



(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)



1-0034246

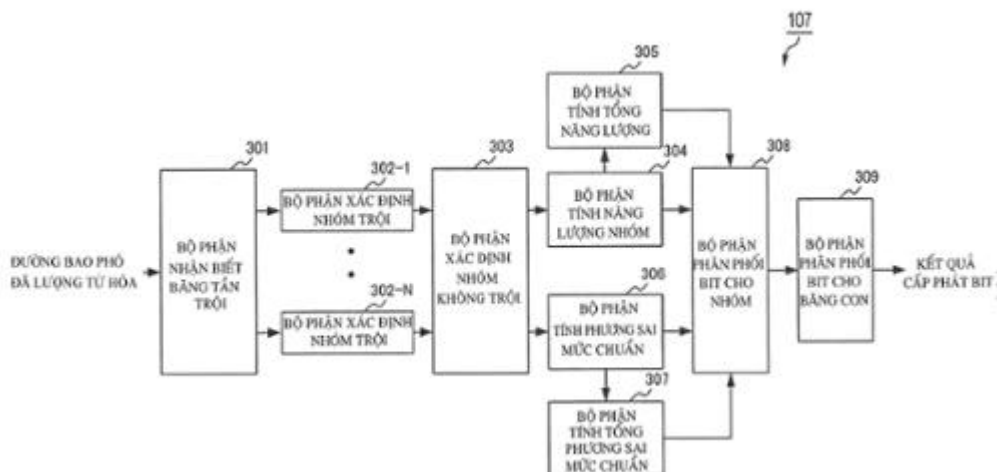
(51)⁷ G01L 19/035

(13) B

- (21) 1-2015-01777 (22) 26/11/2013
(86) PCT/JP2013/006948 26/11/2013 (87) WO 2014/091694 19/06/2014
(30) 2012-272571 13/12/2012 JP
(45) 26/12/2022 417 (43) 26/10/2015 331A
(73) Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V. (DE)
Hansastraße 27c, 80686 München, Germany
(72) Zongxian LIU (SG); Srikanth NAGISETTY (IN); Masahiro OSHIKIRI (JP).
(74) Văn phòng Luật sư Ân Nam (ANNAM IP & LAW)

(54) THIẾT BỊ MÃ HÓA ÂM THANH/GIỌNG NÓI, THIẾT BỊ GIẢI MÃ ÂM THANH/GIỌNG NÓI, PHƯƠNG PHÁP MÃ HÓA ÂM THANH/GIỌNG NÓI, PHƯƠNG PHÁP GIẢI MÃ ÂM THANH/GIỌNG NÓI

(57) Sáng chế đề cập đến thiết bị mã hóa âm thanh hoặc giọng nói, thiết bị giải mã âm thanh hoặc giọng nói, phương pháp mã hóa âm thanh hoặc giọng nói và phương pháp giải mã âm thanh hoặc giọng nói thực hiện hiệu quả sự phân phối bit và cải thiện chất lượng âm thanh. Bộ phận nhận biết băng tần trội (301) nhận biết băng tần trội có giá trị hệ số mức chuẩn là giá trị cao nhất trong phổ của tín hiệu âm thanh hoặc giọng nói đầu vào. Các bộ phận xác định nhóm trội (từ 302-1 đến 302-N) và bộ phận xác định nhóm bình thường (không trội) (303) ghép nhóm tất cả các băng con thành nhóm trội chứa băng tần trội và nhóm bình thường chứa băng tần không trội. Bộ phận phân phối bit cho nhóm (308) phân phối bit cho mỗi nhóm trên cơ sở năng lượng và phương sai mức chuẩn của mỗi nhóm. Bộ phận phân phối bit cho băng con (309) tái phân phối các bit đã được phân phối cho mỗi nhóm cho mỗi băng con theo tỷ lệ mức chuẩn trên năng lượng của các nhóm.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến thiết bị mã hóa âm thanh/giọng nói, thiết bị giải mã âm thanh/giọng nói, phương pháp mã hóa âm thanh/giọng nói và phương pháp giải mã âm thanh/giọng nói sử dụng mô hình mã hóa biến đổi.

Tình trạng kỹ thuật của sáng chế

Dưới dạng mô hình có khả năng mã hóa hiệu quả tín hiệu thoại hoặc tín hiệu âm nhạc trong toàn băng tần (FB) từ 0,02 đến 20 kHz, hiện có kỹ thuật được chuẩn hóa trong ITU-T (International Telecommunication Union Telecommunication Standardization Sector - Lĩnh vực Tiêu chuẩn viễn thông thuộc Tổ chức Viễn thông quốc tế). Kỹ thuật này biến đổi tín hiệu đầu vào thành tín hiệu miền tần số và mã hóa băng tần lên đến 20 kHz (mã hóa biến đổi).

Ở đây, mã hóa biến đổi là mô hình mã hóa biến đổi tín hiệu đầu vào từ miền thời gian thành miền tần số sử dụng sự biến đổi thời gian/tần số chẳng hạn như biến đổi cosin rời rạc (discrete cosine transform - DCT) hoặc biến đổi cosin rời rạc hiệu chỉnh (modified discrete cosine transform - MDCT) để cho phép tín hiệu được ánh xạ tương ứng chính xác với các đặc tính thính giác.

Trong mã hóa biến đổi, hệ số phổ được tách thành các băng con. Trong mã hóa mỗi băng con, việc cấp phát các bit lượng tử hóa nhiều hơn cho các băng tần quan trọng khi cảm nhận đối với tai người làm cho nó có thể cải thiện chất lượng âm thanh tổng thể.

Để đạt được mục đích này, các nghiên cứu đang được thực hiện trên các mô hình

cấp phát bit hiệu quả, và ví dụ, kỹ thuật được bộc lộ trong tài liệu phi sáng chế (sau đây được gọi tắt là "NPL") 1 đã biết. Sau đây, mô hình cấp phát bit được bộc lộ trong tài liệu sáng chế (sau đây, được gọi tắt là "PTL") 1 sẽ được mô tả sử dụng Fig.1 và Fig.2.

Fig.1 là sơ đồ khối minh họa cấu hình của thiết bị mã hóa âm thanh/giọng nói được bộc lộ trong PTL 1. Tín hiệu đầu vào được lấy mẫu ở tần số 48 kHz được nhập vào bộ phát hiện chuyển tiếp 11 và bộ phận biến đổi 12 của thiết bị mã hóa âm thanh/giọng nói.

Bộ phát hiện chuyển tiếp 11 phát hiện, từ tín hiệu đầu vào, khung chuyển tiếp tương ứng với biên mở đầu hoặc biên kết thúc của giọng nói hoặc khung tĩnh tương ứng với phần khác của giọng nói, và bộ phận biến đổi 12 áp dụng, cho khung của tín hiệu đầu vào, sự biến đổi phân giải tần số cao hoặc biến đổi phân giải tần số thấp phụ thuộc vào việc khung được phân tách bởi bộ phát hiện chuyển tiếp 11 là khung chuyển tiếp hay khung tĩnh và thu được hệ số phổ (hoặc hệ số biến đổi).

Bộ phận ước lượng mức chuẩn 13 tách hệ số phổ thu được trong bộ phận biến đổi 12 thành các băng tần có các dải tần khác nhau. Bộ phận ước lượng mức chuẩn 13 ước lượng mức chuẩn (hoặc năng lượng) của mỗi băng tần được tách ra.

Bộ phận lượng tử hóa mức chuẩn 14 xác định đường bao phổ tạo nên các mức chuẩn của tất cả các băng tần dựa vào mức chuẩn của mỗi băng tần được ước lượng bởi bộ phận ước lượng mức chuẩn 13 và lượng tử hóa đường bao phổ xác định được.

Bộ phận chuẩn hóa phổ 15 chuẩn hóa hệ số phổ thu được bởi bộ phận biến đổi 12 theo mức chuẩn được lượng tử hóa bởi bộ phận lượng tử hóa mức chuẩn 14.

Bộ phận điều chỉnh mức chuẩn 16 điều chỉnh mức chuẩn được lượng tử hóa bởi bộ phận lượng tử hóa mức chuẩn 14 dựa trên kỹ thuật lấy trọng số phổ thích ứng.

Bộ phận cấp phát bit 17 cấp phát các bit có sẵn cho mỗi băng tần trong khung sử dụng mức chuẩn lượng tử hóa được điều chỉnh bởi bộ phận điều chỉnh mức chuẩn 16.

Bộ phận mã hóa vector - lưới 18 thực hiện mã hóa vector - lưới trên hệ số phổ được chuẩn hóa bởi bộ phận chuẩn hóa phổ 15 sử dụng các bit được cấp phát cho mỗi băng tần bởi bộ phận cấp phát bit 17.

Bộ phận điều chỉnh mức nhiễu 19 ước lượng mức hệ số phổ trước khi mã hóa trong bộ phận mã hóa vector - lưới 18 và mã hóa mức ước lượng được. Chỉ số điều chỉnh mức nhiễu thu được theo cách này.

Bộ ghép kênh 20 ghép kênh cấu hình khung của tín hiệu đầu vào thu được bởi bộ phận biến đổi 12, đó là, chỉ báo tín hiệu chuyển tiếp chỉ ra khung là khung tĩnh hay khung chuyển tiếp, mức chuẩn được lượng tử hóa bởi bộ phận lượng tử hóa mức chuẩn 14, vector mã hóa lưới thu được bởi bộ phận mã hóa vector - lưới 18 và chỉ số điều chỉnh mức nhiễu thu được bởi bộ phận điều chỉnh mức nhiễu 19, và tạo ra dòng bit và chuyển dòng bit đến thiết bị giải mã âm thanh/giọng nói.

Fig.2 là sơ đồ khối minh họa cấu hình của thiết bị giải mã âm thanh/giọng nói được bộc lộ trong PTL 1. Thiết bị giải mã âm thanh/giọng nói thu dòng bit được chuyển từ thiết bị mã hóa âm thanh/giọng nói và bộ tách kênh 21 tách kênh dòng bit.

Bộ phận giải lượng tử mức chuẩn 22 giải lượng tử mức chuẩn đã lượng tử hóa, thu đường bao phổ tạo nên các mức chuẩn của tất cả các băng tần, và bộ phận điều chỉnh mức chuẩn 23 điều chỉnh mức chuẩn đã được giải lượng tử bởi bộ phận giải lượng tử mức chuẩn 22 dựa trên kỹ thuật lấy trọng số phổ thích ứng.

Bộ phận cấp phát bit 24 cấp phát các bit có sẵn cho mỗi băng tần trong khung sử dụng các mức chuẩn được điều chỉnh bởi bộ phận điều chỉnh mức chuẩn 23. Đó là, bộ phận cấp phát bit 24 tính lại sự cấp phát bit cần thiết để giải mã mã vector - lưới của hệ số phổ đã chuẩn hóa.

Bộ phận giải mã lưới 25 giải mã chỉ báo tín hiệu chuyển tiếp, giải mã vector mã hóa

lưới dựa trên cấu hình khung được chỉ ra bởi chỉ báo tín hiệu chuyển tiếp được giải mã và các bit được cấp phát bởi bộ phận cấp phát bit 24 và thu hệ số phổ.

Bộ tạo điền đầy phổ 26 tái lập hệ số phổ tần số thấp mà không có bit được cấp phát đến sử dụng bảng mã được tạo ra dựa trên hệ số phổ được giải mã bởi bộ phận giải mã lưới 25. Bộ tạo điền đầy phổ 26 điều chỉnh mức hệ số phổ được tái lập sử dụng chỉ số điều chỉnh mức nhiễu. Ngoài ra, bộ tạo điền đầy phổ 26 tái lập hệ số phổ chưa được mã hóa ở tần số cao sử dụng hệ số phổ đã mã hóa ở tần số thấp.

Bộ cộng 27 cộng hệ số phổ đã giải mã và hệ số phổ được tái lập, và tạo ra hệ số phổ chuẩn hóa.

