các phần tử thay đổi theo thời gian và/hoặc tần số của ma trận tăng mức để cho phép tái tạo tín hiệu đa kênh 311 từ nhiều tín hiệu kênh tái tạo 314. Quy trình tăng mức có thể được tiến hành trong miền băng con QMF (Quadrature Mirror Filter - bộ lọc cầu phương). Theo cách khác, một biến đổi thời gian/tần số khác, nhất là biến đổi dựa trên FFT (Fast Fourier Transform - biến đổi Fourier nhanh), có thể được sử dụng để thực hiện quy trình tăng mức. Nhìn chung, phép biến đổi có thể được áp dụng, để cho phép phân tích chọn lọc tần số và việc xử lý (tăng mức). Quy trình tăng mức cũng có thể bao gồm các bộ giải tương quan cho phép cải thiện sự tái tạo hiệp phương sai của tín hiệu đa kênh tái tạo 311, trong đó các bộ giải tương quan có thể được điều khiển bởi siêu dữ liệu mã hóa chung hoặc siêu dữ liệu SPAR 205 bổ sung.

Tín hiệu đa kênh tái tạo 311 có thể bao gồm tín hiệu SR tái tạo và một hoặc nhiều tín hiệu đối tượng tái tạo. Tín hiệu đa kênh tái tạo 311 và siêu dữ liệu đối tượng có thể tạo ra tín hiệu đầu ra 121 (còn được gọi là tín hiệu IA tái tạo 121). Tín hiệu IA tái tạo 121 có thể được sử dụng để kết xuất ra loa 331, để kết xuất ra tai nghe 332 và/hoặc để kết xuất, ví dụ, nội dung VR dựa vào biểu diễn SR 333

Do vậy, đơn vị mã hóa 110, 200 được mô tả được tạo cấu hình để mã hóa tín hiệu đầu vào IA 111 thành

- tín hiệu giảm mức chứa nhiều tín hiệu kênh giảm mức 203, trong đó tín hiệu giảm mức 203 có thể là tín hiệu biểu diễn trường âm thanh (SR); và
- siêu dữ liệu 202, 205 bao gồm SPAR hoặc siêu dữ liệu mã hóa chung 205 và/hoặc siêu dữ liệu đối tượng 202 cho một hoặc nhiều đối tượng.

Siêu dữ liệu 202, 205, nhất là siêu dữ liệu SPAR 205, có thể biểu thị độ phân giải thời gian khác với tín hiệu giảm mức. Cụ thể, siêu dữ liệu 202, 205 có thể được sử dụng cho nhiều khung (ví dụ, cho hai khung) của tín hiệu giảm mức. Vì vậy, siêu khung có thể được xác định cho dòng bit 101, trong đó siêu khung bao gồm nhiều khung của tín hiệu giảm mức cộng với siêu dữ liệu 202, 205 cho nhiều khung của tín hiệu giảm mức SR.

Fig.4 thể hiện siêu khung làm ví dụ 400. Siêu khung 400 có thể bao gồm trường phần đầu gốc (base header - BH) 401 và/hoặc trường thông tin cấu hình (configuration information - CI) 402, có thể bao gồm dữ liệu hợp lệ cho toàn bộ siêu khung 400. Hơn thế nữa, siêu khung 400 bao gồm các trường dữ liệu tín hiệu 411, 412, 421, 422 cho dữ liệu âm thanh mã hóa 206 cho một hoặc nhiều (nhất là cho nhiều) khung của tín hiệu giảm mức. Cụ thể, đối với mỗi tín hiệu kênh giảm mức 203 một hoặc nhiều (nhất là nhiều) trường dữ liệu tín hiệu 411, 412, 421, 422 có thể được tạo ra, ví dụ, các trường dữ liệu tín hiệu 411, 421 cho hai khung của tín hiệu kênh giảm mức 203 thứ nhất và các trường dữ liệu tín hiệu 412, 422 cho hai khung của tín hiệu kênh giảm mức 203 thứ N. Các trường dữ liệu tín hiệu 411, 412, 421, 422 còn được gọi ở đây là trường bit EVS (chẳng hạn bộ mã hóa EVS được sử dụng để mã hóa các tín hiệu kênh giảm mức 203).

Hơn thế nữa, siêu khung 400 bao gồm trường siêu dữ liệu (metadata field - MDF) 403. Trường siêu dữ liệu 403 có thể được tạo cấu hình để tạo ra siêu dữ liệu SPAR hoặc siêu dữ liệu mã hóa chung 205 và/hoặc các hệ số dự đoán (predictive coefficient - PC). Do vậy, trường siêu dữ liệu 403 có thể là trường bit SPAR hoặc trường bit PC (tùy thuộc chế độ mã hóa đang được sử dụng). Ngoài ra, siêu khung 400 có thể bao gồm trường bộ mở rộng khung (frame extender - FE) 404.

Do vậy, siêu khung 400 có thể bao gồm các phần tử báo hiệu được tạo cấu hình để

- chỉ báo một hoặc nhiều chế độ mã hóa-giải mã (EVS) đã được sử dụng để mã hóa (EVS) N tín hiệu giảm mức; mặc định có thể là $N = 4$ kênh, nghĩa là có 4 tín hiệu kênh giảm mức của bộ mã hóa-giải mã (EVS) W, X', Y', Z'.
- chỉ báo chế độ hoạt động được chọn của bộ mã hóa- giải mã (EVS) được hỗ trợ bởi siêu dữ liệu.
- chỉ báo tốc độ bit của siêu dữ liệu.
- cung cấp khả năng báo hiệu các phần mở rộng tiềm năng trong tương lai.

Một hoặc nhiều phần tử báo hiệu (như trường CI 402) có thể chỉ được cung cấp có điều kiện trong bảng trong siêu khung 400. Nếu phần tử báo hiệu tùy ý hoặc có điều kiện được cung cấp, phần tử báo hiệu này có thể được làm thích ứng động và/hoặc được bao gồm trong siêu khung 400. Một hoặc nhiều phần tử báo hiệu có thể được giữ tĩnh và/hoặc

có thể được tạo ra chỉ một lần, chẳng hạn, dưới dạng thông báo ngoài băng. Một hoặc nhiều phần tử báo hiệu có thể là bán động, trong trường hợp đó một hoặc nhiều phần tử báo hiệu được cung cấp trong băng chỉ trong siêu khung được chọn 400.

Siêu khung 400 có thể được thiết kế để cho phép một hoặc nhiều dấu hiệu sau:

- Giải mã và kết xuất hoàn toàn các siêu khung được mã hóa EVS được hỗ trợ bởi siêu dữ liệu.
- Giải mã đơn âm một phần các siêu khung được mã hóa EVS được hỗ trợ bởi siêu dữ liệu.
- Việc trích xuất có độ phức tạp thấp thông tin kích thước siêu khung từ chuỗi siêu khung được ghép nối, mà không cần giải mã các siêu khung, ví dụ, để đặt thông tin kích thước siêu khung vào định dạng thứ cấp (như định dạng tập tin phương tiện gốc ISO, ISOBMFF) cung cấp hoặc yêu cầu thông tin kích thước siêu khung này.
- Việc xác định tốc độ bit có độ phức tạp thấp không cần giải mã siêu khung dữ liệu.
- Sự tiếp thuận ít phức tạp và bỏ qua các siêu khung mà không cần giải mã siêu khung dữ liệu.
- Sự tiếp nghịch ít phức tạp mà không cần giải mã siêu khung dữ liệu (nhất là trong trường hợp hoạt động ở tốc độ bit không đổi).
- Sự tái đồng bộ hóa đơn giản và bỏ qua siêu khung trong trường hợp lỗi bit ở các phần số học và/hoặc EVS được mã hóa entropy và/hoặc dòng bit siêu dữ liệu.
- Siêu khung hiệu chỉnh được, cho phép thay thế khung siêu dữ liệu hoặc khung dữ liệu EVS.

Siêu khung bit mã hóa 400 của bộ mã hóa-giải mã EVS hỗ trợ bởi siêu dữ liệu có thể tương ứng với bước mã hóa là 40ms (ví dụ, chứa hai khung 20ms). Có thể bao gồm các trường bit cơ bản sau:

- Trường phần đầu gốc (BH) 401: Trường này có thể mang chỉ báo sự hiện diện của trường cấu hình (CPI), chỉ báo sự điều chỉnh kích thước trường siêu dữ liệu (MDA) và chỉ báo phần mở rộng (EI). CPI có thể chỉ báo xem trường

thông tin cấu hình (Configuration Information - CI) có được cung cấp trong siêu khung hiện thời 400 không. MDA có thể báo hiệu sự chênh lệch giữa kích thước khung siêu dữ liệu tối đa được báo hiệu và kích thước khung siêu dữ liệu thực tế. EI có thể báo hiệu xem siêu khung 400 có được mở rộng bởi bộ mở rộng khung (Frame Extender - FE) 404 không.

- Trường thông tin cấu hình (CI) 402: Trường này có thể mang thông tin báo hiệu liên quan đến các cấu hình của các công cụ mã hóa EVS, SPAR và hệ số dự đoán, mà đã được sử dụng, như kiểu khung (chế độ mã hóa), tốc độ bit và các thông số cấu hình khác được mô tả trong sáng chế.
- Các trường bit EVS 411, 421, 412, 422: Mỗi trường có thể mang các bit của một khung EVS đơn nhất (nhất là không có phần đầu tải tin EVS), như được quy định trong 3GPP TS 26.445: “Codec for Enhanced Voice Services (EVS); Detailed algorithmic description”, mục 7, mà được đưa vào đây bằng cách viện dẫn.
- Trường bit SPAR (SPAR) 403: Trường này có thể mang các bit thuộc một khung siêu dữ liệu SPAR đơn nhất, có thể được đệm zero ở cuối để được đồng chỉnh byte.
- Trường bit hệ số dự đoán (PC) 403: Trường này có thể mang các bit thuộc một khung siêu dữ liệu hệ số dự đoán đơn nhất, có thể được đệm zero ở cuối để được đồng chỉnh byte.
- Bộ mở rộng khung (FE) 404: Trường này có thể được xác định để sử dụng sau này và có thể mang dữ liệu mở rộng. Ngoại trừ phần tử kích thước chứa trong FE, dữ liệu khác bất kỳ mang bởi FE có thể được dự trữ để sử dụng sau này (RFU).

Tất cả các trường bit cơ bản có thể được đồng chỉnh byte và – nếu cần – được đệm zero ở đầu cuối cho đến kích thước xác định của chúng.

Các trường cơ bản nêu trên có thể được bao gồm trong (một) siêu khung 400 theo thứ tự chuỗi sau. Siêu khung có thể bao gồm

- Một phần đầu gốc (BH) 401 chứa
 - Chỉ báo sự có mặt trường cấu hình (CPI),

- Chỉ báo sự điều chỉnh kích thước trường siêu dữ liệu (MDA)
- và
- Chỉ báo mở rộng (EI).
- Một trường thông tin cấu hình (CI) 402 tùy ý. Sự có mặt của trường CI 402 có thể được báo hiệu bởi CPI.
- Dữ liệu gồm N tín hiệu kênh giảm mức được mã hóa EVS $S_1, \dots, S_N$, hai khung liên tiếp đối với mỗi tín hiệu kênh giảm mức, mà có thể được mang bởi $2 * N$ trường bit EVS cơ bản 411, 421, 412, 422 (còn được gọi ở đây là EVS(.)). Trong hoạt động mặc định với 4 tín hiệu kênh giảm mức, có 8 trường bit EVS liên tiếp 411, 421, 412, 422 biểu diễn hai khung của các tín hiệu kênh giảm mức W, X', Y', Z'.
- Một trường khung siêu dữ liệu (MDF) 403, đối với SPAR hoặc các hệ số dự đoán, do đó là
 - Một trường bit SPAR cơ bản, hoặc
 - Một trường bit PC cơ bản.
- Một bộ mở rộng khung (FE) 404 tùy ý. Sự có mặt của trường FE có thể được chỉ báo bởi EI.

Bảng 1 chỉ báo cấu trúc ví dụ của siêu khung 400.

Bit(MSB-L SB)	Tên	Mô tả
8	BH	Phần đầu gốc, chứa CPI, MDA, EI
Biến	CI	Trường thông tin cấu hình (tùy ý, tùy thuộc CPI)
Biến	EVS ($S_{1,1}$)	Dữ liệu khung EVS cho khung thứ nhất của kênh dmx thứ nhất
Biến	EVS ($S_{2,1}$)	Dữ liệu khung EVS cho khung thứ nhất của kênh dmx thứ hai
Biến	EVS ($S_{...,1}$)	...
Biến	EVS ($S_{N,1}$)	Dữ liệu khung EVS cho khung thứ nhất của kênh dmx thứ N

Biến	EVS ($S_{1,2}$)	Dữ liệu khung EVS cho khung thứ hai của kênh dmx thứ nhất
Biến	EVS ($S_{2,2}$)	Dữ liệu khung EVS cho khung thứ hai của kênh dmx thứ hai
Biến	EVS ($S_{...,2}$)	...
Biến	EVS ($S_{N,2}$)	Dữ liệu khung EVS cho khung thứ hai của kênh dmx thứ N
Biến	MDF	Khung siêu dữ liệu chứa trường SPAR hoặc PC
Biến	FE	Bộ mở rộng khung (tùy ý, tùy thuộc EI)

Bảng 1

Ở trạng thái mặc định, có 4 tín hiệu kênh giảm mức được mã hóa EVS. Cấu trúc siêu khung cho trạng thái mặc định được thể hiện trong 2.

