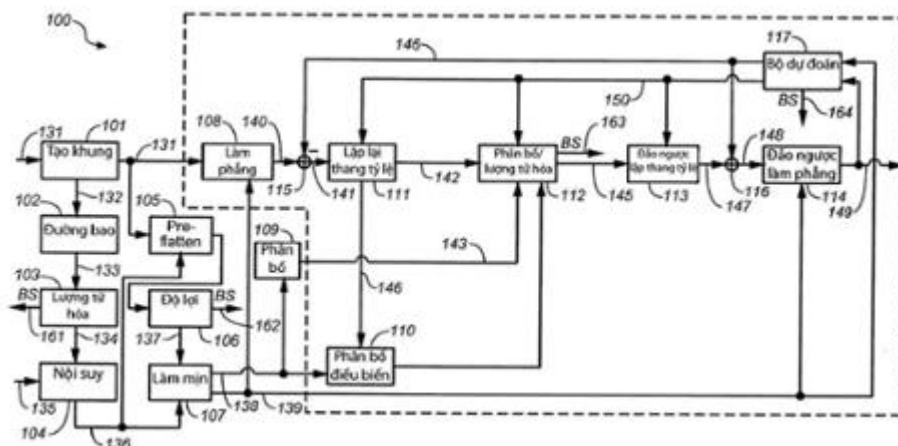


(21) 1-2015-04220 (22) 04/04/2014
(86) PCT/EP2014/056851 04/04/2014 (87) WO 2014/161991 A2 09/10/2014
(30) 61/808,675 05/04/2013 US; 61/875,553 09/09/2013 US
(45) 26/10/2020 391 (43) 25/01/2016 334A
(73) DOLBY INTERNATIONAL AB (NL)
Apollo Building, 3E, Herikerbergweg 1-35, NL-1101 CN Amsterdam, Netherlands
(72) VILLEMOES, Lars (DK); KLEJSA, Janusz (PL); HEDELIN, Per (SE).
(74) Công ty TNHH Quốc tế D & N (D&N INTERNATIONAL CO.,LTD.)

(57) Sáng chế đề cập đến bộ mã hóa và giải mã âm thanh (còn gọi là bộ mã hóa - giải mã âm thanh). Cụ thể là, sáng chế đề cập đến bộ mã hóa - giải mã âm thanh dựa trên biến đổi mà đặc biệt phù hợp để mã hóa/giải mã tiếng nói. Các bộ mã hóa tiếng nói dựa trên biến đổi được tạo cấu hình để mã hóa các tín hiệu tiếng nói thành dòng bit và các bộ giải mã tiếng nói dựa trên biến đổi được tạo cấu hình để giải mã dòng bit nêu trên. Ngoài ra, sáng chế còn đề xuất phương pháp mã hóa các tín hiệu âm thanh và tiếng nói thành dòng bit và giải mã dòng bit đó.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến hệ thống mã hóa và giải mã âm thanh (được gọi là hệ thống mã hóa-giải mã âm thanh). Cụ thể là, sáng chế đề cập đến hệ thống mã hóa-giải mã âm thanh dựa trên biến đổi mà đặc biệt phù hợp để mã hóa/giải mã giọng nói.

Tình trạng kỹ thuật của sáng chế

Các bộ mã hóa âm thanh cảm quan đa dụng đạt được các độ lợi mã hóa tương đối cao nhờ sử dụng các phép biến đổi như biến đổi cosin rời rạc được điều biến (Modified Discrete Cosine Transform - MDCT) với kích cỡ khối các mẫu bao phủ một số một phần mười mili giây (ví dụ, 20 ms). Ví dụ về một hệ thống mã hóa-giải mã âm thanh dựa trên biến đổi như vậy là Mã hóa âm thanh nâng cao (Advanced Audio Coding - AAC) hoặc Hiệu Suất Cao (High Efficiency – HE)- AAC. Tuy nhiên, khi sử dụng các hệ thống mã hóa-giải mã âm thanh dựa trên biến đổi như thế đối với các tín hiệu thoại, chất lượng của các tín hiệu thoại suy giảm nhanh hơn chất lượng của các tín hiệu âm nhạc về các tốc độ bit thấp hơn, đặc biệt đối với các tín hiệu tiếng nói khan (không vang).