Bộ phận định hình đường bao 28 áp dụng đường bao phổ được giải lượng tử bởi bộ phận giải lượng tử mức chuẩn 22 cho hệ số phổ chuẩn hóa được tạo ra bởi bộ cộng 27 và tạo ra hệ số phổ toàn băng tần.

Bộ phận biến đổi ngược 29 áp dụng biến đổi ngược chẳng hạn như biến đổi cosin rời rạc hiệu chỉnh ngược (inverse modified discrete cosine transform - IMDCT) cho hệ số phổ toàn băng tần được tạo ra bởi bộ phận định hình đường bao 28 để biến đổi thành tín hiệu miền thời gian. Ở đây, sự biến đổi ngược với sự phân giải tần số cao được áp dụng cho trường hợp với khung tĩnh và sự biến đổi ngược với sự phân giải tần số thấp được áp dụng cho trường hợp khung chuyển tiếp.

Trong G.719, các hệ số phổ được tách thành các nhóm phổ. Mỗi nhóm phổ được tách thành các băng tần có các vector con có độ dài bằng nhau như được thể hiện trên Fig.3. Các vector con có độ dài khác nhau đối với mỗi nhóm khác nhau và độ dài này tăng khi tần số tăng. Liên quan đến độ phân giải của phép biến đổi, sự phân giải tần số cao hơn được sử dụng cho các tần số thấp, trong khi sự phân giải tần số thấp hơn được sử dụng cho các tần số cao. Như được mô tả trong G.719, việc ghép nhóm cho phép sử dụng hiệu quả một lượng bit có sẵn trong khi mã hóa.

Trong G.719, mô hình cấp phát bit giống như trong thiết bị mã hóa và thiết bị giải mã. Ở đây, mô hình cấp phát bit sẽ được mô tả sử dụng Fig.4.

Như được thể hiện trên Fig.4, ở bước (sau đây được viết tắt là “ST”) 31, các mức chuẩn đã lượng tử hóa được điều chỉnh trước khi cấp phát bit để điều chỉnh trọng số cảm nhận thính giác và các hiệu ứng che chắn.

Trong bước ST32, các băng con có mức chuẩn lớn nhất được nhận biết từ trong số tất cả các băng con và trong bước ST33, một bit được cấp phát cho mỗi hệ số phổ đối với các băng con có mức chuẩn lớn nhất. Tức là, càng nhiều bit là các hệ số phổ được cấp phát.

Trong bước ST34, các mức chuẩn được giảm theo các bit được cấp phát, và trong bước ST35, xác định được là số lượng bit có thể cấp phát còn lại là 8 hoặc nhiều hơn. Khi số bit có thể cấp phát còn lại là 8 hoặc nhiều hơn, tiến trình quay trở lại bước ST32 và khi số bit có thể cấp phát còn lại ít hơn 8, quy trình cấp phát bit chấm dứt.

Do đó, trong mô hình cấp phát bit, các bit có sẵn trong khung được cấp phát giữa các băng con sử dụng các mức chuẩn lượng tử hóa đã được điều chỉnh. Các hệ số phổ đã chuẩn hóa được mã hóa bằng kỹ thuật mã hóa vector-lưới sử dụng các bit được cấp phát cho mỗi băng con.

Danh sách tài liệu tham chiếu

Tài liệu phi sáng chế :

NPL 1: ITU-T Recommendation G.719, “Low-complexity full-band audio coding for high-quality conversational applications (Mã hóa âm toàn dải tần ít phức tạp cho các ứng dụng hội thoại chất lượng cao)”, ITU-T, 2009.

Tuy nhiên, mô hình cấp phát bit ở trên không xem xét đến các đặc tính của tín hiệu

đầu vào khi ghép nhóm các dải phổ, và do đó có vấn đề là không có khả năng thực hiện cấp phát bit hiệu quả và chất lượng âm thanh không thể được cải thiện thêm.

Bản chất kỹ thuật của sáng chế

Mục đích của sáng chế là đề xuất thiết bị mã hóa âm thanh/giọng nói, thiết bị giải mã âm thanh/giọng nói, phương pháp mã hóa âm thanh/giọng nói và phương pháp giải mã âm thanh/giọng nói có khả năng thực hiện cấp phát bit hiệu quả và cải thiện chất lượng âm thanh.

Thiết bị mã hóa âm thanh/giọng nói theo sáng chế bao gồm: bộ phận biến đổi thực hiện biến đổi tín hiệu đầu vào từ miền thời gian thành miền tần số; bộ phận ước lượng thực hiện ước lượng đường bao năng lượng biểu diễn mức năng lượng đối với mỗi trong số các băng con thu được nhờ phân tách phổ tần số của tín hiệu đầu vào; bộ phận lượng tử hóa thực hiện lượng tử hóa các đường bao năng lượng; bộ phận xác định nhóm thực hiện ghép nhóm các đường bao năng lượng đã lượng tử hóa thành các nhóm; bộ phận cấp phát bit thứ nhất thực hiện cấp phát các bit cho các nhóm; bộ phận cấp phát bit thứ hai thực hiện cấp phát các bit đã được cấp phát cho các nhóm cho các băng con trên cơ sở từng nhóm một; và bộ phận mã hóa mã hóa phổ tần số sử dụng các bit đã được cấp phát cho các băng con.

Thiết bị giải mã âm thanh/giọng nói theo sáng chế bao gồm: bộ phận giải lượng tử thực hiện giải lượng tử đường bao phổ đã lượng tử hóa; bộ phận xác định nhóm thực hiện ghép nhóm các đường bao phổ đã lượng tử hóa thành các nhóm; bộ phận cấp phát bit thứ nhất thực hiện cấp phát các bit cho các nhóm; bộ phận cấp phát bit thứ hai thực hiện cấp phát các bit đã được cấp phát cho các nhóm cho các băng con trên cơ sở từng nhóm một; bộ phận giải mã thực hiện giải mã phổ tần số của tín hiệu âm thanh/giọng nói sử dụng các bit được cấp phát cho các băng con; bộ phận định hình đường bao thực hiện áp dụng đường bao phổ đã giải lượng tử cho phổ tần số đã giải mã và tái lập phổ đã giải mã; và bộ

phân biến đổi ngược thực hiện biến đổi ngược phổ đã giải mã từ miền tần số thành miền thời gian.

Phương pháp mã hóa âm thanh/giọng nói theo sáng chế bao gồm: biến đổi tín hiệu đầu vào từ miền thời gian thành miền tần số; ước lượng đường bao năng lượng biểu diễn mức năng lượng cho mỗi băng con trong số các băng con thu được nhờ phân tách phổ tần số của tín hiệu đầu vào; lượng tử hóa các đường bao năng lượng; ghép nhóm các đường bao năng lượng đã lượng tử hóa thành nhiều nhóm; cấp phát các bit cho các nhóm; cấp phát các bit đã được cấp phát cho các nhóm cho các băng con trên cơ sở từng nhóm một; và mã hóa phổ tần số sử dụng các bit đã được cấp phát cho các băng con.

Phương pháp giải mã âm thanh/giọng nói theo sáng chế bao gồm: giải lượng tử đường bao phổ đã lượng tử hóa; ghép nhóm đường bao phổ đã lượng tử hóa thành nhiều nhóm; cấp phát các bit cho các nhóm; cấp phát các bit đã được cấp phát cho các nhóm cho các băng con trên cơ sở từng nhóm một; giải mã phổ tần số của tín hiệu âm thanh/giọng nói sử dụng các bit được cấp phát cho các băng con; áp dụng đường bao phổ đã giải lượng tử cho phổ tần số đã giải mã và tái lập phổ đã giải mã; và biến đổi ngược phổ đã giải mã từ miền tần số thành miền thời gian.

Hiệu quả đạt được của sáng chế

Theo sáng chế, có thể thực hiện cấp phát bit hiệu quả và cải thiện chất lượng âm thanh.

Mô tả vắn tắt các hình vẽ

Fig.1 là sơ đồ khối minh họa cấu hình của thiết bị mã hóa âm thanh/giọng nói được bộc lộ trong PTL 1;

Fig.2 là sơ đồ khối minh họa cấu hình của thiết bị giải mã âm thanh/giọng nói được bộc lộ trong PTL 1;

Fig.3 là sơ đồ minh họa cách ghép nhóm của các hệ số phổ ở chế độ tĩnh được bộc lộ trong PTL 1;

Fig.4 là lưu đồ minh họa mô hình cấp phát bit được bộc lộ trong PTL 1;

Fig.5 là sơ đồ khối minh họa cấu hình của thiết bị mã hóa âm thanh/giọng nói theo phương án của sáng chế;

Fig.6 là sơ đồ khối minh họa cấu hình của thiết bị giải mã âm thanh/giọng nói theo phương án của sáng chế;

Fig.7 là sơ đồ khối minh họa cấu hình bên trong của bộ phận cấp phát bit được thể hiện trên Fig.5;

Từ Fig.8A đến Fig.8C là các sơ đồ được cung cấp để mô tả phương pháp ghép nhóm theo phương án của sáng chế; và

Fig.9 là sơ đồ minh họa phương sai mức chuẩn.

Mô tả chi tiết các phương án thực hiện sáng chế

Sau đây, các phương án của sáng chế sẽ được mô tả chi tiết dựa vào các hình vẽ kèm theo.

Fig.5 là sơ đồ khối minh họa cấu hình của thiết bị mã hóa âm thanh/giọng nói 100 theo phương án của sáng chế. Tín hiệu đầu vào được lấy mẫu ở tần số 48 kHz được nhập vào bộ phát hiện chuyển tiếp 101 và bộ phận biến đổi 102 của thiết bị mã hóa âm thanh/giọng nói 100.

Bộ phát hiện chuyển tiếp 101 phát hiện, từ tín hiệu đầu vào, hoặc là khung chuyển tiếp tương ứng với biên bắt đầu hoặc biên kết thúc của giọng nói hoặc là khung tĩnh tương ứng với phần còn lại của giọng nói, và xuất kết quả phát hiện tới bộ phận biến đổi 102. Bộ

phần biến đổi 102 áp dụng, cho khung tín hiệu đầu vào, sự biến đổi phân giải tần số cao hoặc sự biến đổi phân giải tần số thấp phụ thuộc vào việc kết quả phát hiện được xuất từ bộ phát hiện chuyển tiếp 101 là khung chuyển tiếp hay khung tĩnh, và thu hệ số phổ (hoặc hệ số biến đổi) và xuất hệ số phổ tới bộ phận ước lượng mức chuẩn 103 và bộ phận chuẩn hóa phổ 105. Bộ phận biến đổi 102 xuất cấu hình khung là kết quả phát hiện được xuất từ bộ phát hiện chuyển tiếp 101, tức là, chỉ báo tín hiệu chuyển tiếp chỉ ra khung là khung tĩnh hay khung chuyển tiếp tới bộ ghép kênh 110.

Bộ phận ước lượng mức chuẩn 103 phân tách hệ số phổ được xuất từ bộ phận biến đổi 102 thành các băng tần có băng thông khác nhau và ước lượng mức chuẩn (hoặc năng lượng) của mỗi băng tần tách được. Bộ phận ước lượng mức chuẩn 103 xuất mức chuẩn ước lượng được của mỗi băng tần tới bộ phận lượng tử hóa mức chuẩn 104.

Bộ phận lượng tử hóa mức chuẩn 104 xác định đường bao phổ tạo nên các mức chuẩn của tất cả các băng tần dựa vào các mức chuẩn của các băng tần tương ứng được xuất từ bộ phận ước lượng mức chuẩn 103, lượng tử hóa đường bao phổ xác định được và xuất đường bao phổ đã lượng tử hóa tới bộ phận chuẩn hóa phổ 105 và bộ phận điều chỉnh mức chuẩn 106.

Bộ phận chuẩn hóa phổ 105 chuẩn hóa hệ số phổ được xuất từ bộ phận biến đổi 102 theo đường bao phổ đã lượng tử hóa được xuất từ bộ phận lượng tử hóa mức chuẩn 104 và xuất hệ số phổ đã chuẩn hóa tới bộ phận mã hóa vectơ - lưới 108.