Bit(MSB-L SB)	Tên	Mô tả
8	BH	Phần đầu gốc, chứa CPI, MDA, EI
72	CI	Trường thông tin cấu hình (tùy ý, tùy thuộc CPI)
Biến	EVS ($W(1)$)	Dữ liệu khung EVS cho khung thứ nhất của tín hiệu W
Biến	EVS ($X'(1)$)	Dữ liệu khung EVS cho khung thứ nhất của tín hiệu X'
Biến	EVS ($Y'(1)$)	Dữ liệu khung EVS cho khung thứ nhất của tín hiệu Y'
Biến	EVS ($Z'(1)$)	Dữ liệu khung EVS cho khung thứ nhất của tín hiệu Z'
Biến	EVS($W(2)$)	Dữ liệu khung EVS cho khung thứ hai của tín hiệu W
Biến	EVS ($X'(2)$)	Dữ liệu khung EVS cho khung thứ hai của tín hiệu X'
Biến	EVS ($Y'(2)$)	Dữ liệu khung EVS cho khung thứ hai của tín hiệu Y'

Biến	EVS ($Z'(2)$)	Dữ liệu khung EVS cho khung thứ hai của tín hiệu Z'
Biến	MDF	Khung siêu dữ liệu chứa trường SPAR hoặc PC
Biến	FE	Bộ mở rộng khung (tùy ý, tùy thuộc EI)

Bảng 2

Chi tiết khác liên quan đến các trường bit cơ bản khác nhau được đưa ra dưới đây.

Trường phần đầu gốc (BH) 401 có thể mang chỉ báo sự hiện diện của trường cấu hình (CPI), chỉ báo sự điều chỉnh kích thước trường siêu dữ liệu (MDA) và chỉ báo phần mở rộng (EI). Trường byte này có thể luôn là phần tử thứ nhất trong siêu khung 400.

Cấu trúc của trường BH 401 được thể hiện trong bảng 3.

Bit(MSB-LSB)	Tên	Mô tả
1	CPI	Chỉ báo sự có mặt trường cấu hình
6	MDA	Điều chỉnh kích thước trường siêu dữ liệu
1	EI	Chỉ báo mở rộng

Bảng 3

Chỉ báo sự có mặt trường cấu hình (CPI) có thể là bit đơn được sử dụng để báo hiệu sự có mặt của trường thông tin cấu hình (Configuration Information - CI) trong siêu khung hiện thời 400. CPI có thể có nghĩa sau:

- CPI = '0': Điều này chỉ báo rằng trường thông tin cấu hình không được cung cấp trong siêu khung hiện thời 400. Cần lưu ý rằng thông tin cấu hình thay vào đó có thể được cung cấp dưới dạng thông tin tĩnh ngoài băng hoặc từ khung con 400 nhận được trước đó gần nhất mang trường thông tin cấu hình 402.
- CPI = '1': Điều này chỉ báo rằng trường thông tin cấu hình được cung cấp trong siêu khung hiện thời 400. Thông tin cấu hình được cung cấp trong trường CI 402 là hợp lệ cho siêu khung 400 này và cho siêu khung 400 tương lai bất kỳ cho đến siêu khung 400 tiếp theo mang trường thông tin cấu hình 402 được cung cấp.

Chỉ báo điều chỉnh kích thước trường siêu dữ liệu (MDA) có thể được tạo ra trực tiếp sau bit CPI. Chỉ báo 6-bit này có thể báo hiệu sự chênh lệch giữa độ dài của MDF 403 như được báo hiệu bởi phần tử MDR (được xác định thêm bên dưới) và kích thước thực của MDF 403. Nhờ sử dụng MDA làm chỉ số, chênh lệch được chỉ báo có thể được thu được từ bảng tìm kiếm thể hiện trong bảng 4. Dãy giá trị điều chỉnh trong bảng 4 được xác định theo kiểu Matlab: start-value:step-size:end-value. Các kích thước bước tham số điều chỉnh thay đổi thể hiện trong bảng 4 có thể được thiết kế theo mô hình gần đúng của sự phân bố tổng chiều dài mã entropy của siêu dữ liệu. Điều này cho phép giảm đến mức thấp nhất số lượng bit không được sử dụng trong MDF 403 và do đó giảm chi phí truyền.

MDA	0...47	48...55	56...59	60...61	62	63
Giá trị điều chỉnh	0:1:47	49:2:63	67:4:79	87:8:95	111	143

Bảng 4

Tùy thuộc vào kích thước MDF tối đa, giá trị điều chỉnh biểu thị các đơn vị một byte hoặc hai byte. Đối với các kích thước MDF tối đa lên đến 275 byte, giá trị điều chỉnh biểu thị các đơn vị một byte, nếu không thì biểu thị các đơn vị hai byte.

Chỉ báo MDA có thể được theo sau bởi một bit chỉ báo mở rộng bit (EI). Nếu bit này được thiết lập là 1, siêu khung hiện thời 400 được gắn kèm theo phần tử bộ mở rộng khung (FE).

Trường thông tin cấu hình (CI) 402 được cung cấp tùy ý có thể mang các phần tử báo hiệu sau như được minh họa trong bảng 5. Trường CI 402 có thể gồm có hoặc có thể bao gồm 8 byte dữ liệu (đối với trường hợp hai khung EVS trên một tín hiệu kênh giảm mức và N=4 kênh giảm mức).

Bit(MSB-LS B)	Tên	Mô tả
3	N-I	Chỉ báo về số lượng kênh giảm mức của bộ mã hóa-giải mã EVS
1	MDT	Chỉ báo loại siêu dữ liệu
11	MDC	Cấu hình mã hóa siêu dữ liệu

5	MDR	Báo hiệu tốc độ siêu dữ liệu
3	BND	Số lượng băng mã hóa siêu dữ liệu
1	RES	Được dự trữ để sử dụng sau này
6	FT-1,1	EVS FT cho khung thứ nhất của kênh dmx thứ nhất
6	FT-2,1	EVS FT cho khung thứ nhất của kênh dmx thứ hai
6	...	...
6	FT-N,1	EVS FT cho khung thứ nhất của kênh dmx thứ N
6	FT-1,2	EVS FT cho khung thứ hai của kênh dmx thứ nhất
6	FT-2,2	EVS FT cho khung thứ hai của kênh dmx thứ hai
6	...	...
6	FT-N,2	EVS FT cho khung thứ hai của kênh dmx thứ N
Biến	Đệm zero	Đệm zero để làm đầy byte

Bảng 5

Bảng 6 minh họa trường thông tin cấu hình 402 tùy ý cho trường hợp mặc định với 4 tín hiệu kênh giảm mức được mã hóa EVS. Trong trường hợp này trường CI gồm có 9 byte dữ liệu.

Bit(MSB-L SB)	Tên	Mô tả
3	N-I	'011' biểu thị 4 kênh giảm mức của bộ mã hóa-giải mã EVS
1	MDT	Chỉ báo loại siêu dữ liệu
11	MDC	Cấu hình mã hóa siêu dữ liệu
5	MDR	Báo hiệu tốc độ bit siêu dữ liệu
3	BND	Số lượng băng mã hóa siêu dữ liệu
1	RES	Được dự trữ để sử dụng sau này
6	FT-1,1	EVS FT để mã hóa W(1)
6	FT-2,1	EVS FT để mã hóa X'(1)
6	FT-3,1	EVS FT để mã hóa Y'(1)
6	FT-4,1	EVS FT để mã hóa Z'(1)

6	FT-1,2	EVS FT để mã hóa W(2)
6	FT-2,2	EVS FT để mã hóa X'(2)
6	FT-3,2	EVS FT để mã hóa Y'(2)
6	FT-4,2	EVS FT để mã hóa Z'(2)

Bảng 6

Chỉ báo cho số lượng N của tín hiệu kênh giảm mức được mã hóa EVS (N-I) có thể là phần tử 3 bit để mã hóa số lượng N của tín hiệu kênh giảm mức được mã hóa EVS. N thu được từ chỉ báo N-I bằng cách gia tăng số lượng được biểu diễn bởi phần tử 3-bit thêm 1. Để đạt được hoạt động mặc định với 4 tín hiệu kênh giảm mức EVS, phần tử N-I có thể được thiết lập là 3 ('011').

Bit chỉ báo loại siêu dữ liệu (MDT) có thể có nghĩa sau:

- MDT = '0': chỉ báo rằng MDF mang trường bit PC.
- MDT = '1': chỉ báo rằng MDF mang trường bit SPAR.

Trường cấu hình mã hóa siêu dữ liệu (MDC) có thể bao gồm thông tin cấu hình của công cụ hệ số dự đoán được sử dụng hoặc của công cụ mã hóa SPAR, tùy thuộc chỉ báo của bit MDT. Trường MDC có thể là phần tử 11-bit của trường CI 402. Nghĩa của các bit của trường MDC có thể phụ thuộc vào bit MDT của trường CI 402. Tùy thuộc vào giá trị của bit MDT, các bit MDC có thể có nghĩa sau:

- MDT = '0': Nếu bit MDT là zero, thì 3 MSB của MDC mã hóa tham số cấu hình của sơ đồ mã hóa hệ số dự đoán. 8 bit còn lại của MDC không được sử dụng và được đệm zero. Cấu trúc và nội dung của trường MDC trong trường hợp này được thể hiện trong 7a.
- MDT = '1': Nếu bit MDT bằng 1, thì 11 bit MDC mã hóa cấu hình của bộ mã hóa-giải mã SPAR như minh họa trong bảng 7b. Bậc HOA có thể được tính bằng cách gia tăng `hoa_order_idx` thêm 1.

Bit(MSB-L SB)	Tên	Mô tả	Khoảng giá trị
3	<code>coding_strategy_idx</code>	Chỉ số vào bảng Huffman	0 – 7

8	Đệm zero
---	----------

Bảng 7a

Bit(MSB-LS B)	Tên	Mô tả	Khoảng giá trị
4	nobj	Số lượng đối tượng âm thanh	1 – 15
2	hoa_order_idx	Mã định danh cho bậc HOA	0 - 2
2	ndeco	Số lượng bộ giải tương quan	0 - 3
3	coding_strategy_idx	Chỉ số vào bảng Huffman	0 - 7

Bảng 7a

Trường báo hiệu tốc độ bit siêu dữ liệu (MDR) có thể bao gồm 5 bit và có thể được sử dụng để mã hóa kích thước tối đa của MDF. Kích thước tối đa của MDF có thể thu được bằng cách tra bảng nhờ sử dụng bảng 8, trong đó giá trị MDR là chỉ số của bảng 8. Hơn thế nữa, bảng 8 chỉ báo tốc độ bit siêu dữ liệu (tối đa) theo kbps. Trong bảng 8, kích thước MDF thực tế được báo hiệu dưới dạng kích thước MDF tối đa trừ đi số/giá trị điều chỉnh chỉ báo bởi MDA (từ trường BH 401). Điều này cho phép báo hiệu kích thước MDF thực tế với độ phân giải chi tiết (thường với độ phân giải byte). Cũng cần lưu ý rằng bất kỳ bit nào không được sử dụng trong MDF có thể được đệm zero, điều này có thể xảy ra trong trường hợp kích thước MDF thực tế cung cấp nhiều không gian hơn cần thiết cho siêu dữ liệu mã hóa.

Giá trị MDR	Tốc độ bit siêu dữ liệu (kbps)	Kích thước MDF tối đa (byte)
0	4	20
1	5	25
2	6	30
3	7	35
4	8	40
5	10	50
6	12	60
7	15	75
8	18	90

9	21	105
10	24	120
11	28	140
12	32	160
13	36	180
14	40	200
15	42	210
16	45	225
17	47	235
18	50	250
19	55	275
20	60	300
21	65	325
22	70	350
23	75	375
24	80	400
25	85	425
26	90	450
27	95	475
28	100	500
29	105	525
30	110	550
31	128	640

Bảng 8

Trường số lượng băng (Band Number field -BND) có thể là số 3 bit và có thể chỉ báo số lượng băng con được sử dụng trong việc mã hóa siêu dữ liệu. Số lượng băng thu được từ giá trị BND bằng cách tra cứu trong bảng 9. Trong hoạt động mặc định, trường BND có thể được thiết lập là 5 ('101'), chỉ báo 12 băng con.