Do đó, các hệ thống mã hóa-giải mã âm thanh dựa trên biến đổi vốn không hoàn toàn phù hợp để mã hóa các tín hiệu thoại hoặc mã hóa các tín hiệu âm thanh có thành phần thoại. Nói cách khác, các hệ thống mã hóa-giải mã âm thanh dựa trên biến đổi thể hiện tính không đối xứng về độ lợi mã hóa đạt được đối với các tín hiệu âm nhạc so với độ lợi mã hóa đạt được đối với các tín hiệu thoại. Tính không đối xứng này có thể được giải quyết bằng cách đưa các phần bổ sung vào quá trình mã hóa dựa trên biến đổi, trong đó các phần bổ sung nhằm mục đích cải thiện việc tạo hình phổ hoặc so khớp tín hiệu. Ví dụ về các phần bổ sung như vậy là trước/sau khi tạo hình dạng, tạo dạng nhiễu tạm thời (Temporal Noise Shaping - TNS) và MDCT biến dạng theo thời gian. Hơn thế nữa, tính không đối xứng này có thể được giải quyết bằng cách đưa vào sử dụng bộ mã hóa tiếng nói miền thời gian cổ điển dựa trên lọc dự đoán ngắn hạn (LPC) và dự đoán dài hạn (long term prediction - LTP).

Có thể thấy rằng những cải tiến thu được bằng cách đưa các phần bổ sung vào quá trình mã hóa dựa trên biến đổi thường là không đủ để giảm khoảng cách về hiệu suất giữa việc mã hóa các tín hiệu âm nhạc và các tín hiệu tiếng nói. Mặt khác, việc đưa vào sử dụng bộ mã hóa tiếng nói miền thời gian cổ điển lấp đầy khoảng cách về hiệu suất, tuy nhiên, đến mức mà tính không đối xứng về hiệu suất bị đảo ngược theo hướng ngược lại. Đây là do các bộ mã hóa tiếng nói miền thời gian cổ điển lập mô hình cho hệ thống tạo ra tiếng nói của người và đã được tối ưu hóa để mã hóa các tín hiệu tiếng nói.

Qua xem xét những điều trên, bộ mã hóa-giải mã âm thanh dựa trên biến đổi có thể được sử dụng kết hợp với bộ mã hóa-giải mã tiếng nói miền thời gian cổ điển, trong đó bộ mã hóa-giải mã tiếng nói miền thời gian cổ điển được dùng cho các đoạn tiếng nói của tín hiệu âm thanh và trong đó bộ mã hóa-giải mã dựa trên biến đổi được dùng cho các phân đoạn còn lại của tín hiệu âm thanh. Tuy nhiên, việc đồng tồn tại của bộ mã hóa-giải mã miền thời gian và miền biến đổi trong một hệ thống mã hóa-giải mã âm thanh đòi hỏi nhất đòi hỏi các công cụ đáng tin cậy để chuyển đổi giữa các bộ mã hóa-giải mã khác nhau, dựa trên các đặc tính của tín hiệu âm thanh. Ngoài ra, có thể khó thực hiện chuyển đổi thực giữa bộ mã hóa-giải mã miền thời gian (đối với nội dung tiếng nói) và bộ mã hóa-giải mã miền biến đổi (đối với nội dung còn lại). Cụ thể là, có thể khó đảm bảo sự chuyển tiếp nhịp nhàng giữa bộ mã hóa-giải mã miền thời gian và bộ mã hóa-giải mã miền biến đổi (và ngược lại). Hơn thế nữa, các biến đổi về bộ mã hóa-giải mã miền thời gian có thể được yêu cầu nhằm khiến cho bộ mã hóa-giải mã miền thời gian mạnh mẽ hơn để mã hóa đặc biệt không thể tránh khỏi của các tín hiệu phi tiếng nói, ví dụ để mã hóa giọng hát với nhạc nền.