Bộ phận điều chỉnh mức chuẩn 106 điều chỉnh đường bao phổ đã lượng tử hóa được xuất từ bộ phận lượng tử hóa mức chuẩn 104 dựa trên kỹ thuật lấy trọng số phổ thích ứng và xuất đường bao phổ đã lượng tử hóa đã điều chỉnh tới bộ phận cấp phát bit 107.

Bộ phận cấp phát bit 107 cấp phát các bit có sẵn cho mỗi băng tần trong khung sử dụng đường bao phổ đã lượng tử hóa đã được điều chỉnh được xuất từ bộ phận điều chỉnh mức chuẩn 106 và xuất các bit đã cấp phát tới bộ phận mã hóa vectơ - lưới 108. Chi tiết về

bộ phận cấp phát bit 107 sẽ được mô tả sau.

Bộ phận mã hóa vector - lưới 108 thực hiện mã hóa vectơ-lưới trên hệ số phổ được chuẩn hóa bởi bộ phận chuẩn hóa phổ 105 sử dụng các bit được cấp phát cho mỗi băng tần trong bộ phận cấp phát bit 107 và xuất vectơ mã hóa lưới tới bộ phận điều chỉnh mức nhiễu 109 và bộ ghép kênh 110.

Bộ phận điều chỉnh mức nhiễu 109 ước lượng mức hệ số phổ trước khi mã hóa trong bộ phận mã hóa vectơ - lưới 108 và mã hóa mức ước lượng được. Chỉ số điều chỉnh mức nhiễu được xác định theo cách này. Chỉ số điều chỉnh mức nhiễu được xuất tới bộ ghép kênh 110.

Bộ ghép kênh 110 ghép kênh chỉ báo tín hiệu chuyển tiếp được xuất từ bộ phận biến đổi 102, đường bao phổ đã lượng tử hóa được xuất từ bộ phận lượng tử hóa mức chuẩn 104, vectơ mã hóa lưới được xuất từ bộ phận mã hóa vectơ - lưới 108 và chỉ số điều chỉnh mức nhiễu được xuất từ bộ phận điều chỉnh mức nhiễu 109, và tạo ra dòng bit và truyền dòng bit tới thiết bị giải mã âm thanh/giọng nói.

Fig.6 là sơ đồ khối minh họa cấu hình của thiết bị giải mã âm thanh/giọng nói 200 theo phương án của sáng chế. Dòng bit được truyền từ thiết bị mã hóa âm thanh/giọng nói 100 được thu bởi thiết bị giải mã âm thanh/giọng nói 200 và được giải ghép kênh bằng bộ tách kênh 201.

Bộ phận giải lượng tử mức chuẩn 202 giải lượng tử đường bao phổ đã lượng tử hóa (tức là, mức chuẩn) được xuất từ bộ ghép kênh, thu được đường bao phổ tạo nên các mức chuẩn của tất cả các băng tần và xuất đường bao phổ thu được tới bộ phận điều chỉnh mức chuẩn 203.

Bộ phận điều chỉnh mức chuẩn 203 điều chỉnh đường bao phổ được xuất từ bộ phận giải lượng tử mức chuẩn 202 dựa trên kỹ thuật lấy trọng số phổ thích ứng và xuất

đường bao phổ đã điều chỉnh tới bộ phận cấp phát bit 204.

Bộ phận cấp phát bit 204 cấp phát các bit có sẵn cho mỗi băng tần trong khung sử dụng đường bao phổ được xuất từ bộ phận điều chỉnh mức chuẩn 203. Tức là, bộ phận cấp phát bit 204 tính lại sự cấp phát bit cần thiết để giải mã vector lưới của hệ số phổ đã chuẩn hóa. Các bit được cấp phát được xuất tới bộ phận giải mã lưới 205.

Bộ phận giải mã lưới 205 giải mã vector mã hóa lưới được xuất từ bộ tách kênh 201 dựa trên cấu hình khung được chỉ ra bởi chỉ báo tín hiệu chuyển tiếp được xuất từ bộ tách kênh 201 và các bit được xuất từ bộ phận cấp phát bit 204 và thu hệ số phổ. Hệ số phổ được xuất tới bộ tạo điền đầy phổ 206 và bộ cộng 207.

Bộ tạo điền đầy phổ 206 tái lập hệ số phổ tần số thấp mà không có bit được cấp phát đến sử dụng bảng mã được tạo ra dựa trên hệ số phổ được xuất từ bộ phận giải mã lưới 205. Bộ tạo điền đầy phổ 206 điều chỉnh mức hệ số phổ đã tái lập sử dụng chỉ số điều chỉnh mức nhiễu được xuất từ bộ tách kênh 201. Ngoài ra, bộ tạo điền đầy phổ 206 tái lập hệ số phổ không phải trải qua sự mã hóa tần số cao sử dụng hệ số phổ đã mã hóa tần số thấp. Hệ số phổ tần số thấp đã điều chỉnh mức độ và hệ số phổ tần số cao tái lập được được xuất tới bộ cộng 207.

Bộ cộng 207 cộng hệ số phổ được xuất từ bộ phận giải mã lưới 205 và hệ số phổ được xuất từ bộ tạo điền đầy phổ 206, tạo ra hệ số phổ đã chuẩn hóa và xuất hệ số phổ đã chuẩn hóa tới bộ phận định hình đường bao 208.

Bộ phận định hình đường bao 208 áp dụng đường bao phổ được xuất từ bộ phận giải lượng tử mức chuẩn 202 cho hệ số phổ đã chuẩn hóa được tạo ra bởi bộ cộng 207 và tạo ra hệ số phổ toàn băng tần (tương ứng với phổ giải mã). Hệ số phổ toàn băng tần tạo được được xuất tới bộ phận biến đổi ngược 209.

Bộ phận biến đổi ngược 209 áp dụng biến đổi ngược chẳng hạn như biến đổi cosin

rời rạc hiệu chỉnh ngược (IMDCT) cho hệ số phổ toàn băng tần được xuất từ bộ phận định hình đường bao 208, biến đổi nó thành tín hiệu miền thời gian và xuất tín hiệu đầu ra. Ở đây, biến đổi ngược với sự phân giải tần số cao được áp dụng cho trường hợp khung tĩnh và biến đổi ngược với sự phân giải tần số thấp được áp dụng cho trường hợp khung chuyển tiếp.

Tiếp theo, chi tiết về bộ phận cấp phát bit 107 sẽ được mô tả sử dụng Fig.7. Lưu ý rằng bộ phận cấp phát bit 107 của thiết bị mã hóa âm thanh/giọng nói 100 giống như cấu hình của bộ phận cấp phát bit 204 của thiết bị giải mã âm thanh/giọng nói 200, và do đó chỉ có bộ phận cấp phát bit 107 sẽ được mô tả và phần mô tả bộ phận cấp phát bit 204 sẽ được lược bỏ.

Fig.7 là sơ đồ khối minh họa cấu hình bên trong của bộ phận cấp phát bit 107 được thể hiện trên Fig.5. Bộ phận nhận biết băng tần trội 301 nhận biết, dựa trên đường bao phổ đã lượng tử hóa được xuất từ bộ phận điều chỉnh mức chuẩn 106, băng tần trội là băng con trong đó giá trị hệ số mức chuẩn trong phổ có giá trị cao nhất nội tại, và xuất mỗi băng tần trội nhận biết được tới các bộ phận xác định nhóm trội từ 302-1 đến 302N. Bên cạnh chỉ định băng tần mà giá trị hệ số mức chuẩn có giá trị cao nhất nội tại, các ví dụ về phương pháp xác định băng tần trội có thể bao gồm chỉ định băng tần trong số tất cả các băng con trong đó giá trị hệ số mức chuẩn có giá trị cao nhất là băng tần trội hoặc chỉ định là băng tần trội, băng tần có giá trị hệ số mức chuẩn vượt quá ngưỡng định trước hoặc ngưỡng tính được từ các mức chuẩn của tất cả các băng con.

Các bộ phận xác định nhóm trội từ 302-1 đến 302N xác định thích ứng độ rộng nhóm theo các đặc tính của tín hiệu đầu vào được định tâm trên băng tần trội được xuất từ bộ phận nhận biết băng tần trội 301. Cụ thể hơn, độ rộng nhóm được xác định là độ rộng của nhóm các băng con được định tâm trên và cả hai biên của băng tần trội đến các băng con nơi đường dốc giảm dần của giá trị hệ số mức chuẩn dừng lại. Các bộ phận xác định nhóm trội từ 302-1 đến 302N xác định các băng tần có trong các độ rộng nhóm là các

nhóm trội và xuất các nhóm trội xác định được tới bộ phận xác định nhóm bình thường (không trội) 303. Lưu ý rằng khi băng tần trội được định vị ở biên (biên của tần số có sẵn, chỉ có một phía của đường dốc giảm dần có trong nhóm).

Bộ phận xác định nhóm bình thường 303 xác định các băng con liên tiếp được xuất từ các bộ phận xác định nhóm trội từ 302-1 đến 302N ngoài các nhóm trội là cá nhóm bình thường mà không có các băng tần trội. Bộ phận xác định nhóm bình thường 303 xuất các nhóm trội và các nhóm bình thường tới bộ phận tính năng lượng nhóm 304 và bộ phận tính phương sai mức chuẩn 306.

Bộ phận tính năng lượng nhóm 304 tính năng lượng đặc trưng của nhóm của các nhóm trội và các nhóm bình thường được xuất từ bộ phận xác định nhóm bình thường 303 và xuất năng lượng tính được tới bộ phận tính tổng năng lượng 305 và bộ phận phân phối bit cho nhóm 308. Năng lượng đặc trưng của nhóm được tính theo công thức 1 sau đây.

$$[1] \quad \text{Năng lượng}(G(k)) = \sum_{i=1}^M \text{Mức chuẩn}(i) \quad \dots (\text{Công thức 1})$$

Ở đây, k là chỉ số của mỗi nhóm, Năng lượng(G(k)) là năng lượng của nhóm k, i là chỉ số dải tần con của nhóm 2, M là tổng số băng con của nhóm k và Mức chuẩn(i) là giá trị hệ số mức chuẩn của băng con i của nhóm n.

Bộ phận tính tổng năng lượng 305 cộng gộp tất cả năng lượng đặc trưng của nhóm được xuất từ bộ phận tính năng lượng nhóm 304 và tính tổng năng lượng của tất cả các nhóm. Tổng năng lượng tính được được xuất tới bộ phận phân phối bit cho nhóm 308. Tổng năng lượng được tính theo công thức sau đây.

$$[2] \quad \text{Năng lượng}_{\text{tổng cộng}} = \sum_{k=1}^N \text{Năng lượng}(G(k)) \quad \dots (\text{Công thức 2})$$

Ở đây, Năng lượng_{tổng cộng} là tổng năng lượng của tất cả các nhóm, N là tổng số lượng nhóm trong phổ, k là chỉ số của mỗi nhóm và Năng lượng(G(k)) là năng lượng của

nhóm k.

Bộ phận tính phương sai mức chuẩn 306 tính phương sai mức chuẩn đặc trưng của nhóm đối với các nhóm trội và các nhóm bình thường được xuất từ bộ phận xác định nhóm bình thường 303, và xuất phương sai mức chuẩn tới bộ phận tính tổng phương sai mức chuẩn 307 và bộ phận phân phối bit cho nhóm 308. Phương sai mức chuẩn đẳng trung của nhóm được tính theo công thức 3 sau đây.

$$[3] \text{ Mức chuẩn}_{\text{phương sai}}(G(k)) = \text{Mức chuẩn}_{\text{lớn nhất}}(G(k)) - \text{Mức chuẩn}_{\text{nhỏ nhất}}(G(k))$$

... (Công thức 3)

Ở đây, k là chỉ số của mỗi nhóm, Mức chuẩn_{phương sai}(G(k)) là phương sai mức chuẩn của nhóm k, Mức chuẩn_{lớn nhất}(G(k)) là giá trị hệ số mức chuẩn lớn nhất của nhóm k, và Mức chuẩn_{nhỏ nhất}(G(k)) là giá trị hệ số mức chuẩn nhỏ nhất của nhóm k.