BND (3 bit)	Số lượng băng
0	1

1	3
2	5
3	7
4	9
5	12
6	15
7	23

Bảng 9

Việc sử dụng bit dự trữ (RES) có thể được dự trữ để sử dụng sau này. Trong các cấu hình mặc định, bit này có thể được thiết lập là '0' và có thể được bộ thu bỏ qua.

Trường EVS FT (FT-x,y) có thể biểu diễn kiểu khung EVS (FT) được áp dụng để mã hóa khung thứ y của tín hiệu kênh giảm mức thứ x, trong đó $x=1\dots N$ và trong đó $y=1,2$. Kiểu khung EVS có thể là như được định nghĩa trong 3GPP TS 26.445, mục A2.2.1.2, được đưa vào đây bằng cách viện dẫn. Cần lưu ý rằng trường EVS FT cuối trong trường CI 402 có thể được theo sau bởi tối đa 7 bit đệm zero, mà bảo đảm sự đồng chỉnh octet. Trong trường hợp trường EVS FT cuối kết thúc mà đã được đồng chỉnh octet, thì không có bit đệm zero nào được gắn thêm. Các bit đệm zero sẽ được bộ thu bỏ qua.

Trường bit EVS cơ bản 411, 421, 412, 422 có thể được xác định như được quy định trong 3GPP TS 26.445, mục 7, (được đưa vào đây bằng cách viện dẫn) cho chế độ mã hóa EVS được sử dụng tương ứng. Như được chỉ ra trong tài liệu viện dẫn, không có bit báo hiệu bổ sung nào được xác định là một phần của trường khung EVS cơ bản để chỉ báo tốc độ bit hoặc chế độ hoạt động EVS. Thông tin này có thể là một phần của trường CI 402 tùy ý của siêu khung hiện thời hoặc của siêu khung 400 trước đó hoặc cũng có thể được cung cấp ngoài băng.

Sự phân bổ chi tiết các hệ số cho siêu dữ liệu SPAR được thể hiện trong bảng 10. Bảng 10 thể hiện thứ tự của các bit khi chúng được chèn trong khung. Lưu ý rằng bit có ý nghĩa nhất (most significant bit - MSB) của mỗi tham số luôn được chèn đầu tiên. Vì mỗi trường được lượng tử hóa động, nên sự phân bổ bit là thay đổi.

Tên	Mô tả	Bước mã hóa
obj_pos	Ma trận vị trí đối tượng	Lượng tử hóa phụ thuộc vị trí
Mspar	Ma trận tái tạo HOA SPAR	Được mã hóa Huffman
Pspar	Ma trận SPAR	Được mã hóa Huffman

Bảng 10

Sự phân bố chi tiết các hệ số cho siêu dữ liệu PC được thể hiện trong bảng 11. Bảng 11 thể hiện thứ tự của các bit khi chúng được chèn trong siêu khung 400. Lưu ý rằng bit có ý nghĩa nhất (most significant bit - MSB) của mỗi tham số luôn được chèn đầu tiên. Vì mỗi trường được lượng tử hóa động, nên sự phân bố bit là thay đổi.

Tên	Mô tả	Bước mã hóa
Gfoa	Ma trận các hệ số dự đoán	Được mã hóa Huffman

Bảng 11

Phần tử bộ mở rộng khung (FE) 404 thường mang trong hai byte đầu tiên của nó số nguyên không dấu 16-bit để chỉ báo kích thước của trường FE 404 ở byte. Phần tử này được gọi là FE-kích thước. Số lượng FE-kích thước do đó lớn hơn hoặc bằng 2. Nội dung và nghĩa của phần FE-dữ liệu còn lại của trường FE 404 có thể được dự trữ để sử dụng sau này. Trong hoạt động mặc định, phần tử FE-kích thước có thể được phân tích và phần tử FE-dữ liệu có thể được nhảy qua và bỏ qua. Cấu trúc và nội dung của trường FE 404 được thể hiện trong bảng 12.

Bit(MSB-LSB)	Tên	Mô tả
16	FE-kích thước	Kích thước của trường FE
biến	FE-dữ liệu	các bit dữ liệu mang trong trường FE

Bảng 12

Do vậy, cấu trúc siêu khung được mô tả, cho phép báo hiệu thông tin cấu hình của bộ mã hóa-giải mã EVS hỗ trợ bởi siêu dữ liệu. Cấu trúc siêu khung cho phép bộ thu giải mã dữ liệu của bộ mã hóa-giải mã EVS hỗ trợ bởi siêu dữ liệu.

Ở mức tổng quát, bộ mã hóa-giải mã EVS được hỗ trợ bởi siêu dữ liệu là hệ thống mã hóa đa chế độ và/hoặc đa tốc độ. Bộ mã hóa-giải mã EVS cơ sở có thể được tạo cấu hình để hoạt động ở vô số chế độ mã hóa và/hoặc tốc độ bit khác nhau. Hơn thế nữa, bộ

mã hóa-giải mã siêu dữ liệu không gian có thể đưa ra nhiều chế độ mã hóa và/hoặc tốc độ bit khác nhau. Bộ mã hóa-giải mã siêu dữ liệu không gian sử dụng mã hóa entropy mà thường dẫn đến tốc độ bit thay đổi. Điều này nghĩa là tốc độ bit được sử dụng thực tế thường thấp hơn tốc độ bit đích đã cho. Đối với một số khung nhất định, mức dưới chuẩn của tốc độ bit này có thể nhỏ hơn và đối với một số khung khác, nó có thể lớn hơn.

Để bộ giải mã 120 có thể giải mã dòng bit 101 đã truyền một cách chính xác, chế độ mã hóa và tốc độ bit chính xác được sử dụng bởi bộ mã hóa 110 có thể được đưa ra. Đối với phân mã hóa entropy của siêu dữ liệu, có thể không yêu cầu tốc độ bit được sử dụng chính xác, vì các mã Huffman được sử dụng không có dấu phẩy và có thể được giải mã theo cách độc nhất. Tuy nhiên, bộ thu của dòng bit 101 có thể được cung cấp số lượng bit được sử dụng để mã hóa khung (hoặc siêu khung 400). Điều này là ví dụ mong muốn nếu bộ giải mã 120 cần nhảy qua một số khung nhận được mà không phải giải mã các khung này. Theo sáng chế, cấu trúc siêu khung đã được mô tả là hỗ trợ các đặc tính sau:

- Giải mã toàn bộ khung.
- Chỉ giải mã phần cần thiết để tái tạo đơn âm.
- Trích xuất thông tin chiều dài từ khung 400, để chuyển thông tin này thành định dạng thứ cấp (ISOBMFF) mà cung cấp và/hoặc cần thông tin này.
- Trong dãy ghép khung 400, giải mã chỉ một phần trong dãy khung, và bỏ qua một vài khung đầu một cách hiệu quả.
- Nếu có lỗi bit, tìm vị trí bắt đầu khung tiếp theo (tái đồng bộ hóa).
- Xác định tốc độ bit một cách nhanh chóng và hiệu quả, mà không phải giải mã khung.
- Hiệu chỉnh khung (thay thế siêu dữ liệu, hoặc các phần của các khung EVS)).
- Hoạt động tiếp thuận nhanh mà không giải mã khung.

- Hỗ trợ việc vận chuyển hiệu quả các tải tin dữ liệu của bộ mã hóa-giải mã có độ dài không đổi và biến thiên.

Ngoài ra, cấu trúc siêu khung được mô tả là chứa tất cả các phần tử báo hiệu cần thiết để

- chỉ báo các chế độ của bộ mã hóa-giải mã EVS (bao gồm các tốc độ bit) được sử dụng để mã hóa EVS N tín hiệu kênh giảm mức. Chế độ mặc định có thể là $N = 4$, nghĩa là có 4 kênh giảm mức của bộ mã hóa-giải mã EVS W, X', Y', Z'.
- chỉ báo chế độ hoạt động được chọn của bộ mã hóa-giải mã siêu dữ liệu hỗ trợ.
- chỉ báo tốc độ bit của siêu dữ liệu ở độ phân giải cao mà không có chi phí bit báo hiệu đáng kể.
- cung cấp khả năng báo hiệu các phần mở rộng tiềm năng trong tương lai.

Một số trong số các phần tử báo hiệu của siêu khung 400 có thể không thường xuyên thay đổi trong suốt phiên mã hóa hoặc thậm chí là tĩnh. Một số các phần tử báo hiệu khác như tốc độ bit của siêu dữ liệu có thể thay đổi từ siêu khung này đến siêu khung khác. Vì lý do đó, một số phần tử báo hiệu chỉ được cung cấp có điều kiện trong băng trong siêu khung 400 (như trường CI 402). Nếu chúng được cung cấp, các phần tử báo hiệu này có thể được làm thích ứng động trên cơ sở siêu khung. Cũng có khả năng giữ các phần tử báo hiệu này tĩnh và cung cấp chúng chỉ một lần, chẳng hạn, dưới dạng bản tin ngoài băng. Các phần tử báo hiệu cũng có thể là bán động, trong trường hợp đó chúng được cung cấp trong băng chỉ trong một số siêu khung nhất định.

Liên quan đến sự báo hiệu tốc độ bit siêu dữ liệu, thách thức chính là ở chỗ số lượng bit (hoặc byte) cần thiết trên mỗi siêu khung 400 có thể thay đổi trong khoảng tương đối lớn. Việc chỉ báo hiệu số lượng bit tối đa có thể có trên mỗi khung có thể bỏ qua không sử dụng số lượng bit tương đối lớn, trong trường hợp mã entropy ngắn hơn đáng kể so với độ dài tối đa này. Mặt khác, việc cung cấp phần tử báo hiệu trực tiếp để chỉ báo số lượng bit (hoặc byte) được sử dụng thực tế trong siêu khung 400 sẽ đòi hỏi số lượng bit báo hiệu tương đối lớn. Theo sáng chế, sơ đồ được mô tả giữ số lượng bit báo

hiệu cho số lượng bit (hoặc byte) được sử dụng thực tế trong siêu khung 400 ở mức thấp nhất, trong khi vẫn cho phép bao phủ một khoảng tương đối lớn của các tốc độ bit siêu dữ liệu có thể có.

Từ góc độ hệ thống, các siêu khung 400 của bộ mã hóa-giải mã EVS hỗ trợ bởi siêu dữ liệu được tạo ra ở đầu cuối mã hóa. Đầu cuối mã hóa này có thể là máy chủ trong mạng có quyền truy cập vào dữ liệu âm thanh nổi sống động không được mã hóa hoặc dữ liệu âm thanh thực tế ảo VR (Virtual Reality). Đầu cuối mã hóa này cũng có thể là điện thoại di động thu nhận các tín hiệu âm thanh nổi sống động. Khung mã hóa 400 có thể được chèn vào tệp tin được tải về đầu cuối nhận hoặc được truyền theo giao thức tạo dòng giống DASH (Tạo dòng thích ứng động trên HTTP) hoặc RTSP/RTP (Giao thức tạo dòng theo thời gian thực /Giao thức truyền theo thời gian thực). Nếu các siêu khung 400 được mã hóa được lưu trữ trong tệp tin, thì các siêu khung 400 có thể được chèn vào tệp tin được định dạng theo ISO/BMFF. Trong trường hợp một số thông tin cấu hình là tĩnh và trong trường hợp thông tin này không được truyền dưới dạng một phần của siêu khung 400, thay vào đó thông tin này có thể được cung cấp từ đầu mã hóa đến đầu giải mã bằng phương tiện ngoài băng giống giao thức mô tả phiên (session description protocol - SDP).

Các sơ đồ được mô tả theo sáng chế có thể sử dụng bộ mã hóa-giải mã EVS làm bộ mã hóa-giải mã cơ sở và có thể cung cấp các bản tin đa chế độ/đa tốc độ (kiểu khung) trong băng trong siêu khung 400 hoặc ngoài băng nhờ sử dụng, ví dụ, SDP. Sơ đồ này có thể được kết hợp với khung sườn mã hóa siêu dữ liệu nổi sống động đa chế độ mà có thể được tạo cấu hình hiệu quả với tập tham số cấu hình mà cũng có thể được truyền trong băng hoặc ngoài băng. Hơn thế nữa, có khả năng kết hợp việc mã hóa siêu dữ liệu nổi sống động đa chế độ với sơ đồ cho phép các tốc độ bit (hoặc số lượng bit trong khung/siêu khung) tối đa kết hợp trong băng hoặc ngoài băng.