Sáng chế giải quyết các vấn đề kỹ thuật nêu trên của các hệ thống mã hóa-giải mã âm thanh. Cụ thể là, sáng chế mô tả hệ thống mã hóa-giải mã âm thanh mà chỉ dịch các đặc điểm tới hạn của bộ mã hóa-giải mã tiếng nói và qua đó đạt được chất lượng đồng đều đối với tiếng nói và âm nhạc, trong khi vẫn nằm trong cấu trúc bộ mã hóa-giải mã dựa trên biến đổi. Nói cách khác, sáng chế mô tả bộ mã hóa-giải mã âm thanh dựa trên biến đổi mà đặc biệt thích hợp để mã hóa tín hiệu tiếng nói hoặc tín hiệu thoại.

Bản chất kỹ thuật của sáng chế

Theo một khía cạnh, bộ mã hóa tiếng nói dựa trên biến đổi được mô tả. Bộ mã hóa tiếng nói được tạo cấu hình để mã hóa tín hiệu tiếng nói thành dòng bit. Cần lưu ý rằng sau đây, nhiều khía cạnh khác nhau của bộ mã hóa tiếng nói dựa trên biến đổi sẽ được mô tả. Chỉ ra rõ ràng rằng các khía cạnh này có thể được kết hợp với một khía cạnh khác theo nhiều cách khác nhau. Cụ thể là, các khía cạnh được mô tả theo sự phụ thuộc vào các điểm yêu cầu bảo hộ độc lập khác nhau có thể được kết hợp với các điểm yêu cầu bảo hộ độc lập khác. Hơn nữa, các khía cạnh được mô tả trong bối cảnh của bộ mã hóa là có thể ứng dụng được theo cách tương tự với bộ giải mã tương ứng.

Bộ mã hóa tiếng nói có thể bao gồm đơn vị tạo khung được tạo cấu hình để nhận tập hợp khối. Tập hợp khối có thể tương ứng với tập hợp chuyển đổi của các khối được mô tả trong phần mô tả chi tiết của bản mô tả này. Theo cách khác, tập hợp khối có thể tương ứng với tập hợp hiện tại của các khối được mô tả trong phần mô tả chi tiết của bản mô tả này. Tập hợp khối bao gồm nhiều khối liên tiếp gồm các hệ số biến đổi, và nhiều khối liên tiếp chỉ báo các mẫu của tín hiệu tiếng nói. Cụ thể là, tập hợp khối có thể bao gồm bốn khối hoặc nhiều hơn các hệ số biến đổi. Khối của nhiều khối liên tiếp có thể đã được xác định từ tín hiệu tiếng nói nhờ sử dụng đơn vị biến đổi mà được tạo cấu hình để biến đổi số lượng mẫu định trước của tín hiệu tiếng nói từ miền thời gian thành miền tần số. Cụ thể là, đơn vị biến đổi có thể được tạo cấu hình để thực hiện biến đổi miền thời gian thành miền tần số như biến đổi cosin rời rạc được điều biến (MDCT). Thực chất, khối hệ số biến đổi có thể bao gồm nhiều hệ số biến đổi (còn được gọi là các hệ số tần số hoặc các hệ số phổ) đối với nhiều bin tần số tương ứng. Cụ thể là, khối hệ số biến đổi có thể bao gồm các hệ số MDCT.

Số lượng bin tần số hoặc kích thước khối thường phụ thuộc vào kích thước của phép biến đổi được thực hiện bởi đơn vị biến đổi. Trong một ví dụ được ưu tiên, các khối từ nhiều khối liên tiếp tương ứng với cái gọi là các khối ngắn, bao gồm, ví dụ, 256 bin tần số. Ngoài các khối ngắn, đơn vị biến đổi có thể được tạo cấu hình để để tạo ra cái gọi là các khối dài, bao gồm, ví dụ, 1024 bin tần số. Các khối dài có thể được sử dụng bởi bộ mã hóa âm thanh để mã hóa các phân đoạn cố định của tín hiệu âm thanh đầu vào. Tuy nhiên, nhiều khối liên tiếp được sử dụng để mã hóa tín hiệu tiếng nói (hoặc đoạn tiếng nói được bao gồm trong tín hiệu âm thanh đầu vào) có thể

chỉ bao gồm các khối ngắn. Cụ thể là, các khối hệ số biến đổi có thể bao gồm 256 hệ số biến đổi trong 256 bin tần số.