Bộ phận tính tổng phương sai mức chuẩn 307 tính tổng phương sai mức chuẩn của tất cả các nhóm dựa trên phương sai mức chuẩn đặc trưng của nhóm được xuất từ bộ phận tính phương sai mức chuẩn 306. Tổng phương sai mức chuẩn tính được được xuất tới bộ phận phân phối bit cho nhóm 308. Tổng phương sai mức chuẩn được tính theo công thức 4 sau đây.

$$[4] \text{ Mức chuẩn}_{\text{tổng phương sai}} = \sum_{i=1}^N \text{Mức chuẩn}_{\text{phương sai}}(G(k)) \quad \dots \text{ (Công thức 4)}$$

Ở đây, Mức chuẩn_{tổng phương sai} là tổng phương sai mức chuẩn của tất cả các nhóm, N là tổng số nhóm trong phổ, k là chỉ số của mỗi nhóm, và Mức chuẩn_{phương sai}(G(k)) là phương sai mức chuẩn của nhóm k.

Bộ phận phân phối bit cho nhóm 308 (tương ứng với bộ phận cấp phát bit thứ nhất) phân phối các bit trên cơ sở từng nhóm một dựa trên năng lượng đặc trưng của nhóm được xuất từ bộ phận tính năng lượng nhóm 304, tổng năng lượng của tất cả các nhóm được

xuất từ bộ phận tính tổng năng lượng 305, phương sai mức chuẩn đặc trưng của nhóm được xuất từ bộ phận tính phương sai mức chuẩn 306 và tổng phương sai mức chuẩn của tất cả các nhóm được xuất từ bộ phận tính tổng phương sai mức chuẩn 307, và xuất các bit được phân phối trên cơ sở từng nhóm một tới bộ phận phân phối bit cho băng con 309. Các bit được phân phối trên cơ sở từng nhóm một được tính theo công thức sau đây 5.

$$[5] \quad \text{Bit}(G(k)) = \text{Bit}_{\text{tổng}} \times \left(\text{tỷ số } 1 \times \frac{\text{Năng lượng}(G(k))}{\text{Năng lượng}_{\text{tổng}}} + (1 - \text{tỷ số } 1) \times \frac{\text{Mức chuẩn}_{\text{phương sai}}(G(k))}{\text{Mức chuẩn}_{\text{tổng phương sai}}} \right)$$

... (Công thức 5)

Ở đây, k là chỉ số của mỗi nhóm, $\text{Bit}(G(k))$ là số bit được phân phối cho nhóm k , $\text{Bit}_{\text{tổng}}$ là tổng số bit có sẵn, tỷ số 1 là tỷ lệ bit được cấp phát theo năng lượng, $\text{Năng lượng}(G(k))$ là năng lượng của nhóm k , $\text{Năng lượng}_{\text{tổng}}$ là tổng năng lượng của tất cả các nhóm, và $\text{Mức chuẩn}_{\text{phương sai}}(G(k))$ là phương sai mức chuẩn của nhóm k .

Ngoài ra, tỷ số 1 trong công thức 5 ở trên có giá trị nằm trong đoạn $[0, 1]$ và điều chỉnh tỷ lệ bit được cấp phát theo năng lượng hoặc phương sai mức chuẩn. Giá trị tỷ số 1 càng lớn thì số bit được cấp phát theo năng lượng càng nhiều, và trong trường hợp cực hạn, nếu giá trị này bằng 1, tất cả các bit được cấp phát theo năng lượng. Giá trị tỷ số 1 càng nhỏ thì số bit được cấp phát theo phương sai mức chuẩn càng nhiều, và trong trường hợp cực hạn, nếu giá trị này bằng 0, tất cả các bit được cấp phát theo phương sai mức chuẩn.

Bằng cách phân phối các bit trên cơ sở từng nhóm một như được mô tả ở trên, bộ phận phân phối bit cho nhóm 308 có thể phân phối nhiều bit hơn cho nhóm trội và phân phối ít bit hơn cho các nhóm bình thường.

Do đó, bộ phận phân phối bit cho nhóm 308 có thể xác định độ quan trọng cảm nhận của mỗi nhóm theo năng lượng và phương sai mức chuẩn và làm nổi bật các nhóm trội hơn. Phương sai mức chuẩn phù hợp với lý thuyết che chắn và có thể xác định độ

quan trọng cảm nhận chính xác hơn.

Bộ phận phân phối bit cho băng con 309 (tương ứng với bộ phận cấp phát bit thứ hai) phân phối các bit cho các băng con trong mỗi nhóm dựa trên các bit đặc trưng của nhóm được xuất từ bộ phận phân phối bit cho nhóm 308 và xuất các bit được cấp phát cho các băng con đặc trưng của nhóm tới bộ phận mã hóa vector - lưới 108 dưới dạng kết quả cấp phát bit. Ở đây, có nhiều bit hơn được phân phối cho các băng con quan trọng khi cảm nhận và có ít bit hơn được phân phối cho các băng con ít quan trọng hơn. Số bit được phân phối cho mỗi băng con trong nhóm được tính theo công thức sau đây 6.

$$[6] \quad \text{Bit}_{G(k)\text{dải tần con}(i)} = \text{Bit}(G(k)) \times \frac{\text{Mức chuẩn}(i)}{\text{Năng lượng}(G(k))} \quad \dots \text{(Công thức 6)}$$

Ở đây, $\text{Bit}_{G(k)\text{băng con}(i)}$ là bit được cấp phát cho băng con i của nhóm k , i là chỉ số của nhóm k , $\text{Bit}_{G(k)}$ là bit được cấp phát cho nhóm k , $\text{Năng lượng}(G(k))$ là năng lượng của nhóm k , và $\text{Mức chuẩn}(i)$ là giá trị hệ số mức chuẩn của băng con i của nhóm k .

Tiếp theo, phương pháp ghép nhóm sẽ được mô tả sử dụng các hình vẽ từ Fig.8A đến Fig.8C. Giả thiết rằng đường bao phổ đã lượng tử hóa được thể hiện trên Fig.8A được nhập vào bộ phận nhận biết băng tần trội 301. Bộ phận nhận biết băng tần trội 301 nhận biết các băng tần trội 9 và 20 dựa trên đường bao phổ đã lượng tử hóa được nhập vào (tham khảo Fig.8B).

Các bộ phận tạo nhóm trội từ 302-1 đến 302-N xác định các băng con được tập trung trên và trên cả hai phía của các băng tần trội 9 và 20 đến các băng con nơi đường dốc giảm dần của giá trị hệ số mức chuẩn dừng lại là nhóm trội đồng nhất. Trong các ví dụ trên các hình vẽ từ Fig.8A đến Fig.8C, đối với băng tần trội 9, các băng con từ 6 đến 12 được xác định là nhóm trội (nhóm 2), trong khi với băng tần trội 20, các băng con từ 17 đến 22 được xác định là nhóm trội (nhóm 4) (xem Fig.8C).

Bộ phận xác định nhóm bình thường 303 xác định các băng tần tiếp theo ngoài các

nhóm trội là các nhóm bình thường mà không phải là các băng tần trội. Trong ví dụ trên các hình vẽ từ Fig.8A đến Fig.8C, các băng con từ 1 đến 5 (nhóm 1), các băng con từ 13 đến 16 (nhóm 3) và các băng con từ 23 đến 25 (nhóm 5) được xác định là các nhóm bình thường tương ứng (xem Fig.8C).

Kết quả là, đường bao phổ đã lượng tử hóa được tách thành năm nhóm, đó là, hai nhóm trội (các nhóm 2 và 4) và ba nhóm bình thường (các nhóm 1, 3 và 5).

Sử dụng phương pháp như thế có thể xác định thích ứng độ rộng nhóm theo các đặc điểm của tín hiệu đầu vào. Theo phương pháp này, thiết bị giải mã âm thanh/giọng nói cũng sử dụng các hệ số mức chuẩn đã lượng tử hóa có sẵn, và do đó thông tin bổ sung không cần phải được truyền tới thiết bị giải mã âm thanh/giọng nói.

Lưu ý rằng bộ phận tính phương sai mức chuẩn 306 tính phương sai mức chuẩn đặc trưng của nhóm. Trong các ví dụ trên các hình vẽ từ Fig.8A đến Fig.8C, phương sai mức chuẩn Năng lượng_{phương sai} (G(2)) trong nhóm 2 được thể hiện trên Fig.9 là mức quy chiếu.

Tiếp theo, độ quan trọng cảm nhận sẽ được mô tả. Phổ của tín hiệu âm thanh/giọng nói thường bao gồm nhiều đỉnh (núi) và các chỗ uốn lõm (thung lũng). Đỉnh được tạo nên bởi thành phần phổ được đặt ở tần số trội của tín hiệu âm thanh/giọng nói (thành phần âm trội). Đỉnh rất quan trọng khi cảm nhận. Độ quan trọng cảm nhận của đỉnh có thể được xác định nhờ sự chênh lệch giữa năng lượng của đỉnh và năng lượng của chỗ uốn lõm, tức là, nhờ phương sai mức chuẩn. Theo lý thuyết, khi đỉnh có năng lượng đủ lớn so với các băng tần bên cạnh, đỉnh cần được mã hóa với số bit vừa đủ nhưng nếu đỉnh được mã hóa với số bit không đủ, nhiều mã hóa có trong đó trở nên nổi bật, dẫn đến chất lượng âm thanh bị suy giảm. Mặt khác, chỗ uốn lõm không tạo nên thành phần âm thanh trội nào của tín hiệu âm thanh/giọng nói và không quan trọng khi cảm nhận.

Theo phương pháp ghép nhóm băng tần của phương án này, băng tần trội tương

ứng với đỉnh của phổ và việc ghép nhóm các băng tần có nghĩa là tách các đỉnh (các nhóm trội bao gồm các băng tần trội) ra khỏi các chỗ uốn lồi (các nhóm bình thường không phải là các băng tần trội).

Bộ phận phân phối bit cho nhóm 308 xác định độ quan trọng cảm nhận của đỉnh. Ngược lại với kỹ thuật G.719 trong đó độ quan trọng cảm nhận được xác định chỉ theo năng lượng, phương án này xác định độ quan trọng cảm nhận dựa trên cả năng lượng và sự phân bố mức chuẩn (năng lượng) và xác định các bit được phân phối cho mỗi nhóm dựa trên độ quan trọng cảm nhận xác định được.

Trong bộ phận phân phối bit cho băng con 309, khi phương sai mức chuẩn trong nhóm lớn, điều này có nghĩa là nhóm này là một trong số các đỉnh, đỉnh quan trọng hơn khi cảm nhận và hệ số mức chuẩn có giá trị cao nhất cần được mã hóa chính xác. Với lý do này, có nhiều bit hơn được phân phối cho mỗi năng con của đỉnh này. Mặt khác, khi phương sai mức chuẩn trong nhóm là rất nhỏ, điều này có nghĩa là nhóm này là một trong số các chỗ uốn lồi, và chỗ uốn lồi không quan trọng khi cảm nhận và không cần được mã hóa chính xác. Với lý do này, có ít bit hơn được phân phối cho mỗi băng con của nhóm này.

Do đó, phương án này nhận biết băng tần trội trong đó giá trị hệ số mức chuẩn trong phổ của tín hiệu âm thanh/giọng nói được nhập vào có giá trị cao nhất nội tại, ghép tất cả các băng con thành các nhóm trội chứa băng tần trội và các nhóm bình thường không chứa băng tần trội nào, phân phối các bit cho mỗi nhóm dựa trên năng lượng đặc trưng của nhóm và các phương sai mức chuẩn, và còn phân phối các bit đã được phân phối trên cơ sở từng nhóm một tới mỗi băng con theo tỷ lệ mức chuẩn trên năng lượng của mỗi nhóm. Theo cách này, có thể cấp phát nhiều bit hơn cho các nhóm quan trọng cảm nhận và các băng con và thực hiện phân phối bit hiệu quả. Kết quả là, chất lượng âm thanh có thể được cải thiện.