Cấu trúc siêu khung được mô tả theo sáng chế báo hiệu kích thước trường siêu dữ liệu được sử dụng thực tế là số lớn nhất (mà được báo hiệu tùy ý ngoài băng) trừ đi tham số điều chỉnh mà chỉ báo cho nó được truyền dưới dạng một phần của mỗi siêu khung 400. Việc mã hóa các tham số điều chỉnh tốt hơn là được thực hiện ở các kích thước bước thay đổi, điều này cho phép bao phủ khoảng tăng của các tham số điều chỉnh có thể có nhờ sử dụng số lượng giảm của các bit báo hiệu cho các tham số điều chỉnh này. Hơn thế nữa, các kích thước bước tham số điều chỉnh thay đổi có thể được thiết kế nhờ sử dụng

mô hình gần đúng của sự phân bố tổng độ dài mã entropy của siêu dữ liệu. Điều này cho phép giảm đến mức thấp nhất số lượng bit không được sử dụng trong trường siêu dữ liệu và do đó giảm chi phí truyền. Ngoài ra, chi phí cho tốc độ bit siêu dữ liệu (kích thước) có thể được báo hiệu, trong khi giữ số lượng bit không được sử dụng trong trường siêu dữ liệu ở mức thấp nhất. Do đó, tốc độ bit truyền tổng thể được làm giảm.

Thông tin cấu hình (CI) trong trường CI 402 có thể liên quan đến các kiểu khung EVS được chọn để mã hóa EVS bốn tín hiệu kênh giảm mức W, X', Y', Z'. Thông tin cấu hình có thể còn liên quan đến (i) chế độ hoạt động được chọn của mã EVS được hỗ trợ bởi siêu dữ liệu, FOA hoặc HIQ; (ii) tốc độ bit của siêu dữ liệu SPAR trong trường hợp hoạt động HIQ; (iii) tốc độ bit của siêu dữ liệu hệ số dự đoán trong trường hợp hoạt động FOA. Chỉ báo nếu thông tin cấu hình có thể là (1) động và được cung cấp trong bảng cùng với tải tin; (2) bán động và được cung cấp trong bảng cùng với tải tin trước đó; hoặc (3) tĩnh và được cung cấp ngoài bảng dưới dạng chuỗi thập lục phân cùng với thuộc tính mã hóa-giải mã của các tập thích ứng DASH.

Chế độ FOA (First Order Ambisonics) là chế độ hoạt động ở tốc độ bit thấp (ví dụ, hoạt động ở ~128 kbps) dựa vào siêu dữ liệu hệ số dự đoán. FOA thường biểu thị chất lượng tương đối hạn chế do độ phân giải không gian tương đối thấp. Chế độ HIQ (Chất lượng nổi sống động cao - High Immersive Quality) là chế độ hoạt động với tốc độ trung bình đến cao (ví dụ, hoạt động ở 128 đến 512 kbps). Chế độ này dựa trên siêu dữ liệu SPAR và có khả năng mang lại chất lượng nổi sống động rất cao vì nó tập trung vào việc tái tạo tín hiệu SR gốc.

Fig.5 thể hiện phương pháp 500 để tạo dòng bit 101, trong đó dòng bit 101 bao gồm chuỗi siêu khung 400 cho chuỗi các khung (cơ bản) của tín hiệu âm thanh nổi sống động 111. Tín hiệu âm thanh nổi sống động (IA) 111 có thể bao gồm tín hiệu biểu diễn trường âm thanh (SR) để có thể mô tả trường âm thanh ở vị trí tham chiếu. Vị trí tham chiếu có thể là vị trí nghe của người nghe và/hoặc vị trí thu âm thanh của micro. Tín hiệu SR có thể bao gồm nhiều kênh (hoặc dạng sóng) cho nhiều hướng tới khác nhau của trường âm thanh ở vị trí tham chiếu. Theo cách khác, hoặc ngoài ra, tín hiệu IA 111 có thể bao gồm một hoặc nhiều đối tượng âm thanh và/hoặc tín hiệu đa kênh.

Tín hiệu IA 111, nhất là tín hiệu SR được bao gồm trong tín hiệu IA, có thể bao gồm hoặc có thể là tín hiệu ambisonics bậc L, với L lớn hơn hoặc bằng 1. Theo cách khác, hoặc ngoài ra, tín hiệu SR có thể biểu thị định dạng tổ ong (BH) với nhiều hướng tới được bố trí trong nhiều vòng khác nhau trên hình cầu quanh vị trí tham chiếu. Nhiều vòng có thể bao gồm vòng giữa, vòng trên, vòng dưới và/hoặc đỉnh. Theo cách khác, hoặc ngoài ra, tín hiệu SR có thể biểu thị định dạng không gian trung gian, được gọi là ISF, nhất là định dạng ISF như được định nghĩa trong công nghệ Dolby Atmos.

Do vậy, tín hiệu IA 111 có thể bao gồm nhiều kênh khác nhau. Mỗi kênh được bao gồm trong tín hiệu IA 111 thường bao gồm chuỗi mẫu âm thanh cho chuỗi khoảng khắc thời gian hoặc cho chuỗi các khung. Nói cách khác, “các tín hiệu” được mô tả theo sáng chế thường bao gồm chuỗi mẫu âm thanh cho các chuỗi khoảng khắc hoặc khung thời gian ứng (ví dụ, ở khoảng cách theo thời gian là 20ms hoặc nhỏ hơn).

Phương pháp 500 có thể bao gồm bước trích xuất một hoặc nhiều đối tượng âm thanh từ tín hiệu IA 111. Đối tượng âm thanh thường bao gồm tín hiệu đối tượng (với chuỗi mẫu âm thanh cho chuỗi khoảng khắc hoặc khung thời gian tương ứng). Hơn thế nữa, đối tượng âm thanh thường bao gồm siêu dữ liệu đối tượng 202 biểu thị vị trí của đối tượng âm thanh. Vị trí của đối tượng âm thanh có thể thay đổi theo thời gian, sao cho siêu dữ liệu đối tượng 202 của đối tượng âm thanh có thể chỉ báo chuỗi vị trí cho chuỗi khoảng khắc hoặc khung thời gian này.

Hơn thế nữa, phương pháp 500 có thể bao gồm bước xác định tín hiệu dư dựa trên tín hiệu IA 111 và dựa trên một hoặc nhiều đối tượng âm thanh. Tín hiệu dư có thể mô tả tín hiệu IA ban đầu mà từ đó một hoặc nhiều đối tượng âm thanh 103, 303 đã được trích xuất và/hoặc loại bỏ. Tín hiệu dư có thể là tín hiệu SR được bao gồm trong tín hiệu IA 111. Theo cách khác, hoặc ngoài ra, tín hiệu dư có thể bao gồm hoặc có thể là tín hiệu âm thanh đa kênh và/hoặc tầng tín hiệu âm thanh. Theo cách khác, hoặc ngoài ra, tín hiệu dư có thể bao gồm nhiều đối tượng âm thanh ở vị trí và/hoặc định vị đối tượng cố định (ví dụ, các đối tượng âm thanh mà được gán cho các loa cụ thể của bố trí loa xác định).

Ngoài ra, phương pháp 500 có thể bao gồm bước tạo ra và/hoặc cung cấp tín hiệu giảm mức dựa trên tín hiệu IA 111 (ví dụ, nhờ sử dụng môđun giảm mức 210). Số lượng kênh của tín hiệu giảm mức thường nhỏ hơn số lượng kênh của tín hiệu IA 111. Hơn thế

nữa, phương pháp 500 có thể bao gồm bước xác định siêu dữ liệu mã hóa chung hoặc SPAR 205 cho phép việc tăng mức tín hiệu giảm mức (tức là một hoặc nhiều tín hiệu kênh giảm mức 203) thành tín hiệu đối tượng của một hoặc nhiều đối tượng âm thanh tái tạo cho một hoặc nhiều đối tượng âm thanh tương ứng. Hơn thế nữa, siêu dữ liệu mã hóa chung hoặc SPAR 205 có thể cho phép việc tăng mức tín hiệu giảm mức thành tín hiệu dư tái tạo cho tín hiệu dư tương ứng.

Tín hiệu giảm mức chứa một hoặc nhiều tín hiệu kênh giảm mức 203, siêu dữ liệu SPAR 205 và siêu dữ liệu đối tượng 202 có thể được chèn vào dòng bit 101. Cụ thể, phương pháp 500 có thể bao gồm bước thực hiện việc mã hóa dạng sóng tín hiệu giảm mức để tạo ra dữ liệu âm thanh mã hóa 206 cho chuỗi các khung của một hoặc nhiều tín hiệu kênh giảm mức 203. Việc mã hóa dạng sóng có thể được thực hiện nhờ sử dụng ví dụ, kỹ thuật mã hóa dịch vụ thoại nâng cao (Enhanced Voice Service - EVS). Hơn thế nữa, phương pháp 500 có thể bao gồm bước thực hiện việc mã hóa entropy siêu dữ liệu SPAR 205 và/hoặc siêu dữ liệu đối tượng 202 của một hoặc nhiều đối tượng âm thanh để tạo ra siêu dữ liệu (mã hóa) 207 để được chèn vào dòng bit 101.

Phương pháp 500 có thể bao gồm, theo cách lặp lại cho chuỗi siêu khung 400, bước chèn 501 dữ liệu âm thanh mã hóa 206 cho một hoặc nhiều (nhất là nhiều) khung (ví dụ, cho hai hoặc nhiều khung) của một hoặc nhiều tín hiệu kênh giảm mức 203 thu được từ tín hiệu âm thanh nổi sóng động 111, vào các trường dữ liệu 411, 421, 412, 422 của siêu khung 400. Khung (cơ sở) của tín hiệu kênh giảm mức 203 có thể trải qua 20ms của tín hiệu kênh giảm mức 203. Siêu khung 400 có thể trải qua nhiều chiều dài của khung (cơ sở), ví dụ, 40ms.

Hơn thế nữa, phương pháp 500 có thể bao gồm bước chèn 502 siêu dữ liệu 202, 205 (nhất là siêu dữ liệu mã hóa 207) để tái tạo một hoặc nhiều (nhất là nhiều) khung của tín hiệu âm thanh nổi sóng động 111 từ dữ liệu âm thanh mã hóa 206, vào (một) trường siêu dữ liệu 403 của siêu khung 400. Do vậy, siêu khung 400 có thể cung cấp siêu dữ liệu 202, 205 cho một hoặc nhiều (nhất là nhiều) khung của một hoặc nhiều tín hiệu kênh giảm mức 203, nhờ đó cho phép cuộc truyền hiệu quả của tín hiệu IA 111.

Cụ thể, khung của tín hiệu kênh giảm mức 203 có thể được tạo ra nhờ sử dụng bộ mã hóa-giải mã âm thanh hoặc thoại nhiều chế độ và/hoặc nhiều tốc độ. Hơn thế nữa,

siêu dữ liệu 202, 205 có thể được tạo ra nhờ sử dụng sơ đồ mã hóa siêu dữ liệu nổi sống động nhiều chế độ và/hoặc nhiều tốc độ. Thông tin cấu hình biểu thị hoạt động của bộ mã hóa-giải mã âm thanh hoặc thoại nhiều chế độ và/hoặc nhiều tốc độ (đã được sử dụng cho tín hiệu kênh giảm mức 203) và/hoặc hoạt động của sơ đồ mã hóa siêu dữ liệu nổi sống động nhiều chế độ và/hoặc nhiều tốc độ có thể được bao gồm trong trường thông tin cấu hình 402 của siêu khung (hiện thời) 400, có thể được bao gồm trong trường thông tin cấu hình 402 của siêu khung 400 trước đó của chuỗi siêu khung 400 hoặc có thể được truyền nhờ sử dụng sơ đồ báo hiệu ngoài băng. Kết quả là, sơ đồ hiệu quả và linh hoạt để mã hóa tín hiệu âm thanh nổi sống động 111 có thể được tạo ra.

Siêu khung 400 có thể bao gồm dữ liệu âm thanh mã hóa 206 kết hợp với nhiều tín hiệu kênh giảm mức 203. Dữ liệu âm thanh mã hóa 206 của khung của tín hiệu kênh giảm mức 203 thứ nhất có thể được tạo ra nhờ sử dụng phiên bản thứ nhất của bộ mã hóa-giải mã âm thanh hoặc thoại nhiều chế độ và/hoặc nhiều tốc độ. Hơn thế nữa, dữ liệu âm thanh mã hóa 206 của khung của tín hiệu kênh giảm mức 203 thứ hai có thể được tạo ra nhờ sử dụng phiên bản thứ hai của bộ mã hóa-giải mã âm thanh hoặc thoại nhiều chế độ và/hoặc nhiều tốc độ, trong đó phiên bản thứ nhất và thứ hai của bộ mã hóa-giải mã âm thanh hoặc thoại nhiều chế độ và/hoặc nhiều tốc độ có thể khác nhau. Thông tin cấu hình (được bao gồm trong siêu khung hiện thời 400, siêu khung 400 trước đó hoặc được truyền ngoài băng) có thể là biểu thị hoạt động của phiên bản thứ nhất và thứ hai (nhất là của mỗi phiên bản) của bộ mã hóa-giải mã âm thanh hoặc thoại nhiều chế độ và/hoặc nhiều tốc độ. Nhờ làm vậy, tính linh hoạt và hiệu quả để mã hóa tín hiệu âm thanh nổi sống động 111 có thể được tăng thêm.