Nói chung hơn, số lượng bin tần số hoặc kích thước khối có thể là sao cho khối hệ số biến đổi trải dài trong khoảng từ 3 đến 7 mili giây của tín hiệu tiếng nói (ví dụ, 5ms của tín hiệu tiếng nói). Kích thước của khối có thể được lựa chọn sao cho bộ mã hóa tiếng nói có thể hoạt động đồng bộ với các khung hình được mã hóa bởi bộ mã hóa hình ảnh. Đơn vị biến đổi có thể được tạo cấu hình để tạo ra các khối hệ số biến đổi có số lượng bin tần số khác nhau. Ví dụ, đơn vị biến đổi có thể được tạo cấu hình để tạo ra các khối có 1920, 960, 480, 240, 120 bin tần số ở tốc độ lấy mẫu 48 kHz. Kích thước khối trải dài trong khoảng từ 3 đến 7ms của tín hiệu tiếng nói có thể được sử dụng cho bộ mã hóa tiếng nói. Trong ví dụ nêu trên, khối bao gồm 240 bin tần số có thể được sử dụng cho bộ mã hóa tiếng nói.

Bộ mã hóa tiếng nói còn có thể bao gồm đơn vị ước lượng đường bao được tạo cấu hình để xác định đường bao hiện thời dựa trên nhiều khối liên tiếp gồm các hệ số biến đổi. Đường bao hiện thời có thể được xác định dựa trên nhiều khối liên tiếp của tập hợp khối. Các khối bổ sung có thể được xem xét, ví dụ các khối của tập hợp khối đứng liền trước tập hợp khối. Theo cách khác hoặc thêm vào đó, chính là các khối đứng phía trước có thể được đưa vào xem xét. Nói chung, điều này có thể có lợi cho việc tạo ra tính liên tục giữa các tập hợp khối liên tiếp. Đường bao hiện thời có thể chỉ báo nhiều trị số năng lượng phổ đối với nhiều bin tần số tương ứng. Nói cách khác, đường bao hiện thời có thể có kích thước giống như mỗi khối trong nhiều khối liên tiếp. Nói theo cách khác nữa, một đường bao hiện thời duy nhất có thể được xác định cho nhiều (tức là cho nhiều hơn một) khối tín hiệu tiếng nói. Đây này là có lợi nhằm cung cấp số liệu thống kê ý nghĩa về dữ liệu phổ được bao gồm trong nhiều khối liên tiếp.

Đường bao hiện thời có thể chỉ báo nhiều trị số năng lượng phổ đối với nhiều dải tần số tương ứng. Dải tần số có thể bao gồm một hoặc nhiều bin tần số. Cụ thể là, một hoặc nhiều trong số các dải tần số có thể bao gồm nhiều hơn một bin tần số. Số lượng bin tần số trên mỗi dải tần số có thể tăng lên khi tần số tăng. Nói cách khác, số lượng bin tần số trên mỗi dải tần số có thể phụ thuộc vào các xem xét về mặt cảm quan. Đơn vị ước lượng đường bao có thể được tạo cấu hình để xác định trị số năng lượng phổ đối với dải tần số cụ thể dựa trên các hệ số biến đổi của nhiều khối liên tiếp

nằm trong dải tần số cụ thể. Cụ thể là, đơn vị ước lượng đường bao có thể được tạo cấu hình để xác định trị số năng lượng phổ đối với dải tần số cụ thể dựa trên trị số căn quân phương của các hệ số biến đổi của nhiều khối liên tiếp nằm trong dải tần số cụ thể. Như vậy, đường bao hiện thời có thể chỉ báo đường bao phổ trung bình của các đường bao phổ của nhiều khối liên tiếp. Hơn thế nữa, đường bao hiện thời có thể có độ phân giải tần số dạng dải.