Lưu ý rằng hệ số mức chuẩn theo phương án này là năng lượng bằng con và cũng có thể được gọi là “đường bao năng lượng.”

Nội dung của đơn sáng chế Nhật Bản số 2012-272571 nộp ngày 13 tháng 12 năm 2012, bao gồm bản mô tả, các hình vẽ và tóm tắt được kết hợp toàn bộ ở đây nhằm mục đích tham khảo.

Khả năng ứng dụng trong công nghiệp

Thiết bị mã hóa âm thanh/giọng nói, thiết bị giải mã âm thanh/giọng nói, phương pháp mã hóa âm thanh/giọng nói và phương pháp giải mã âm thanh/giọng nói theo sáng chế có khả năng áp dụng cho thiết bị đầu cuối truyền thông vô tuyến, thiết bị trạm cố định truyền thông vô tuyến, thiết bị đầu cuối hội nghị qua điện thoại, thiết bị đầu cuối hội thảo qua video và thiết bị đầu cuối truyền giọng nói trên giao thức Internet (voice over Internet protocol (VoIP)) hoặc tương tự.

YÊU CẦU BẢO HỘ

1. Thiết bị mã hóa âm thanh hoặc giọng nói bao gồm:

bộ phận biến đổi (102) thực hiện biến đổi tín hiệu đầu vào từ miền thời gian thành miền tần số;

bộ phận ước lượng (103) thực hiện ước lượng đường bao năng lượng biểu diễn mức năng lượng cho mỗi băng con trong số các băng con đạt được bằng cách tách phổ tần số của tín hiệu đầu vào;

bộ phận lượng tử hóa (104) thực hiện lượng tử hóa các đường bao năng lượng để thu được các đường bao năng lượng đã được lượng tử hóa;

bộ phận xác định nhóm (302-1, 302-N, 303) thực hiện ghép nhóm các đường bao năng lượng đã lượng tử hóa thành các nhóm;

bộ phận cấp phát bit thứ nhất (308) thực hiện cấp phát các bit cho mỗi nhóm trong số các nhóm để thu được số lượng bit cụ thể cho mỗi nhóm trong số các nhóm;

bộ phận cấp phát bit thứ hai (309) thực hiện cấp phát, cho mỗi nhóm trong số các nhóm, số lượng bit cụ thể đã được cấp phát cho nhóm tương ứng trong các nhóm cho các băng con thuộc về nhóm tương ứng; và

bộ phận mã hóa (108) thực hiện mã hóa, cho mỗi băng con trong số các băng con, phổ tần số trong băng con tương ứng bằng cách sử dụng các bit được cấp phát cho băng con tương ứng.

2. Thiết bị mã hóa âm thanh hoặc giọng nói theo điểm 1, còn bao gồm bộ phận nhận biết băng tần trội (301) thực hiện nhận biết băng tần trội là băng tần trong đó đường bao năng lượng của phổ tần số có giá trị cao nhất nội tại, trong đó:

bộ phận xác định nhóm (302-1, 302-N, 303) xác định băng tần trội và các băng con trên cả hai phía của băng tần trội tạo ra đường dốc giảm dần của đường bao năng lượng là nhóm trội và xác định tiếp các băng con không phải băng tần trội là nhóm không trội.

3. Thiết bị mã hóa âm thanh hoặc giọng nói theo điểm 1, còn bao gồm:

bộ phận tính năng lượng (304) thực hiện tính năng lượng đặc trưng của nhóm; và

bộ phận tính phân bố (306) thực hiện tính sự phân bố đường bao năng lượng đặc trưng của nhóm, trong đó:

bộ phận cấp phát bit thứ nhất (308) cấp phát, dựa trên năng lượng đặc trưng của nhóm tính được và sự phân bố đường bao năng lượng đặc trưng của nhóm, nhiều bit hơn cho nhóm khi ít nhất một trong số năng lượng và sự phân bố đường bao năng lượng lớn hơn và cấp phát ít bit hơn cho nhóm khi ít nhất một trong số năng lượng và sự phân bố đường bao năng lượng nhỏ hơn.

4. Thiết bị mã hóa âm thanh hoặc giọng nói theo điểm 1, trong đó bộ phận cấp phát bit thứ hai (309) cấp phát nhiều bit hơn cho băng con có đường bao năng lượng lớn hơn và cấp phát ít bit hơn cho băng con có đường bao năng lượng nhỏ hơn.

5. Thiết bị mã hóa âm thanh hoặc giọng nói theo điểm 1, trong đó bộ phận cấp phát bit thứ hai (309) được tạo cấu hình để cấp phát nhiều bit tới băng con quan trọng hơn khi cảm nhận và ít bit cho băng con ít quan trọng khi cảm nhận.

6. Thiết bị mã hóa âm thanh hoặc giọng nói theo điểm 1, trong đó bộ phận cấp phát bit thứ hai (309) được tạo cấu hình để cấp phát nhiều bit cho các băng con trong nhóm có phương sai năng lượng cao hơn và cấp phát ít bit hơn cho các băng con trong nhóm có phương sai năng lượng thấp hơn.

7. Thiết bị mã hóa âm thanh hoặc giọng nói theo điểm 1, trong đó bộ phận cấp phát bit thứ hai (309) được tạo cấu hình để cấp phát nhiều bit hơn cho các băng con trong nhóm có đỉnh nằm trong phổ tần số và cấp phát ít bit hơn cho các băng con trong nhóm có chỗ uốn lồi nằm trong phổ tần số.

8. Thiết bị mã hóa âm thanh hoặc giọng nói theo điểm 1, trong đó bộ phận cấp phát bit thứ hai (309) được tạo cấu hình để hoạt động dựa trên biểu thức sau:

$$[6] \quad \text{Bit}_{G(k)\text{đãi tần con}(i)} = \text{Bit}(G(k)) \times \frac{\text{Mức chuẩn}(i)}{\text{Năng lượng}(G(k))} \quad \dots \text{ (Công thức 6)}$$

trong đó, $\text{Bit}_{G(k)\text{đãi tần con}(i)}$ là bit được cấp phát cho băng con i của nhóm k , i là chỉ số của nhóm k , $\text{Bit}(G(k))$ là bit được cấp phát cho nhóm k , $\text{Năng lượng}(G(k))$ là năng lượng của nhóm k , và $\text{Mức chuẩn}(i)$ là giá trị hệ số mức chuẩn của băng con i của nhóm k .

9. Thiết bị mã hóa âm thanh hoặc giọng nói theo điểm 1, trong đó bộ phận cấp phát bit thứ nhất (308) được tạo cấu hình để cấp phát nhiều bit hơn cho nhóm trội và ít bit hơn cho nhóm không trội.

10. Thiết bị mã hóa âm thanh hoặc giọng nói theo điểm 1, trong đó bộ phận cấp phát bit thứ nhất (308) được tạo cấu hình để cấp phát các bit trên cơ sở từng nhóm một dựa trên năng lượng cụ thể của mỗi nhóm, tổng năng lượng của tất cả các nhóm, phương sai năng lượng của nhóm cụ thể và phương sai năng lượng tổng cộng của tất cả các nhóm.

11. Thiết bị mã hóa âm thanh hoặc giọng nói theo điểm 1, trong đó bộ phận cấp phát bit thứ nhất (308) được tạo cấu hình để hoạt động dựa trên biểu thức sau:

$$[5] \quad \text{Bit}(G(k)) = \text{Bit}_{\text{tổng}} \times \left(\text{tỷ số } 1 \times \frac{\text{Năng lượng}(G(k))}{\text{Năng lượng}_{\text{tổng}}} + (1 - \text{tỷ số } 1) \times \frac{\text{Mức chuẩn}_{\text{phương sai}}(G(k))}{\text{Mức chuẩn}_{\text{tổng phương sai}}} \right)$$

trong đó, k là chỉ số của mỗi nhóm, $\text{Bit}(G(k))$ là số bit được phân phối cho nhóm k , $\text{Bit}_{\text{tổng}}$ là tổng số bit có sẵn, tỷ số 1 là tỷ lệ bit được cấp phát theo năng lượng, $\text{Năng lượng}(G(k))$ là năng lượng của nhóm k , $\text{Năng lượng}_{\text{tổng}}$ là tổng năng lượng của tất cả các nhóm, và $\text{Mức chuẩn}_{\text{phương sai}}(G(k))$ là phương sai mức chuẩn của nhóm k .

12. Thiết bị mã hóa âm thanh hoặc giọng nói theo điểm 11, trong đó giá trị của tỷ số 1 là

giữa 0 và 1.

13. Thiết bị mã hóa âm thanh hoặc giọng nói theo điểm 1, trong đó bộ phận cấp phát bit thứ nhất (308) được tạo cấu hình để xác định độ quan trọng cảm nhận của mỗi nhóm bằng cách sử dụng năng lượng và phương sai năng lượng của nhóm và để tăng cường nhóm trội.

14. Thiết bị mã hóa âm thanh hoặc giọng nói theo điểm 1, trong đó bộ phận cấp phát bit thứ nhất (308) được tạo cấu hình để xác định độ quan trọng cảm nhận của nhóm dựa trên năng lượng của nhóm và phân bố năng lượng và để xác định các bit sẽ được cấp phát cho mỗi nhóm dựa trên độ quan trọng cảm nhận đối với mỗi nhóm tương ứng.

15. Thiết bị mã hóa âm thanh hoặc giọng nói theo điểm 1, trong đó bộ phận xác định nhóm (302-1, 302-N, 303) được tạo cấu hình để xác định thích ứng độ rộng nhóm của các nhóm theo các đặc tính của tín hiệu đầu vào.

16. Thiết bị mã hóa âm thanh hoặc giọng nói theo điểm 1, trong đó bộ phận xác định nhóm (302-1, 302-N, 303) được tạo cấu hình để sử dụng các năng lượng băng con đã được lượng tử hóa.

17. Thiết bị mã hóa âm thanh hoặc giọng nói theo điểm 1, trong đó bộ phận xác định nhóm (302-1, 302-N, 303) được tạo cấu hình để tách các đỉnh của phổ tần số từ các chỗ uốn lồi của phổ tần số, trong đó đỉnh của phổ tần số được đặt trong nhóm trội và chỗ uốn lồi của phổ tần số được đặt trong nhóm không trội.

18. Thiết bị mã hóa âm thanh hoặc giọng nói theo điểm 1,

trong đó bộ phận xác định nhóm (302-1, 302-N, 303) được tạo cấu hình để xác

định các băng tần trội, trong đó các giá trị năng lượng băng con trong phổ tần số của tín hiệu đầu vào có các giá trị cao nhất nội tại, và để nhóm các băng con bao gồm các băng tần trội vào trong các nhóm trội và các băng con khác vào trong các nhóm không trội,

trong đó bộ phận cấp phát bit thứ nhất (308) được tạo cấu hình để cấp phát các bit cho nhóm tương ứng dựa trên năng lượng của nhóm tương ứng và phương sai năng lượng của nhóm tương ứng, và

trong đó, bộ phận cấp phát bit thứ hai (309) được tạo cấu hình để cấp phát các bit, được cấp phát trên cơ sở từng nhóm tới nhóm tương ứng, tới băng con tương ứng trong nhóm tương ứng theo tỷ lệ năng lượng của băng con tương ứng với năng lượng của nhóm tương ứng.

19. Thiết bị mã hóa âm thanh hoặc giọng nói theo điểm 1,

trong đó bộ phận cấp phát bit thứ nhất (308) được tạo cấu hình để cấp phát nhiều bit cho nhóm quan trọng hơn khi cảm nhận và ít bit cho nhóm ít quan trọng hơn khi cảm nhận, và

trong đó bộ phận cấp phát bit thứ hai (309) được tạo cấu hình để cấp phát nhiều bit cho băng con quan trọng hơn khi cảm nhận và ít bit hơn cho băng con ít quan trọng hơn khi cảm nhận.