Nói cách khác, phương pháp 500 có thể bao gồm bước chèn dữ liệu âm thanh mã hóa 206 cho một hoặc nhiều khung của tín hiệu kênh giảm mức 203 thứ nhất và tín hiệu kênh giảm mức 203 thứ hai thu được từ tín hiệu âm thanh nổi sống động 111, lần lượt vào một hoặc nhiều trường dữ liệu 411, 421 thứ nhất và một hoặc nhiều trường dữ liệu 412, 422 thứ hai của siêu khung 400. Tín hiệu kênh giảm mức 203 thứ nhất có thể được mã hóa nhờ sử dụng bộ mã hóa (âm thanh hoặc thoại) thứ nhất, và tín hiệu kênh giảm mức thứ hai có thể được mã hóa nhờ sử dụng bộ mã hóa (âm thanh hoặc thoại) thứ hai. Bộ mã hóa thứ nhất và thứ hai có thể khác nhau hoặc có thể được vận hành nhờ sử dụng cấu hình khác. Hơn thế nữa, phương pháp 500 có thể bao gồm bước cung cấp thông tin

cấu hình liên quan đến bộ mã hóa thứ nhất và bộ mã hóa thứ hai trong siêu khung 400, trong siêu khung 400 trước đó của chuỗi siêu khung 400 hoặc nhờ sử dụng sơ đồ báo hiệu ngoài băng. Nhờ làm vậy, tính linh hoạt và hiệu quả để mã hóa tín hiệu âm thanh nổi sống động 111 có thể được tăng thêm.

Phương pháp 500 có thể bao gồm bước chèn trường phần đầu 401 vào siêu khung 400. Trường phần đầu 401 có thể là biểu thị kích thước của trường siêu dữ liệu 403 của siêu khung 400, nhờ đó cho phép kích thước của siêu khung 400 được làm thích ứng theo cách linh hoạt để thay đổi chiều dài của siêu dữ liệu (được mã hóa entropy và/hoặc không tổn hao) 207.

Trường siêu dữ liệu 403 có thể biểu thị kích thước tối đa có thể có (mà có thể ví dụ, được chỉ báo trong trường thông tin cấu hình tùy ý 402 của siêu khung 400). Trường phần đầu 401 có thể biểu thị giá trị điều chỉnh, và kích thước của trường siêu dữ liệu 403 của siêu khung 400 có thể tương ứng với kích thước lớn nhất có thể có trừ đi giá trị điều chỉnh, nhờ đó cho phép kích thước của trường siêu dữ liệu 403 được báo hiệu theo cách hiệu quả và chính xác.

Trường phần đầu 401 có thể bao gồm chỉ báo kích thước (ví dụ, giá trị điều chỉnh) cho kích thước của trường siêu dữ liệu 403. Chỉ báo kích thước có thể biểu thị độ phân giải hoặc kích thước bước khác nhau (liên quan đến các khoảng cách kích thước) cho các khoảng kích thước khác nhau của kích thước của trường siêu dữ liệu 403. Độ phân giải và/hoặc kích thước bước của chỉ báo kích thước có thể phụ thuộc vào sự phân bố kích thước thống kê của siêu dữ liệu (được mã hóa entropy). Bằng cách cung cấp cho chỉ báo kích thước độ phân giải thay đổi, hiệu suất tốc độ bit để báo hiệu kích thước của trường siêu dữ liệu 403 có thể được cải thiện.

Trường phần đầu 401 có thể biểu thị việc siêu khung 400 có bao gồm trường thông tin cấu hình 402 hay không. Nói cách khác, trường phần đầu 401 có thể biểu thị sự có mặt của trường thông tin cấu hình 402. Trường thông tin cấu hình 402 có thể chỉ được chèn vào siêu khung 400 nếu cần (ví dụ, nếu cấu hình của bộ mã hóa của tín hiệu IA 111 đã thay đổi). Kết quả là, hiệu suất tốc độ bit của chuỗi siêu khung 400 có thể được cải thiện.

Trường phần đầu 401 có thể chỉ báo rằng không có trường thông tin cấu hình 402 có mặt trong siêu khung hiện thời 400. Phương pháp 500 có thể bao gồm bước truyền thông tin cấu hình trong siêu khung 400 trước đó của chuỗi siêu khung 400 hoặc nhờ sử dụng sơ đồ báo hiệu ngoài băng. Kết quả là, thông tin cấu hình (mà ít nhất là tĩnh tạm thời) có thể được truyền theo cách hiệu quả.

Theo cách khác, hoặc ngoài ra, trường phần đầu 401 có thể là biểu thị việc siêu khung 400 có bao gồm trường mở rộng 404 cho thông tin bổ sung liên quan đến tín hiệu âm thanh nổi sống động 111 hay không. Kết quả là, cấu trúc siêu khung có thể được làm thích ứng theo cách linh hoạt với các phần mở rộng trong tương lai.

Phương pháp 500 có thể bao gồm bước chèn trường thông tin cấu hình 402 vào siêu khung 400 (nếu cần). Trường thông tin cấu hình 402 có thể là biểu thị số lượng tín hiệu kênh giảm mức 203 được bao gồm trong các trường dữ liệu 411, 421, 412, 422 của siêu khung 400. Theo cách khác, hoặc ngoài ra, trường thông tin cấu hình 402 có thể biểu thị thứ tự của tín hiệu biểu diễn trường âm thanh được bao gồm trong tín hiệu IA 111. Kết quả là, các loại tín hiệu IA 111 khác nhau (với các loại tín hiệu SR khác nhau) có thể được mã hóa và truyền.

Trường thông tin cấu hình 402 có thể biểu thị kích thước lớn nhất có thể có của trường siêu dữ liệu 403. Theo cách khác, hoặc ngoài ra, trường thông tin cấu hình 402 có thể biểu thị loại khung và/hoặc chế độ mã hóa được sử dụng để mã hóa mỗi tín hiệu trong số một hoặc nhiều tín hiệu kênh giảm mức 203. Việc cung cấp thông tin này cho phép việc sử dụng các sơ đồ mã hóa khác nhau để mã hóa tín hiệu IA 111.

Dữ liệu âm thanh mã hóa 206 của khung của tín hiệu kênh giảm mức 203 có thể được tạo ra nhờ sử dụng bộ mã hóa-giải mã âm thanh hoặc thoại nhiều chế độ và/hoặc nhiều tốc độ. Theo cách khác, hoặc ngoài ra, siêu dữ liệu (mã hóa) 207 có thể được tạo ra nhờ sử dụng sơ đồ mã hóa siêu dữ liệu nổi sống động nhiều chế độ và/hoặc nhiều tốc độ. Kết quả là, các tín hiệu IA 111 có thể được mã hóa có chất lượng tương đối cao và ở tốc độ dữ liệu tương đối thấp.

Siêu khung 400 của chuỗi siêu khung 400 có thể cấu thành ít nhất một phần của phần tử dữ liệu được truyền nhờ sử dụng giao thức truyền, nhất là DASH, RTSP hoặc RTP, hoặc lưu trữ trong tập tin theo định dạng lưu trữ, nhất là ISOBMFF. Nói cách khác,

dòng bit 101 chứa chuỗi siêu khung 400 có thể sử dụng một hoặc nhiều phần từ dữ liệu của giao thức truyền hoặc của định dạng lưu trữ. Nhờ đó cho phép dòng bit 101 được truyền hoặc lưu trữ theo cách hiệu quả và tin cậy.

Fig.6 thể hiện lưu đồ của phương pháp làm ví dụ 600 để thu được dữ liệu liên quan đến tín hiệu âm thanh nổi sóng động 111 từ dòng bit 101. Dòng bit 101 bao gồm chuỗi siêu khung 400 cho chuỗi các khung của tín hiệu âm thanh nổi sóng động 111. Theo một ví dụ ưu tiên, nhiều khung (cơ bản) của tín hiệu IA 111 được bao gồm trong một siêu khung 400. Cần lưu ý rằng tất cả các dấu hiệu được mô tả trong ngữ cảnh của phương pháp 500 tạo dòng bit 101 có thể được áp dụng theo cách tương tự cho phương pháp 600 để thu được dữ liệu từ dòng bit 101.

Tín hiệu IA 111 có thể bao gồm tín hiệu SR, tín hiệu đa kênh và/hoặc một hoặc nhiều đối tượng âm thanh. Các khía cạnh và/hoặc dấu hiệu mà được mô tả trong ngữ cảnh của phương pháp 500 và/hoặc trong ngữ cảnh của thiết bị mã hóa 110 cũng có thể được áp dụng theo cách tương tự và/hoặc bổ sung cho phương pháp 600 và/hoặc cho thiết bị giải mã 120 (và ngược lại).

Phương pháp 600 bao gồm, theo cách lặp lại cho chuỗi siêu khung 400, bước trích xuất 601 dữ liệu âm thanh mã hóa 206 cho một hoặc nhiều (nhất là nhiều) khung của một hoặc nhiều tín hiệu kênh giảm mức 203 thu được từ tín hiệu âm thanh nổi sóng động 111, từ các trường dữ liệu 411, 421, 412, 422 của siêu khung 400. Hơn thế nữa, phương pháp 600 bao gồm bước trích xuất 602 siêu dữ liệu (mã hóa) 207 để tái tạo một hoặc nhiều (nhất là nhiều) khung của tín hiệu âm thanh nổi sóng động 111 từ dữ liệu âm thanh mã hóa 206 từ trường siêu dữ liệu 403 của siêu khung 400.

Phương pháp 600 có thể bao gồm bước thu được một hoặc nhiều đối tượng âm thanh tái tạo từ dữ liệu âm thanh mã hóa 206 và từ siêu dữ liệu 207 (nhất là từ siêu dữ liệu đối tượng 202). Như chỉ ra ở trên, đối tượng âm thanh thường bao gồm tín hiệu đối tượng và siêu dữ liệu đối tượng 202 chỉ báo vị trí (thay đổi theo thời gian) của đối tượng âm thanh. Hơn thế nữa, phương pháp 600 có thể bao gồm bước thu được tín hiệu dư tái tạo từ dữ liệu âm thanh mã hóa 206 và từ siêu dữ liệu 202, 205. Một hoặc nhiều đối tượng âm thanh tái tạo và tín hiệu dư tái tạo có thể mô tả và/hoặc có thể biểu thị tín hiệu IA 111. Cụ thể, dữ liệu (như thứ tự của tín hiệu SR được bao gồm trong tín hiệu IA 111)

có thể được trích xuất từ dòng bit 101, mà cho phép việc xác định tín hiệu IA tái tạo 121, trong đó tín hiệu IA tái tạo 121 là dạng gần đúng của tín hiệu IA 111 ban đầu.

Như chỉ ra ở trên, phương pháp 600 để thu được dữ liệu liên quan đến tín hiệu âm thanh nổi sóng động 111 từ dòng bit 101 có thể bao gồm các dấu hiệu tương ứng với phương pháp 500 tạo dòng bit 101. Cụ thể, phương pháp 600 có thể bao gồm bước trích xuất trường phần đầu 401 từ siêu khung 400 cho trước. Kích thước của trường siêu dữ liệu 403 của siêu khung 400 cho trước có thể được thu được từ trường phần đầu 401.

Kích thước của trường siêu dữ liệu 403 có thể được chỉ báo như được nêu trong ngữ cảnh của phương pháp 500. Trường siêu dữ liệu 403 có thể biểu thị kích thước tối đa có thể có, và trường phần đầu 401 có thể biểu thị giá trị điều chỉnh, trong đó kích thước của trường siêu dữ liệu 403 của siêu khung 400 có thể tương ứng với kích thước lớn nhất có thể có trừ đi giá trị điều chỉnh. Cụ thể, trường phần đầu 401 có thể bao gồm chỉ báo kích thước cho kích thước của trường siêu dữ liệu 403, trong đó chỉ báo kích thước có thể biểu thị độ phân giải khác nhau cho các khoảng kích thước khác nhau của kích thước của trường siêu dữ liệu 403. Kết quả là, kích thước của trường siêu dữ liệu 403 có thể được báo hiệu theo cách có hiệu quả về tốc độ bit.

Phương pháp 600 có thể bao gồm bước xác định, dựa trên trường phần đầu 401, xem siêu khung 400 có bao gồm trường thông tin cấu hình 402 không và/hoặc xem trường thông tin cấu hình 402 có mặt trong siêu khung 400 không. Nếu trường thông tin cấu hình 402 không có mặt, thì thông tin cấu hình đã được cung cấp trong siêu khung 400 trước đó và/hoặc đã được cung cấp ngoài băng có thể được sử dụng để xử lý một hoặc nhiều khung của một hoặc nhiều tín hiệu kênh giảm mức 203 được bao gồm trong siêu khung 400. Nếu trường thông tin cấu hình 402 có mặt, thì thông tin cấu hình được bao gồm trong siêu khung 400 có thể được sử dụng để xử lý một hoặc nhiều khung của một hoặc nhiều tín hiệu kênh giảm mức 203 được bao gồm trong siêu khung 400.