Bộ mã hóa tiếng nói còn có thể bao gồm đơn vị nội suy đường bao được tạo cấu hình để xác định nhiều đường bao được nội suy đối với nhiều khối liên tiếp gồm các hệ số biến đổi một cách lần lượt, dựa trên đường bao hiện thời. Cụ thể là, nhiều đường bao được nội suy có thể được xác định dựa trên đường bao hiện thời được lượng tử hóa, mà cũng có sẵn ở bộ giải mã tương ứng. Bằng cách làm như vậy, nhiều đường bao được nội suy được bảo đảm có thể được xác định theo cùng một cách ở bộ mã hóa tiếng nói và ở bộ giải mã tiếng nói tương ứng. Do đó, các đặc điểm của đơn vị nội suy đường bao được mô tả trong phạm vi của bộ giải mã tiếng nói cũng có thể áp dụng được cho bộ mã hóa tiếng nói và ngược lại. Nói chung, đơn vị nội suy đường bao có thể được tạo cấu hình để xác định trị số gần đúng của đường bao phổ của mỗi khối tiếp trong số nhiều khối liên tiếp (tức là đường bao được nội suy), dựa trên đường bao hiện thời.

Bộ mã hóa tiếng nói còn có thể bao gồm đơn vị làm phẳng được tạo cấu hình để xác định nhiều khối hệ số biến đổi được làm phẳng bằng cách làm phẳng nhiều khối tương ứng của các hệ số biến đổi nhờ lần lượt sử dụng nhiều đường bao được nội suy tương ứng. Cụ thể là, đường bao được nội suy đối với khối cụ thể (hoặc đường bao được lấy từ đó) có thể được sử dụng để làm phẳng, tức là để loại bỏ hình dạng phổ của, các hệ số biến đổi được bao gồm trong khối cụ thể đó. Cần lưu ý rằng quá trình làm phẳng này khác với thao tác tẩy trắng được áp dụng cho khối cụ thể của các hệ số biến đổi. Nghĩa là, các hệ số biến đổi được làm phẳng không thể được hiểu là các hệ số biến đổi của tín hiệu được tẩy trắng miền thời gian như thường được tạo ra bởi phép phân tích LPC (linear predictive coding - mã hóa dự đoán tuyến tính) của bộ mã hóa tiếng nói cổ điển. Chỉ có chung khía cạnh tạo ra tín hiệu có phổ năng lượng tương đối phẳng. Tuy nhiên, quy trình thu được phổ năng lượng phẳng như vậy lại khác nhau. Như sẽ được nêu trong bản mô tả này, việc sử dụng đường bao phổ được

ước lượng để làm phẳng khối hệ số biến đổi là có lợi, vì đường bao phổ được ước lượng có thể được sử dụng cho các mục đích phân bố bit.

Bộ mã hóa tiếng nói dựa trên biến đổi còn có thể bao gồm đơn vị xác định độ lợi đường bao được tạo cấu hình để xác định nhiều độ lợi đường bao đối với nhiều khối hệ số biến đổi một cách lần lượt. Hơn thế nữa, bộ mã hóa tiếng nói dựa trên biến đổi có thể bao gồm đơn vị làm mịn đường bao được tạo cấu hình để xác định nhiều đường bao được điều chỉnh bằng cách thay đổi nhiều đường bao được nội suy theo nhiều độ lợi đường bao một cách lần lượt. Đơn vị xác định độ lợi đường bao có thể được tạo cấu hình để xác định độ lợi đường bao thứ nhất đối với khối thứ nhất của các hệ số biến đổi (từ nhiều khối liên tiếp), sao cho phương sai của các hệ số biến đổi được làm phẳng của khối thứ nhất tương ứng gồm các hệ số biến đổi được làm phẳng mà thu được nhờ sử dụng đường bao được điều chỉnh thứ nhất được làm giảm đi so với phương sai của các hệ số biến đổi được làm phẳng của khối thứ nhất tương ứng gồm các hệ số biến đổi được làm phẳng mà thu được nhờ sử dụng đường bao được nội suy thứ nhất. Đường bao được điều chỉnh thứ nhất có thể được xác định bằng cách thay đổi đường bao được nội suy thứ nhất nhờ sử dụng độ lợi đường bao thứ nhất. Đường bao được nội suy thứ nhất có thể là đường bao được nội suy từ nhiều đường bao được nội suy đối với khối thứ nhất của các hệ số biến đổi từ nhiều khối hệ số biến đổi.