20. Thiết bị giải mã âm thanh hoặc giọng nói bao gồm:

bộ phận giải lượng tử (202) thực hiện giải lượng tử đường bao phổ đã lượng tử hóa để thu được đường bao phổ đã được giải lượng tử;

bộ phận xác định nhóm (302-1, 302-N, 303) thực hiện ghép nhóm các đường bao phổ đã lượng tử hóa thành nhiều nhóm;

bộ phận cấp phát bit thứ nhất (308) thực hiện cấp phát các bit cho mỗi nhóm trong số các nhóm để thu được số lượng bit của nhóm cụ thể cho mỗi nhóm trong số các nhóm;

bộ phận cấp phát bit thứ hai (309) thực hiện cấp phát, cho mỗi nhóm trong số các nhóm, số lượng bit cụ thể cho mỗi nhóm được cấp phát cho nhóm tương ứng trong số các nhóm tới các băng con thuộc về nhóm tương ứng;

bộ phận giải mã (205) thực hiện giải mã, đối với mỗi băng con trong số các băng con, phổ tần số của tín hiệu âm thanh hoặc giọng nói bằng cách sử dụng các bit đã được cấp phát cho băng con tương ứng để thu được phổ tần số đã được giải mã;

bộ phận định hình đường bao (208) thực hiện áp dụng đường bao phổ đã giải lượng tử cho phổ tần số đã giải mã và tái lập phổ đã giải mã; và

bộ phận biến đổi ngược (209) thực hiện biến đổi ngược phổ đã giải mã từ miền tần số sang miền thời gian.

21. Thiết bị giải mã âm thanh hoặc giọng nói theo điểm 20, còn bao gồm bộ phận nhận biết băng tần trội (301) thực hiện nhận biết băng tần trội là băng con trong đó đường bao năng lượng của phổ tần số có giá trị cao nhất nội tại, trong đó:

bộ phận xác định nhóm (302-1, 302-N, 303) xác định băng tần trội và các băng con trên cả hai phía của băng tần trội tạo ra đường dốc giảm dần của đường bao năng lượng là các nhóm trội và xác định các băng con tiếp theo ngoài băng tần trội là các nhóm không trội.

22. Thiết bị giải mã âm thanh hoặc giọng nói theo điểm 20, còn bao gồm:

bộ phận tính năng lượng (304) thực hiện tính năng lượng đặc trưng của nhóm; và

bộ phận tính phân bố (306) thực hiện tính đường bao năng lượng đặc trưng của nhóm, trong đó:

bộ phận cấp phát bit thứ nhất (308) thực hiện cấp phát, dựa trên năng lượng đặc trưng của nhóm tính được và sự phân bố đường bao năng lượng đặc trưng của nhóm, nhiều bit hơn cho nhóm khi ít nhất một trong số năng lượng và sự phân bố đường bao

năng lượng lớn hơn và cấp phát ít bit hơn cho các nhóm khi ít nhất một trong số năng lượng và sự phân bố đường bao năng lượng nhỏ hơn.

23. Thiết bị giải mã âm thanh hoặc giọng nói theo điểm 20, trong đó bộ phận cấp phát bit thứ hai (309) cấp phát nhiều bit hơn cho băng con có đường bao năng lượng lớn hơn và cấp phát ít bit hơn cho băng con có đường bao năng lượng nhỏ hơn.

24. Phương pháp mã hóa âm thanh hoặc giọng nói bao gồm các bước:

biến đổi tín hiệu đầu vào từ miền thời gian thành miền tần số;

ước lượng đường bao năng lượng biểu diễn mức năng lượng của mỗi băng con trong số nhiều băng con thu được bằng cách phân tách phổ tần số của tín hiệu đầu vào;

lượng tử hóa các đường bao năng lượng để thu được đường bao năng lượng đã được lượng tử hóa;

ghép nhóm các đường bao năng lượng đã lượng tử hóa thành các nhóm, mỗi nhóm có nhiều tập ít nhất hai băng con;

cấp phát, cho mỗi nhóm trong số các nhóm, các bit cho mỗi nhóm trong số các nhóm để thu được số lượng bit cụ thể cho mỗi nhóm cho mỗi nhóm trong số các nhóm;

cấp phát, cho mỗi nhóm trong số các nhóm, số lượng bit cụ thể được cấp phát cho mỗi nhóm cho nhóm tương ứng trong số các nhóm tới các băng con thuộc về nhóm tương ứng; và

mã hoá, mỗi băng con trong số các băng con, phổ tần số có trong băng con tương ứng bằng cách sử dụng các bit đã được cấp phát cho băng con tương ứng.

25. Phương pháp giải mã âm thanh hoặc giọng nói bao gồm các bước:

giải lượng tử đường bao phổ đã lượng tử hóa để thu được đường bao phổ đã được giải lượng tử;

ghép nhóm đường bao phổ đã lượng tử hóa thành các nhóm;

cấp phát, cho mỗi nhóm trong số các nhóm, các bit cho mỗi nhóm trong số các

nhóm để thu được số lượng bit cụ thể cho mỗi nhóm cho mỗi nhóm trong số các nhóm;

cấp phát, cho mỗi nhóm trong số các nhóm, số lượng bit cụ thể được cấp phát cho mỗi nhóm cho nhóm tương ứng trong số các nhóm tới các băng con thuộc về nhóm tương ứng;

giải mã, cho mỗi băng con trong số các băng con, phổ tần số của tín hiệu âm thanh hoặc giọng nói bằng cách sử dụng các bit đã được cấp phát cho băng con tương ứng để thu được phổ tần đã được giải mã;

áp dụng đường bao phổ đã giải lượng tử cho phổ tần số đã giải mã và tái lập phổ giải mã; và

biến đổi ngược phổ giải mã từ miền tần số thành miền thời gian.

26. Thiết bị đầu cuối truyền thông vô tuyến bao gồm thiết bị mã hóa âm thanh hoặc giọng nói theo điểm 1, thiết bị giải mã âm thanh hoặc giọng nói theo điểm 20, phương pháp mã hóa âm thanh hoặc giọng nói theo điểm 24 hoặc phương pháp giải mã âm thanh hoặc giọng nói theo điểm 25.

27. Thiết bị trạm gốc truyền thông vô tuyến bao gồm thiết bị mã hóa âm thanh hoặc giọng nói theo điểm 1, thiết bị giải mã âm thanh hoặc giọng nói theo điểm 20, phương pháp mã hóa âm thanh hoặc giọng nói theo điểm 24 hoặc phương pháp giải mã âm thanh hoặc giọng nói theo điểm 25.

28. Thiết bị đầu cuối hội nghị từ xa bao gồm thiết bị mã hóa âm thanh hoặc giọng nói theo điểm 1, thiết bị giải mã âm thanh hoặc giọng nói theo điểm 20, phương pháp mã hóa âm thanh hoặc giọng nói theo điểm 24 hoặc phương pháp giải mã âm thanh hoặc giọng nói theo điểm 25.

29. Thiết bị đầu cuối hội nghị truyền hình bao gồm thiết bị mã hóa âm thanh hoặc giọng

nói theo điểm 1, thiết bị giải mã âm thanh hoặc giọng nói theo điểm 20, trong đó thiết bị này bị này sử dụng phương pháp mã hóa âm thanh hoặc giọng nói theo điểm 24 hoặc phương pháp giải mã âm thanh hoặc giọng nói theo điểm 25.

30. Thiết bị đầu cuối truyền giọng nói trên nền tảng giao thức Internet (VoIP) bao gồm thiết bị mã hóa âm thanh hoặc giọng nói theo điểm 1, thiết bị giải mã âm thanh hoặc giọng nói theo điểm 20, trong đó thiết bị này bị này sử dụng phương pháp mã hóa âm thanh hoặc giọng nói theo điểm 24 hoặc phương pháp giải mã âm thanh hoặc giọng nói theo điểm 25.