Ngoài ra, phương pháp 600 có thể bao gồm bước xác định, dựa trên trường phần đầu 401, liệu siêu khung 400 có bao gồm trường mở rộng 404 cho thông tin bổ sung liên quan đến tín hiệu âm thanh nổi sóng động 111 hay không, nhờ đó tạo ra phương tiện hiệu quả và linh hoạt để truyền thông tin trong dòng bit 101.

Phương pháp 600 có thể bao gồm bước trích xuất trường thông tin cấu hình 402 từ siêu khung 400. Hơn thế nữa, phương pháp 600 có thể bao gồm bước xác định, dựa trên trường thông tin cấu hình 402, số lượng tín hiệu kênh giảm mức 203 được biểu diễn bởi các trường dữ liệu 411, 421, 412, 422 của siêu khung 400, nhờ đó cho phép việc xử lý chính xác một hoặc nhiều khung của một hoặc nhiều tín hiệu kênh giảm mức 203 được bao gồm trong siêu khung 400.

Hơn thế nữa, phương pháp 600 có thể bao gồm bước xác định, dựa trên trường thông tin cấu hình 402, kích thước lớn nhất có thể có của trường siêu dữ liệu 403.

Ngoài ra, phương pháp 600 có thể bao gồm bước xác định, dựa trên trường thông tin cấu hình 402, bậc của tín hiệu âm thanh nổi sóng động 111, để cho phép sự tái tạo chính xác của tín hiệu IA 111.

Phương pháp 600 cũng có thể bao gồm bước xác định, dựa trên trường thông tin cấu hình 402, loại khung và/hoặc chế độ mã hóa được sử dụng để mã hóa mỗi tín hiệu trong số một hoặc nhiều tín hiệu kênh giảm mức, nhờ đó cho phép việc xử lý chính xác một hoặc nhiều khung của một hoặc nhiều tín hiệu kênh giảm mức 203 được bao gồm trong siêu khung 400.

Các phương án ví dụ khác nhau của sáng chế có thể được thực hiện ở phần cứng hoặc mạch chuyên dụng, phần mềm, logic hoặc tổ hợp bất kỳ của chúng. Một số khía cạnh có thể được thực hiện ở phần cứng, trong khi các khía cạnh khác có thể được thực hiện trong firmware hoặc phần mềm, mà có thể được thực thi bởi bộ điều khiển, bộ vi xử lý hoặc thiết bị máy tính khác. Nhìn chung, sáng chế được hiểu là còn bao gồm thiết bị thích hợp để thực hiện các phương pháp mô tả ở trên, chẳng hạn thiết bị (bộ kết xuất không gian) có bộ nhớ và bộ xử lý được nối với bộ nhớ, trong đó bộ xử lý được tạo cấu hình để thực thi các lệnh và để thực hiện các phương pháp theo các phương án của sáng chế.

Mặc dù các khía cạnh khác nhau của các phương án ví dụ của sáng chế được minh họa và mô tả dưới dạng các sơ đồ khối, lưu đồ, hoặc nhờ sử dụng một dạng biểu diễn hình ảnh khác nào đó, nhưng sẽ hiểu rằng các khối, thiết bị, hệ thống, kỹ thuật hoặc các phương pháp được mô tả ở đây có thể được thực hiện trong, dưới dạng các ví dụ không giới hạn, phần cứng, phần mềm, firmware, mạch chuyên dụng hoặc logic, phần cứng đa

dụng hoặc bộ điều khiển, hoặc các thiết bị máy tính khác, hoặc một tổ hợp nào đó của chúng.

Ngoài ra, các khối khác nhau thể hiện trong lưu đồ có thể được xem là các bước của phương pháp, và/hoặc là các hoạt động là kết quả từ việc thực hiện mã chương trình máy tính, và/hoặc là nhiều phần tử mạch logic ghép nối được tạo cấu trúc để thực hiện (các) chức năng liên quan. Chẳng hạn, các phương án theo sáng chế bao gồm sản phẩm chương trình máy tính chứa chương trình máy tính hữu hình trên phương tiện đọc được bằng máy, trong đó chương trình máy tính chứa các mã chương trình được tạo cấu trúc để thực hiện các phương pháp như được mô tả ở trên.

Trong ngữ cảnh của sáng chế, phương tiện đọc được bằng máy có thể là phương tiện hữu hình bất kỳ mà có thể chứa, hoặc lưu trữ, chương trình để sử dụng bởi hoặc kết nối với hệ thống, thiết bị hoặc máy thực thi lệnh. Phương tiện đọc được bằng máy có thể là phương tiện tín hiệu đọc được bằng máy hoặc phương tiện lưu trữ đọc được bằng máy. Phương tiện đọc được bằng máy có thể bao gồm nhưng không chỉ giới hạn ở hệ thống, thiết bị hoặc máy điện tử, từ tính, quang, điện từ, hồng ngoại hoặc bán dẫn, hoặc tổ hợp thích hợp bất kỳ của các loại trên. Các ví dụ cụ thể về phương tiện lưu trữ đọc được bằng máy sẽ bao gồm kết nối điện có một hoặc nhiều dây, đĩa máy tính di chuyển được, đĩa cứng, bộ nhớ truy cập ngẫu nhiên (RAM), bộ nhớ chỉ đọc (ROM), bộ nhớ chỉ đọc lập trình được xóa được (EPROM hoặc bộ flash), sợi quang, bộ nhớ chỉ đọc đĩa CD di chuyển được (CD-ROM), thiết bị lưu trữ quang, thiết bị lưu trữ từ, hoặc tổ hợp thích hợp bất kỳ của các loại trên.

Mã chương trình máy tính để thực hiện các phương pháp của sáng chế có thể được viết bằng tổ hợp bất kỳ của một hoặc nhiều ngôn ngữ lập trình. Các mã chương trình máy tính này có thể được cung cấp cho bộ xử lý của máy tính đa dụng, máy tính chuyên dụng, hoặc thiết bị xử lý dữ liệu lập trình được khác, sao cho các mã chương trình, khi được thực thi bởi bộ xử lý của máy tính hoặc thiết bị xử lý dữ liệu lập trình được khác, khiến cho các chức năng/hoạt động được xác định trong các lưu đồ và/hoặc sơ đồ khối được thực hiện. Mã chương trình có thể thực thi toàn bộ trên máy tính, một phần trên máy tính, dưới dạng gói phần mềm độc lập, một phần trên máy tính và một phần trên máy tính từ xa hoặc toàn bộ trên máy tính từ xa hoặc máy chủ.

Hơn nữa, mặc dù các hoạt động được mô tả theo một thứ tự cụ thể, nhưng việc này không nên được hiểu là yêu cầu các hoạt động như vậy phải được thực hiện theo thứ tự cụ thể được thể hiện hoặc theo thứ tự hoặc tất cả các hoạt động được minh họa phải được thực hiện, để đạt được kết quả mong muốn. Trong một số trường hợp, xử lý đa nhiệm và xử lý song song có thể là một ưu điểm. Tương tự, mặc dù một vài chi tiết phương án thực hiện cụ thể được bao gồm trong phần mô tả trên, nhưng các chi tiết này không nên được hiểu là các giới hạn đối với phạm vi của sáng chế bất kỳ, hoặc của những gì có thể được yêu cầu bảo hộ, mà các chi tiết này chỉ là phần mô tả các dấu hiệu mà có thể là đặc trưng cho các phương án cụ thể của sáng chế cụ thể. Các dấu hiệu được mô tả trong bản mô tả này trong ngữ cảnh của các phương án riêng biệt cũng có thể được thực hiện ở dạng kết hợp trong một phương án duy nhất. Ngược lại, các dấu hiệu khác nhau được mô tả trong ngữ cảnh một phương án duy nhất cũng có thể được thực hiện theo nhiều phương án riêng biệt hoặc theo dạng tổ hợp con thích hợp bất kỳ.

Cần lưu ý rằng phần mô tả và hình vẽ chỉ minh họa các nguyên lý của phương pháp và thiết bị được đề xuất. Do đó người có hiểu biết trong lĩnh vực này sẽ hiểu là có thể nghĩ ra các bố trí khác nhau mà, mặc dù không được mô tả và thể hiện rõ ràng ở đây, thực thi các nguyên lý của sáng chế và được bao gồm trong ý tưởng và phạm vi của sáng chế. Hơn thế nữa, tất cả các ví dụ được đề cập về nguyên tắc được dự tính rõ ràng để chỉ cho các mục đích sư phạm để giúp người đọc hiểu rõ các nguyên lý của các phương pháp và thiết bị được đề xuất và các khái niệm được đóng góp bởi các tác giả sáng chế để cải tiến lĩnh vực này, và cần được hiểu là không giới hạn ở các ví dụ và điều kiện được đề cập cụ thể này. Ngoài ra, tất cả các nội dung trình bày ở đây đề cập đến các nguyên lý, khía cạnh và phương án của sáng chế, cũng như các ví dụ cụ thể của chúng, được dự tính bao gồm các dạng biến đổi tương đương của chúng.

YÊU CẦU BẢO HỘ

1. Phương pháp tạo dòng bit; trong đó dòng bit bao gồm chuỗi siêu khung cho chuỗi các khung của tín hiệu âm thanh nổi sống động ambisonic; trong đó phương pháp này bao gồm các bước:

theo cách lặp lại cho chuỗi siêu khung,

- chèn dữ liệu âm thanh mã hóa cho hai hoặc nhiều khung của một hoặc nhiều tín hiệu kênh giảm mức thu được từ tín hiệu âm thanh nổi sống động ambisonic, vào trường dữ liệu của siêu khung trong chuỗi các siêu khung trong dòng bit, trong đó mỗi khung trong số hai hoặc nhiều khung trong siêu khung được mã hóa với trường dữ liệu tương ứng được chỉ định để hỗ trợ các hoạt động tái đồng bộ hóa mức khung bởi thiết bị nhận của tín hiệu âm thanh nổi sống động ambisonic trong trường hợp có lỗi bit trong tín hiệu âm thanh nổi sống động ambisonic; và
- chèn siêu dữ liệu để tái tạo một hoặc nhiều khung của tín hiệu âm thanh nổi sống động ambisonic từ dữ liệu âm thanh mã hóa vào trường siêu dữ liệu của siêu khung.

2. Phương pháp theo điểm 1, trong đó

- a. phương pháp này bao gồm bước chèn trường phần đầu vào siêu khung; và
- b. trường phần đầu biểu thị kích thước của trường siêu dữ liệu của siêu khung.

3. Phương pháp theo điểm 2, trong đó

- a. trường siêu dữ liệu có kích thước dữ liệu không lớn hơn kích thước lớn nhất;
- b. trường phần đầu biểu thị giá trị điều chỉnh; và
- c. kích thước của trường siêu dữ liệu của siêu khung tương ứng với kích thước lớn nhất có trừ đi giá trị điều chỉnh.

4. Phương pháp theo điểm 2, trong đó

- a. trường phần đầu bao gồm chỉ báo kích thước cho kích thước của trường siêu dữ liệu; và
- b. chỉ báo kích thước biểu thị độ phân giải khác nhau cho các khoảng kích thước khác nhau của kích thước của trường siêu dữ liệu.

5. Phương pháp theo điểm 4, trong đó

- a. siêu dữ liệu để tái tạo một hoặc nhiều khung của tín hiệu âm thanh nổi sống động ambisonic biểu thị sự phân bố kích thước thống kê của kích thước của siêu dữ liệu; và
- b. độ phân giải của chỉ báo kích thước phụ thuộc vào sự phân bố kích thước của siêu dữ liệu.

6. Phương pháp theo điểm 1, trong đó

- a. phương pháp này bao gồm bước chèn trường phần đầu vào siêu khung; và
- b. trường phần đầu biểu thị việc siêu khung có bao gồm trường thông tin cấu hình hay không, và/hoặc
- c. trường phần đầu biểu thị sự có mặt của trường thông tin cấu hình.

7. Phương pháp theo điểm 1, trong đó

- a. phương pháp này bao gồm bước chèn trường thông tin cấu hình vào siêu khung; và
- b. trường thông tin cấu hình biểu thị một số tín hiệu kênh giảm mức được biểu diễn bởi các trường dữ liệu của siêu khung.

8. Phương pháp theo điểm 1, trong đó

- a. phương pháp này bao gồm bước chèn trường thông tin cấu hình vào siêu khung; và
- b. trường thông tin cấu hình biểu thị kích thước lớn nhất của trường siêu dữ liệu.