Cụ thể là, đơn vị xác định độ lợi đường bao có thể được tạo cấu hình để xác định độ lợi đường bao thứ nhất đối với khối thứ nhất của các hệ số biến đổi, sao cho phương sai của các hệ số biến đổi được làm phẳng của khối thứ nhất tương ứng gồm các hệ số biến đổi được làm phẳng thu được nhờ sử dụng đường bao được điều chỉnh thứ nhất là một. Đơn vị làm phẳng có thể được tạo cấu hình để xác định nhiều khối hệ số biến đổi được làm phẳng bằng cách làm phẳng nhiều khối tương ứng của các hệ số biến đổi nhờ sử dụng nhiều đường bao được điều chỉnh tương ứng một cách lần lượt. Kết quả là, mỗi khối hệ số biến đổi được làm phẳng có thể có phương sai là một.

Đơn vị xác định độ lợi đường bao có thể được tạo cấu hình để chèn dữ liệu độ lợi chỉ báo nhiều độ lợi đường bao vào dòng bit. Kết quả là, bộ giải mã tương ứng được cho phép để xác định nhiều đường bao được điều chỉnh theo cùng một cách như bộ mã hóa.

Bộ mã hóa tiếng nói có thể được tạo cấu hình để xác định dòng bit dựa trên nhiều khối hệ số biến đổi được làm phẳng. Cụ thể là, bộ mã hóa tiếng nói có thể được tạo cấu hình để xác định dữ liệu hệ số dựa trên nhiều khối hệ số biến đổi được làm phẳng, trong đó dữ liệu hệ số được chèn vào dòng bit. Các phương thức ví dụ để xác định dữ liệu hệ số dựa trên nhiều khối hệ số biến đổi được làm phẳng được mô tả dưới đây.

Bộ mã hóa tiếng nói dựa trên biến đổi có thể bao gồm đơn vị lượng tử hóa đường bao được tạo cấu hình để xác định đường bao hiện thời được lượng tử hóa bằng cách lượng tử hóa đường bao hiện thời. Hơn thế nữa, đơn vị lượng tử hóa đường bao có thể được tạo cấu hình để chèn dữ liệu đường bao vào dòng bit, trong đó dữ liệu đường bao chỉ báo đường bao hiện thời được lượng tử hóa. Kết quả là, bộ giải mã tương ứng có thể được làm cho nhận ra đường bao hiện thời được lượng tử hóa bằng cách giải mã dữ liệu đường bao. Đơn vị nội suy đường bao có thể được tạo cấu hình để xác định nhiều đường bao được nội suy, dựa trên đường bao hiện thời được lượng tử hóa. Bằng cách thực hiện như vậy, bộ mã hóa và bộ giải mã có thể đảm bảo được tạo cấu hình để xác định nhiều đường bao được nội suy giống nhau.

Bộ mã hóa tiếng nói dựa trên biến đổi có thể được tạo cấu hình để hoạt động trong nhiều chế độ khác nhau. Các chế độ khác nhau có thể bao gồm chế độ bước ngắn và chế độ bước dài. Đơn vị tạo khung, đơn vị ước lượng đường bao và đơn vị nội suy đường bao có thể được tạo cấu hình để xử lý tập hợp khối bao gồm nhiều khối liên tiếp gồm các hệ số biến đổi, khi bộ mã hóa tiếng nói dựa trên biến đổi được hoạt động trong chế độ bước ngắn. Do đó, khi ở trong chế độ bước ngắn, bộ mã hóa có thể được tạo cấu hình để chia nhỏ phân đoạn / khung của tín hiệu âm thanh thành dãy khối liên tiếp mà được xử lý bởi bộ mã hóa theo cách liên tiếp.