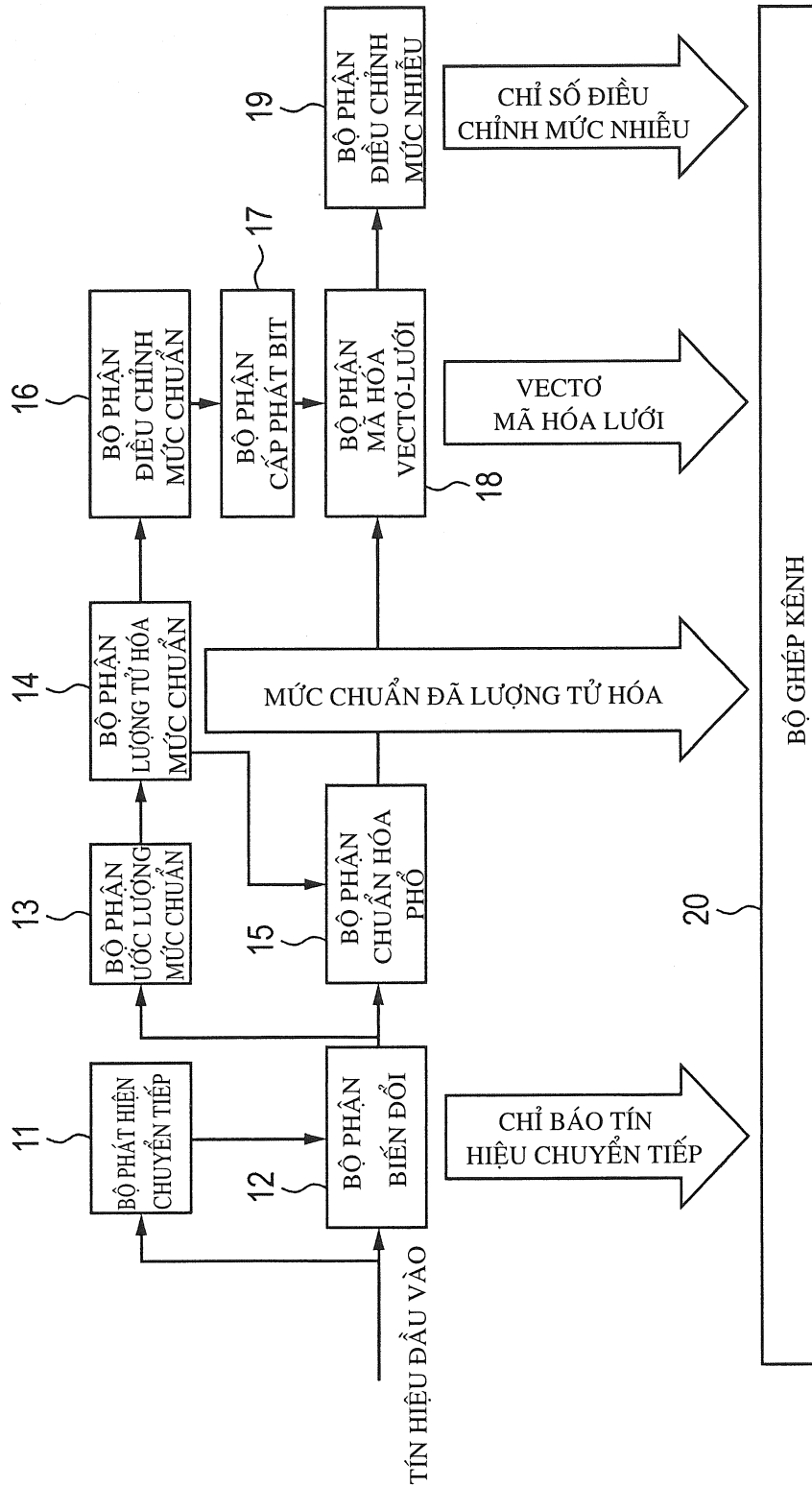


FIG. 1

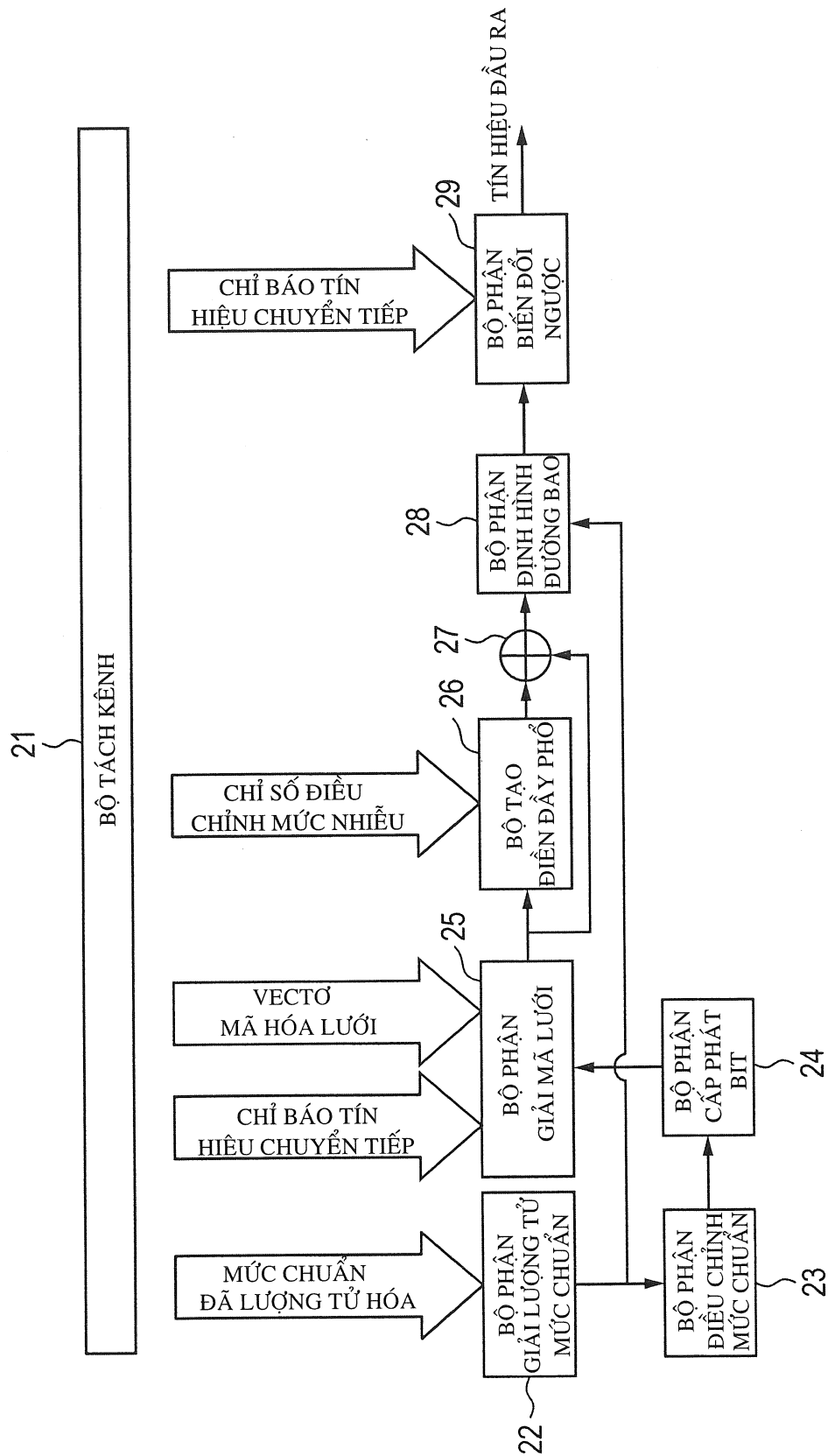


FIG. 2

NHÓM	ĐỘ DÀI CỦA VECTƠ CON	SỐ LƯỢNG VECTƠ CON	SỐ LƯỢNG HỆ SỐ	BĂNG THÔNG (Hz)	BẮT ĐẦU (Hz)	KẾT THÚC (Hz)
I	8	16	128	3 200	0	3 200
II	16	8	128	3 200	3 200	6 400
III	24	12	288	7 200	6 400	13 600
IV	32	8	256	6 400	13 600	20 000
TỔNG		44	800	20 000		