9. Phương pháp theo điểm 1, trong đó

- a. phương pháp này bao gồm bước chèn trường thông tin cấu hình vào siêu khung; và
- b. trường thông tin cấu hình biểu thị bậc của tín hiệu biểu diễn trường âm thanh được bao gồm trong tín hiệu âm thanh nổi sống động ambisonic.

10. Phương pháp theo điểm 1, trong đó

- a. phương pháp này bao gồm bước chèn trường thông tin cấu hình vào siêu khung; và
- b. trường thông tin cấu hình biểu thị loại khung và/hoặc chế độ mã hóa được sử dụng để mã hóa mỗi tín hiệu trong số một hoặc nhiều tín hiệu kênh giảm mức.

11. Phương pháp theo điểm 1, trong đó

- a. phương pháp này bao gồm bước chèn trường phân đầu vào siêu khung; và
- b. trường phân đầu biểu thị việc siêu khung có bao gồm trường mở rộng cho thông tin bổ sung liên quan đến tín hiệu âm thanh nổi sống động ambisonic hay không.

12. Phương pháp theo điểm 1, trong đó

- a. dữ liệu âm thanh mã hóa của khung của tín hiệu kênh giảm mức được tạo ra nhờ sử dụng bộ mã hóa-giải mã âm thanh hoặc thoại nhiều chế độ và/hoặc nhiều tốc độ; và/hoặc
- b. siêu dữ liệu được tạo ra nhờ sử dụng sơ đồ mã hóa siêu dữ liệu nổi sống động nhiều chế độ và/hoặc nhiều tốc độ.

13. Phương pháp theo điểm 1, trong đó dữ liệu âm thanh mã hóa của khung của tín hiệu kênh giảm mức được mã hóa nhờ sử dụng bộ mã hóa dịch vụ thoại nâng cao.

14. Phương pháp theo điểm 1, trong đó siêu khung cấu thành ít nhất một phần của phần tử dữ liệu được truyền nhờ sử dụng giao thức truyền, nhất là DASH, RTSP hoặc RTP, hoặc được lưu trữ trong tập tin theo định dạng lưu trữ, nhất là ISOBMFF.

15. Phương pháp theo điểm 1, trong đó

- a. trường phân đầu biểu thị rằng không trường thông tin cấu hình nào có mặt; và
- b. phương pháp này bao gồm bước truyền thông tin cấu hình trong siêu khung trước đó của chuỗi siêu khung hoặc nhờ sử dụng sơ đồ báo hiệu ngoài băng.

16. Phương pháp theo điểm 1, trong đó một hoặc nhiều tín hiệu kênh giảm mức bao gồm tín hiệu kênh giảm mức thứ nhất và tín hiệu kênh giảm mức thứ hai; trong đó tín hiệu kênh giảm mức thứ nhất và tín hiệu kênh giảm mức thứ hai được suy ra từ tín hiệu âm thanh nổi sống động ambisonic; trong đó tín hiệu kênh giảm mức thứ nhất được mã hóa nhờ sử dụng bộ mã hóa thứ nhất, và trong đó tín hiệu kênh giảm mức thứ hai được mã hóa nhờ sử dụng bộ mã hóa thứ hai; trong đó phương pháp này bao gồm bước:

- cung cấp thông tin cấu hình liên quan đến bộ mã hóa thứ nhất và bộ mã hóa thứ hai trong siêu khung, trong siêu khung trước đó của chuỗi siêu khung hoặc nhờ sử dụng sơ đồ báo hiệu ngoài băng.

17. Phương pháp theo điểm 1, trong đó phương pháp này bao gồm các bước:

- a. trích xuất một hoặc nhiều đối tượng âm thanh từ tín hiệu âm thanh nổi sống động, được gọi là IA; trong đó đối tượng âm thanh bao gồm tín hiệu đối tượng và siêu dữ liệu đối tượng biểu thị vị trí của đối tượng âm thanh;
- b. xác định tín hiệu dư dựa trên tín hiệu IA và dựa trên một hoặc nhiều đối tượng âm thanh;
- c. cung cấp tín hiệu giảm mức dựa trên tín hiệu IA, nhất là sao cho số lượng tín hiệu kênh giảm mức của tín hiệu giảm mức nhỏ hơn số lượng tín hiệu kênh của tín hiệu IA;
- d. xác định siêu dữ liệu mã hóa chung để cho phép tăng mức tín hiệu giảm mức thành một hoặc nhiều tín hiệu đối tượng âm thanh tái tạo tương ứng với một hoặc nhiều đối tượng âm thanh và/hoặc với tín hiệu dư tái tạo tương ứng với tín hiệu dư;
- e. thực hiện việc mã hóa dạng sóng tín hiệu giảm mức để tạo ra dữ liệu âm thanh mã hóa cho chuỗi các khung của một hoặc nhiều tín hiệu kênh giảm mức; và
- f. thực hiện việc mã hóa entropy siêu dữ liệu mã hóa chung và siêu dữ liệu đối tượng của một hoặc nhiều đối tượng âm thanh để tạo ra siêu dữ liệu để được chèn vào các trường siêu dữ liệu của chuỗi siêu khung.

18. Phương pháp để thu được dữ liệu liên quan đến tín hiệu âm thanh nổi sống động ambisonic từ dòng bit; trong đó dòng bit bao gồm chuỗi siêu khung cho chuỗi các khung của tín hiệu âm thanh nổi sống động ambisonic; trong đó phương pháp bao gồm các bước, theo cách lặp lại cho chuỗi siêu khung,

- a. trích xuất dữ liệu âm thanh mã hóa cho hai hoặc nhiều khung của một hoặc nhiều tín hiệu kênh giảm mức thu được từ tín hiệu âm thanh nổi sống động ambisonic, từ trường dữ liệu của siêu khung trong chuỗi các siêu khung trong dòng bit, trong đó mỗi khung trong số hai hoặc nhiều khung trong siêu khung được mã hóa với trường dữ liệu tương ứng được chỉ định để hỗ trợ các hoạt động tái đồng bộ hóa mức khung bởi thiết bị nhận của tín hiệu âm thanh nổi sống động ambisonic trong trường hợp có lỗi bit trong tín hiệu âm thanh nổi sống động ambisonic; và

b. trích xuất siêu dữ liệu để tái tạo một hoặc nhiều khung của tín hiệu âm thanh nổi sống động ambisonic từ dữ liệu âm thanh mã hóa từ trường siêu dữ liệu của siêu khung.

19. Phương pháp theo điểm 18, phương pháp này còn bao gồm các bước:

- a. thu được một hoặc nhiều đối tượng âm thanh tái tạo từ dữ liệu âm thanh mã hóa và từ siêu dữ liệu; trong đó đối tượng âm thanh bao gồm tín hiệu đối tượng và siêu dữ liệu đối tượng biểu thị vị trí của đối tượng âm thanh; và
- b. thu được tín hiệu dư tái tạo từ dữ liệu âm thanh mã hóa và từ siêu dữ liệu; trong đó một hoặc nhiều đối tượng âm thanh tái tạo và tín hiệu dư tái tạo mô tả tín hiệu âm thanh nổi sống động ambisonic.

20. Phương pháp theo điểm 18, trong đó phương pháp này bao gồm các bước:

- a. trích xuất trường phần đầu từ siêu khung; và
- b. thu được kích thước của trường siêu dữ liệu của siêu khung từ trường phần đầu.

21. Phương pháp theo điểm 20, trong đó

- a. trường siêu dữ liệu có kích thước dữ liệu không lớn hơn kích thước lớn nhất;
- b. trường phần đầu biểu thị giá trị điều chỉnh; và
- c. kích thước của trường siêu dữ liệu của siêu khung tương ứng với kích thước lớn nhất có trừ đi giá trị điều chỉnh.

22. Phương pháp theo điểm 20, trong đó

- a. trường phần đầu bao gồm chỉ báo kích thước cho kích thước của trường siêu dữ liệu; và
- b. chỉ báo kích thước biểu thị độ phân giải khác nhau cho các khoảng kích thước khác nhau của kích thước của trường siêu dữ liệu.

23. Phương pháp theo điểm 20, trong đó phương pháp này bao gồm các bước:

- a. trích xuất trường phần đầu từ siêu khung; và
- b. xác định, dựa trên trường phần đầu, việc siêu khung có bao gồm trường thông tin cấu hình hay không; và/hoặc
- c. xác định, dựa trên trường phần đầu, xem trường thông tin cấu hình có mặt trong siêu khung hay không.

24. Phương pháp theo điểm 20, trong đó phương pháp này bao gồm các bước:

- a. trích xuất trường thông tin cấu hình từ siêu khung; và

b. xác định, dựa vào trường thông tin cấu hình, một số tín hiệu kênh giảm mức được biểu diễn bởi các trường dữ liệu của siêu khung.

25. Phương pháp theo điểm 20, trong đó phương pháp này bao gồm các bước:

- a. trích xuất trường thông tin cấu hình từ siêu khung; và
- b. xác định, dựa vào trường thông tin cấu hình, kích thước lớn nhất của trường siêu dữ liệu.

26. Phương pháp theo điểm 20, trong đó phương pháp này bao gồm các bước:

- a. trích xuất trường thông tin cấu hình từ siêu khung; và
- b. xác định, dựa vào trường thông tin cấu hình, bậc của tín hiệu biểu diễn trường âm thanh bao gồm trong tín hiệu âm thanh nổi sống động ambisonic.

27. Phương pháp theo điểm 18, trong đó phương pháp này bao gồm các bước:

- a. trích xuất trường thông tin cấu hình từ siêu khung; và
- b. xác định, dựa vào trường thông tin cấu hình, loại khung và/hoặc chế độ mã hóa được sử dụng để mã hóa mỗi tín hiệu trong số một hoặc nhiều tín hiệu kênh giảm mức.

28. Phương pháp theo điểm 18, trong đó phương pháp này bao gồm các bước:

- a. trích xuất trường phần đầu từ siêu khung; và
- b. xác định, dựa vào trường phần đầu, xem siêu khung có bao gồm trường mở rộng cho thông tin bổ sung liên quan đến tín hiệu âm thanh nổi sống động ambisonic hay không.

29. Thiết bị mã hóa được tạo cấu hình để tạo dòng bit; trong đó dòng bit bao gồm chuỗi siêu khung cho chuỗi các khung của tín hiệu âm thanh nổi sống động ambisonic;

trong đó thiết bị mã hóa được tạo cấu hình để, theo cách lặp lại cho chuỗi siêu khung,

- a. chèn dữ liệu âm thanh mã hóa cho hai hoặc nhiều khung của một hoặc nhiều tín hiệu kênh giảm mức thu được từ tín hiệu âm thanh nổi sống động ambisonic, vào trường dữ liệu của siêu khung trong chuỗi các siêu khung trong dòng bit, trong đó mỗi khung trong số hai hoặc nhiều khung trong siêu khung được mã hóa với trường dữ liệu tương ứng được chỉ định để hỗ trợ các hoạt động tái đồng bộ hóa mức khung bởi thiết bị nhận của tín hiệu âm thanh nổi sống động ambisonic trong trường hợp có lỗi bit trong tín hiệu âm thanh nổi sống động ambisonic; và

- chèn siêu dữ liệu để tái tạo một hoặc nhiều khung của tín hiệu âm thanh nổi sống động ambisonic từ dữ liệu âm thanh mã hóa vào trường siêu dữ liệu của siêu khung.

30. Thiết bị giải mã được tạo cấu hình để thu được dữ liệu liên quan đến tín hiệu âm thanh nổi sống động ambisonic từ dòng bit; trong đó dòng bit bao gồm chuỗi siêu khung cho chuỗi các khung của tín hiệu âm thanh nổi sống động ambisonic; trong đó thiết bị giải mã được tạo cấu hình để, theo cách lặp lại cho chuỗi siêu khung,

a. trích xuất dữ liệu âm thanh mã hóa cho hai hoặc nhiều khung của một hoặc nhiều tín hiệu kênh giảm mức thu được từ tín hiệu âm thanh nổi sống động ambisonic, từ trường dữ liệu của siêu khung trong chuỗi các siêu khung trong dòng bit, trong đó mỗi khung trong số hai hoặc nhiều khung trong siêu khung được mã hóa với trường dữ liệu tương ứng được chỉ định để hỗ trợ các hoạt động tái đồng bộ hóa mức khung bởi thiết bị nhận của tín hiệu âm thanh nổi sống động ambisonic trong trường hợp có lỗi bit trong tín hiệu âm thanh nổi sống động ambisonic; và

b. trích xuất siêu dữ liệu để tái tạo một hoặc nhiều khung của tín hiệu âm thanh nổi sống động ambisonic từ dữ liệu âm thanh mã hóa từ trường siêu dữ liệu của siêu khung.