Mặt khác, đơn vị tạo khung, đơn vị ước lượng đường bao và đơn vị nội suy đường bao có thể được tạo cấu hình để xử lý tập hợp khối chỉ bao gồm một khối duy nhất của các hệ số biến đổi, khi bộ mã hóa tiếng nói dựa trên biến đổi được hoạt động trong chế độ bước dài. Do đó, khi ở trong chế độ bước dài, bộ mã hóa có thể được tạo cấu hình để xử lý phân đoạn / khung hoàn chỉnh của tín hiệu âm thanh, mà không bị chia nhỏ thành các khối. Việc này có thể có lợi đối với các phân đoạn / khung ngắn của tín hiệu âm thanh, và/hoặc đối với các tín hiệu âm nhạc. Khi ở trong chế độ bước dài, đơn vị ước lượng đường bao có thể được tạo cấu hình để xác định đường bao hiện

thời của một khối duy nhất của các hệ số biến đổi được bao gồm trong tập hợp khối. Đơn vị nội suy đường bao có thể được tạo cấu hình để xác định đường bao được nội suy đối với một khối duy nhất của các hệ số biến đổi làm đường bao hiện thời của một khối duy nhất của các hệ số biến đổi. Nói cách khác, sự nội suy đường bao được nêu trong bản mô tả này có thể được bỏ qua, khi ở trong chế độ bước dài, và đường bao hiện thời của một khối duy nhất có thể được thiết lập là đường bao được nội suy (để xử lý thêm).

Theo một khía cạnh khác, sáng chế đề xuất bộ giải mã tiếng nói dựa trên biến đổi được tạo cấu hình để giải mã dòng bit nhằm cung cấp tín hiệu tiếng nói được tái tạo. Như đã nêu trên, bộ giải mã có thể bao gồm các bộ phận tương tự với các bộ phận của bộ mã hóa tương ứng. Bộ giải mã có thể bao gồm đơn vị giải mã đường bao được tạo cấu hình để xác định đường bao hiện thời được lượng tử hóa từ dữ liệu đường bao được bao gồm trong dòng bit. Như nêu trên, đường bao hiện thời được lượng tử hóa thường chỉ báo nhiều trị số năng lượng phổ đối với nhiều bin tần số tương ứng của các dải tần số. Hơn nữa, dòng bit có thể bao gồm dữ liệu (ví dụ, dữ liệu hệ số) chỉ báo nhiều khối liên tiếp của các hệ số biến đổi đã làm phẳng được tái tạo. Nhiều khối liên tiếp của các hệ số biến đổi đã làm phẳng được tái tạo thường liên kết với nhiều khối liên tiếp tương ứng của các hệ số biến đổi được làm phẳng ở bộ mã hóa. Nhiều khối liên tiếp này có thể tương ứng với nhiều khối liên tiếp của tập hợp khối, ví dụ của tập hợp khối được chuyển đổi được mô tả dưới đây. Khối hệ số biến đổi đã làm phẳng được tái tạo có thể bao gồm nhiều hệ số biến đổi đã làm phẳng được tái tạo đối với nhiều bin tần số tương ứng.

Bộ giải mã còn có thể bao gồm đơn vị nội suy đường bao được tạo cấu hình để xác định nhiều đường bao được nội suy đối với nhiều khối hệ số biến đổi đã làm phẳng được tái tạo một cách lần lượt, dựa trên đường bao hiện thời được lượng tử hóa. Đơn vị nội suy đường bao của bộ giải mã thường hoạt động theo cùng một cách với đơn vị nội suy đường bao của bộ mã hóa. Đơn vị nội suy đường bao có thể được tạo cấu hình để xác định nhiều đường bao được nội suy còn dựa trên đường bao trước đó được lượng tử hóa. Đường bao trước đó được lượng tử hóa có thể được liên kết với nhiều khối trước đó của các hệ số biến đổi được tái tạo, đứng ngay trước nhiều khối hệ số biến đổi được tái tạo. Như thế, đường bao trước đó được lượng tử hóa có thể được nhận bởi bộ giải mã dưới dạng dữ liệu đường bao đối với tập hợp trước đó của các

khối hệ số biến đổi (ví dụ, trong trường hợp của khung P). Theo cách khác hoặc ngoài ra, dữ liệu đường bao đối với tập hợp khối có thể chỉ báo đường bao trước đó được lượng tử hóa bên cạnh chỉ báo đường bao hiện thời được lượng tử hóa (ví dụ, trong trường hợp của khung I). Điều này cho phép khung I được giải mã mà không cần biết dữ liệu trước đó.