FIG. 3

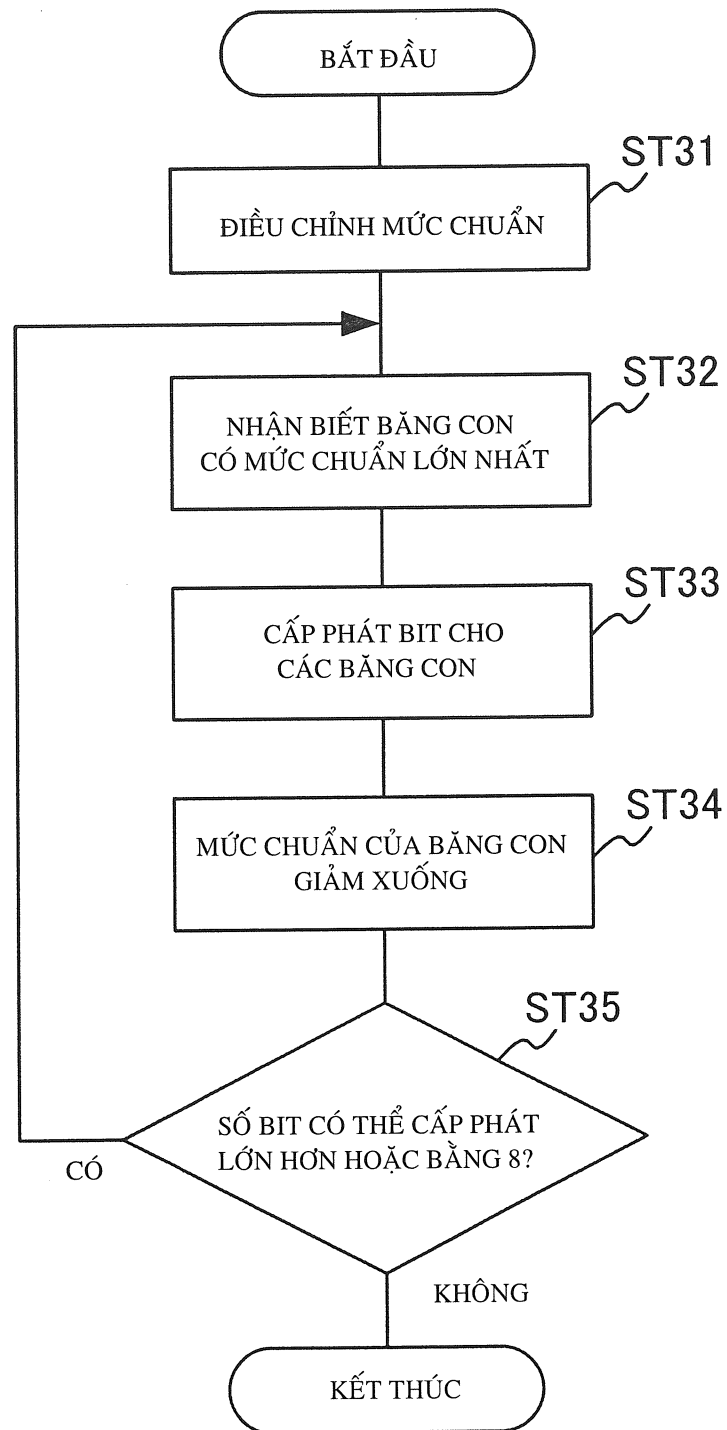


FIG. 4

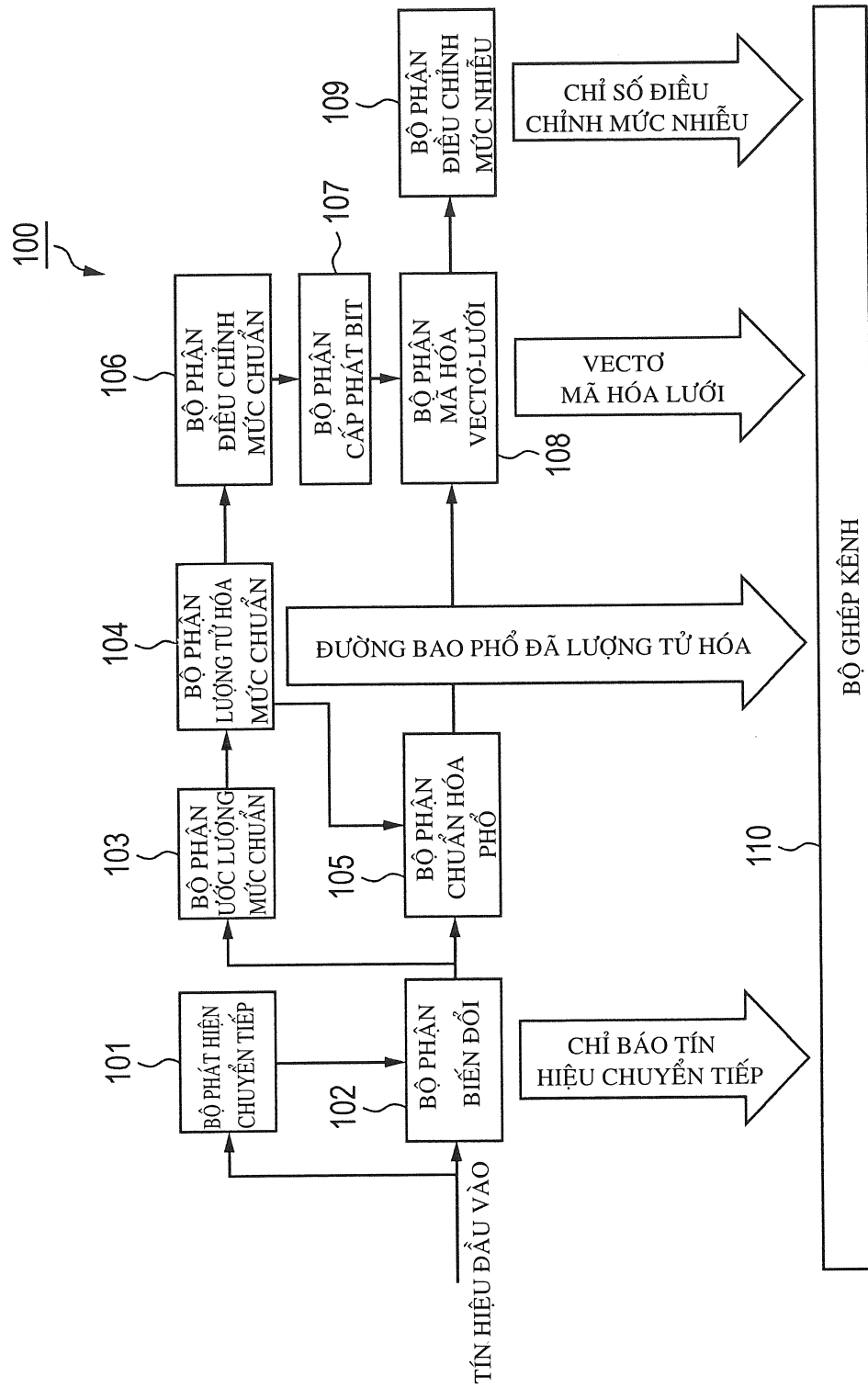


FIG. 5

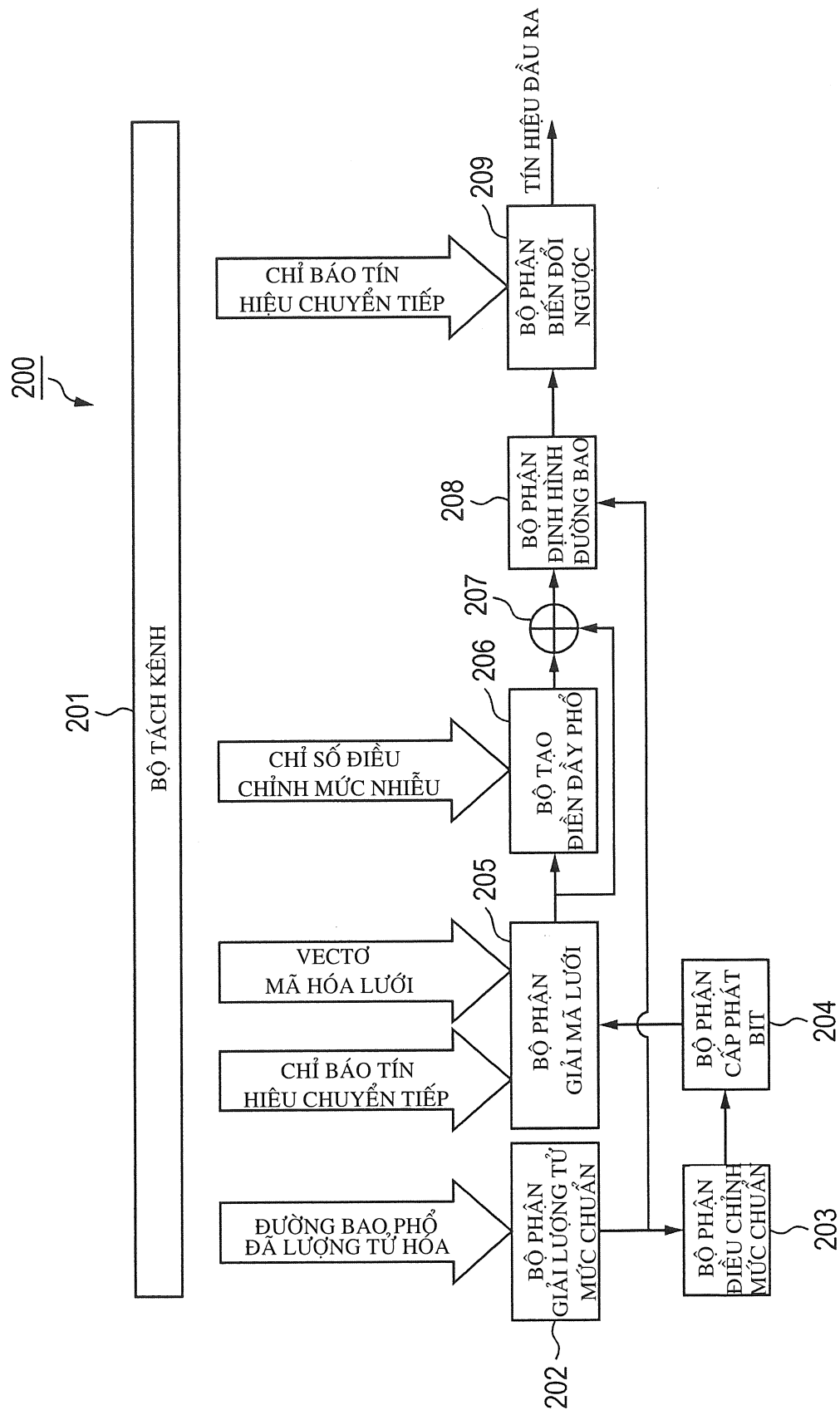


FIG. 6

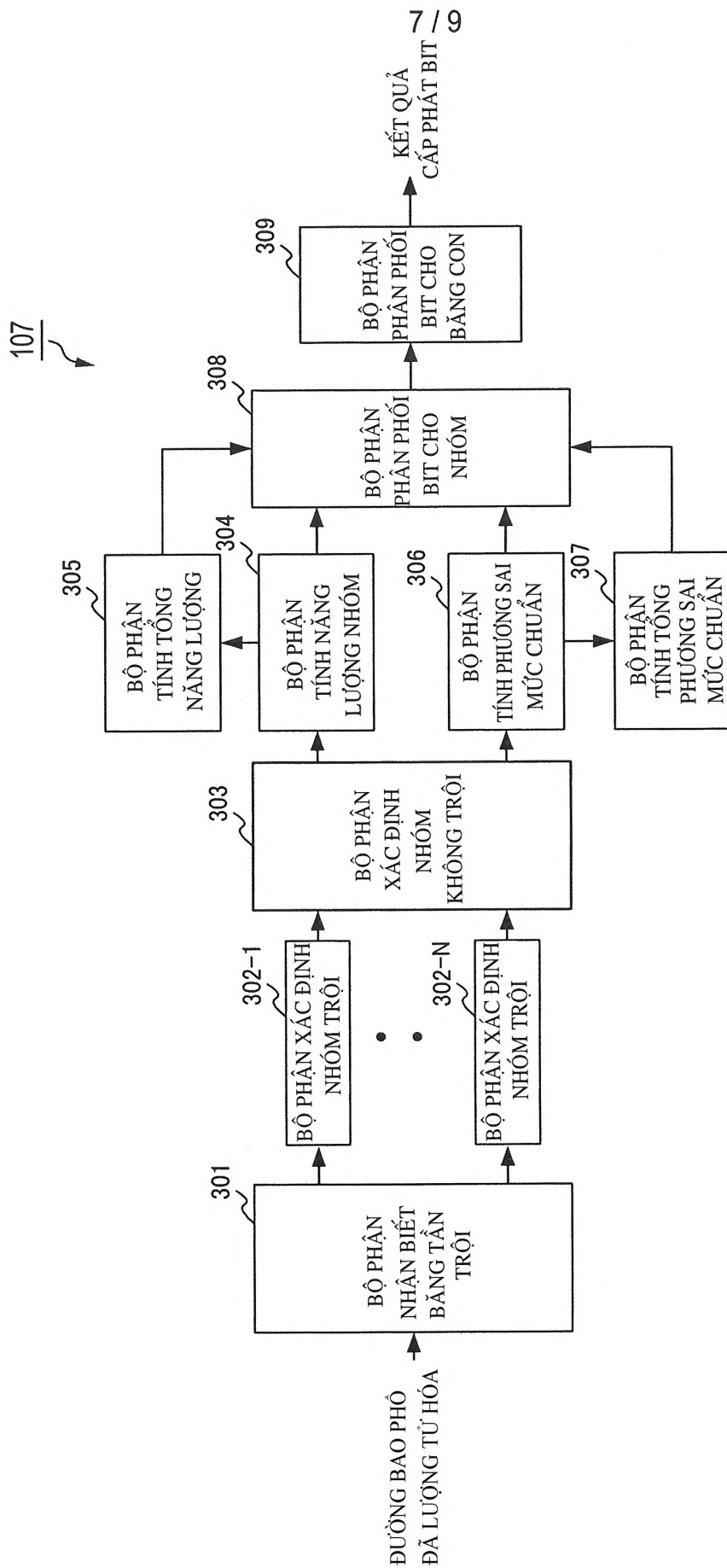


FIG. 7

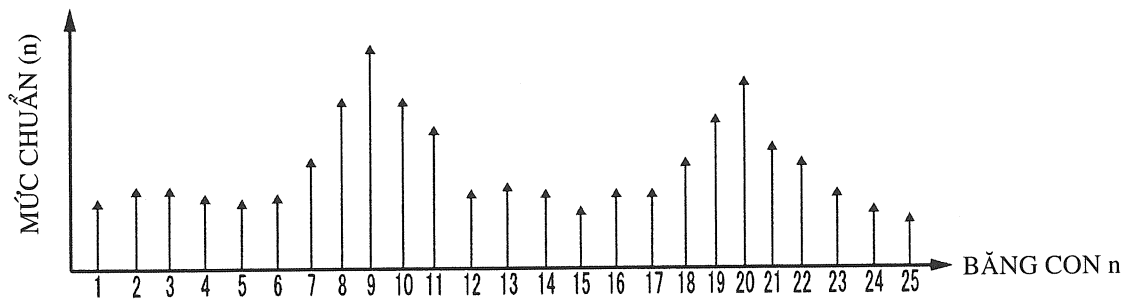


FIG. 8A

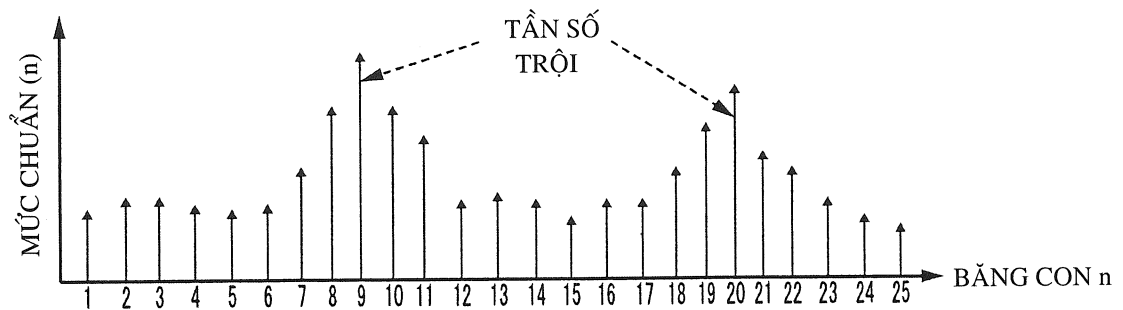


FIG. 8B

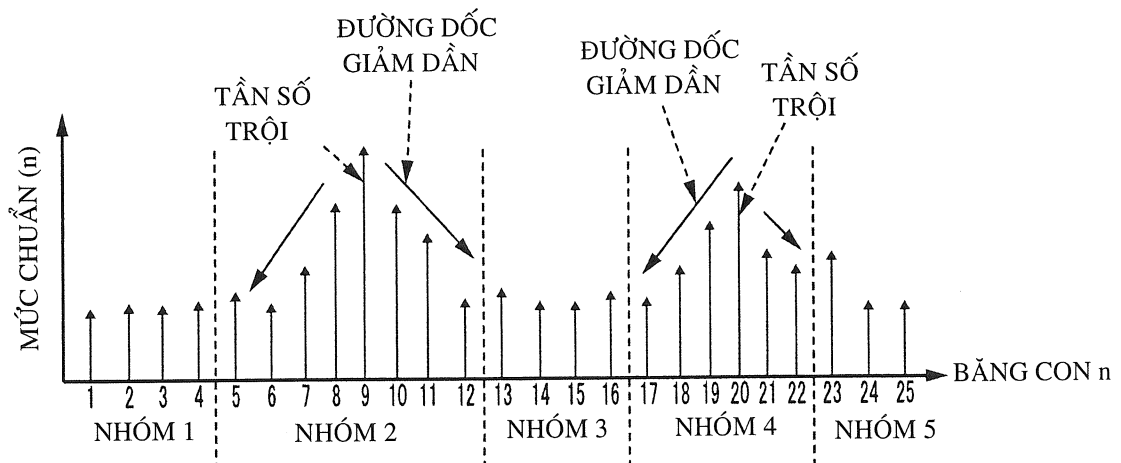
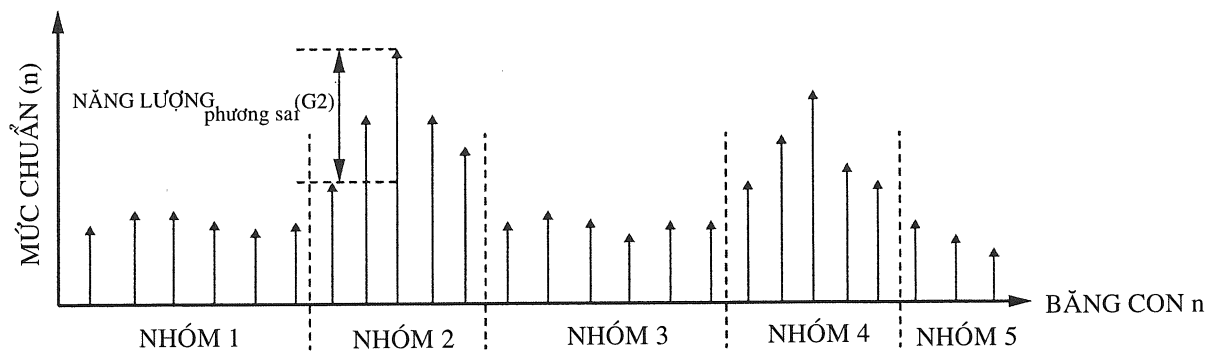


FIG. 8C

*FIG. 9*