1/3

HỆ THỐNG MÃ HÓA ẨM THANH

100

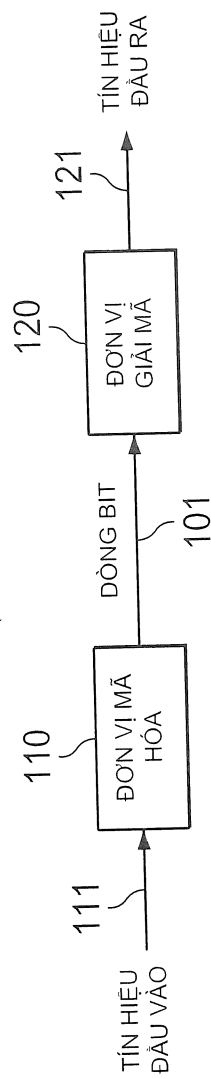


FIG. 1

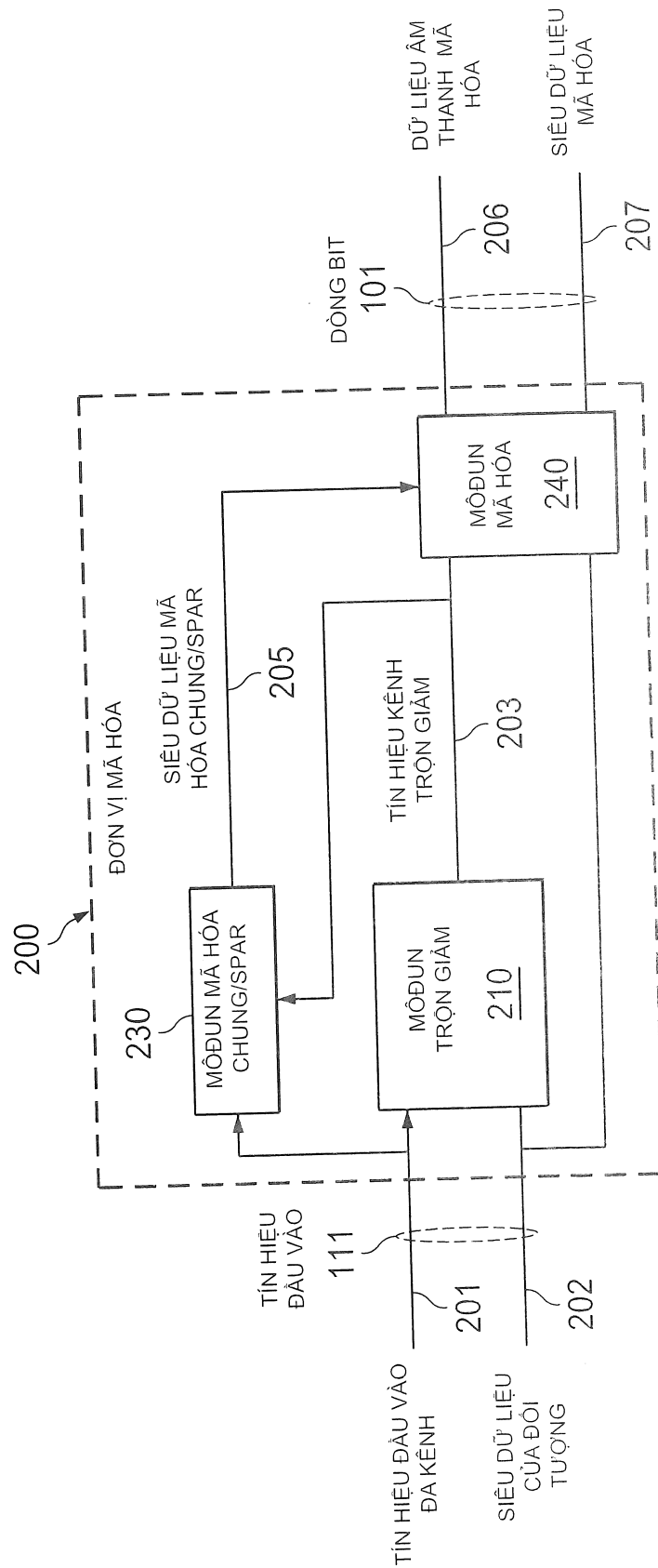


FIG. 2

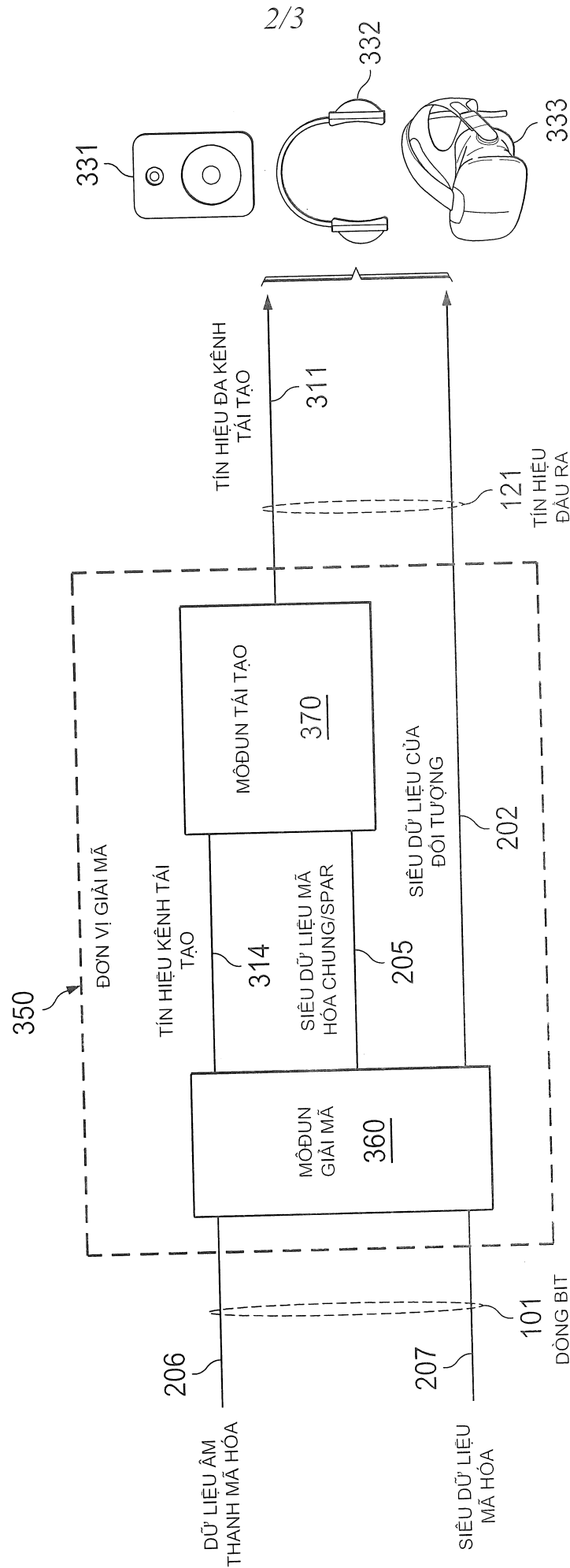


FIG. 3

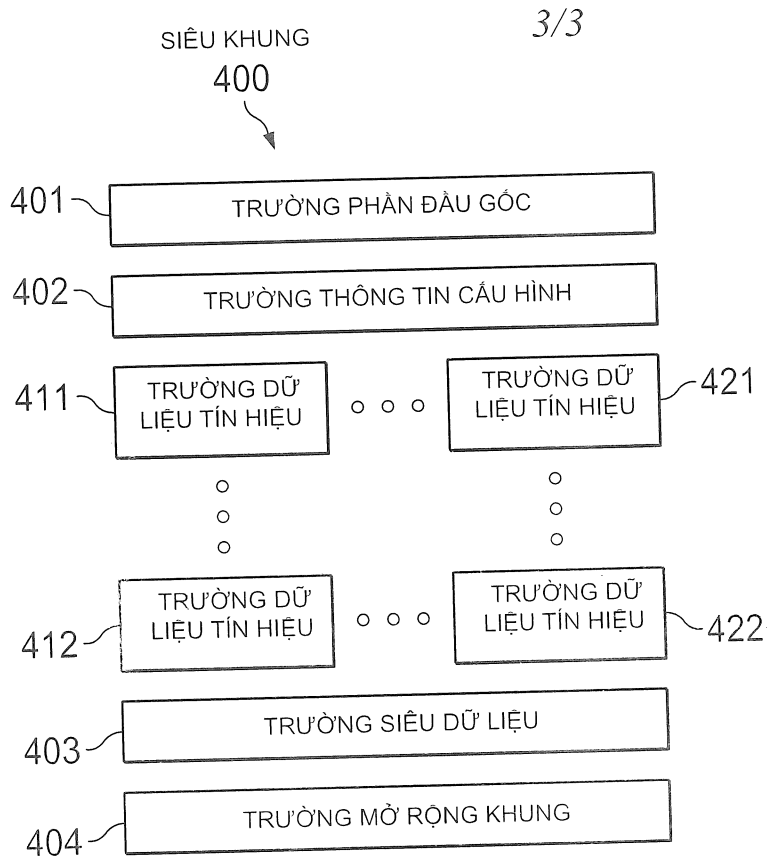


FIG. 4

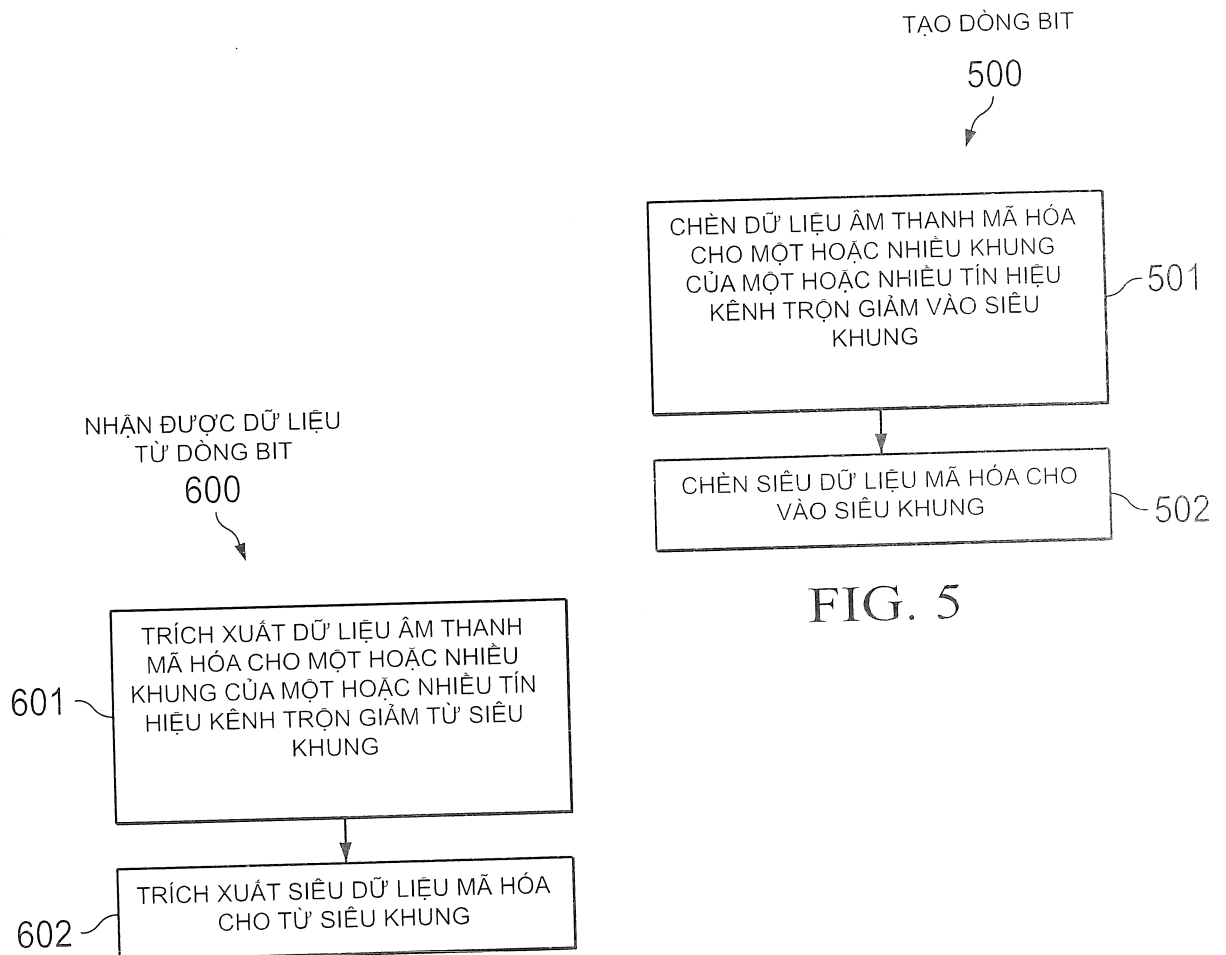


FIG. 5

FIG. 6