Đơn vị nội suy đường bao có thể được tạo cấu hình để xác định trị số năng lượng phổ đối với bin tần số cụ thể của đường bao được nội suy thứ nhất bằng cách nội suy các trị số năng lượng phổ đối với bin tần số cụ thể của đường bao hiện thời được lượng tử hóa và của đường bao trước đó được lượng tử hóa ở thời điểm trung gian thứ nhất. Đường bao được nội suy thứ nhất liên kết với hoặc tương ứng với khối thứ nhất của nhiều khối liên tiếp của các hệ số biến đổi đã làm phẳng được tái tạo. Như đã nêu trên, các đường bao trước đó và hiện tại được lượng tử hóa thường là các đường bao dạng dải. Các trị số năng lượng phổ đối với dải tần số cụ thể thường là không đổi đối với tất cả các bin tần số được bao gồm trong dải tần số.

Đơn vị nội suy đường bao có thể được tạo cấu hình để xác định trị số năng lượng phổ đối với bin tần số cụ thể của đường bao được nội suy thứ nhất bằng cách lượng tử hóa phép nội suy giữa các trị số năng lượng phổ đối với bin tần số cụ thể của đường bao hiện thời được lượng tử hóa và của đường bao trước đó được lượng tử hóa. Như vậy, nhiều đường bao được nội suy có thể là các đường bao được nội suy được lượng tử hóa.

Đơn vị nội suy đường bao có thể được tạo cấu hình để xác định trị số năng lượng phổ đối với bin tần số cụ thể của đường bao được nội suy thứ hai bằng cách nội suy các trị số năng lượng phổ đối với bin tần số cụ thể của đường bao hiện thời được lượng tử hóa và của đường bao trước đó được lượng tử hóa ở thời điểm trung gian thứ hai. Đường bao được nội suy thứ hai có thể được liên kết với hoặc có thể tương ứng với khối thứ hai của nhiều khối hệ số biến đổi đã làm phẳng được tái tạo. Khối thứ hai của các hệ số biến đổi đã làm phẳng được tái tạo có thể theo sau khối thứ nhất của các hệ số biến đổi đã làm phẳng được tái tạo và thời điểm trung gian thứ hai có thể theo sau thời điểm trung gian thứ nhất. Cụ thể là, hiệu số giữa thời điểm trung gian thứ hai và thời điểm trung gian thứ nhất có thể tương ứng với khoảng thời gian giữa khối thứ hai của các hệ số biến đổi đã làm phẳng được tái tạo và khối thứ nhất của các hệ số biến đổi đã làm phẳng được tái tạo.

Đơn vị nội suy đường bao có thể được tạo cấu hình để thực hiện một hoặc nhiều trong số: nội suy tuyến tính, nội suy hình học, và nội suy điều hòa. Hơn nữa, đơn vị nội suy đường bao có thể được tạo cấu hình để thực hiện nội suy trong miền lôgarit.

Hơn nữa, bộ giải mã có thể bao gồm đơn vị làm phẳng ngược được tạo cấu hình để xác định nhiều khối hệ số biến đổi được tái tạo bằng cách cung cấp dạng phổ cho nhiều khối tương ứng của các hệ số biến đổi đã làm phẳng được tái tạo, sử dụng nhiều đường bao được nội suy tương ứng một cách lần lượt.

Như nêu trên, dòng bit có thể chỉ báo nhiều độ lợi đường bao (trong dữ liệu độ lợi) đối với nhiều khối hệ số biến đổi đã làm phẳng được tái tạo một cách lần lượt. Bộ giải mã tiếng nói dựa trên biến đổi còn có thể bao gồm đơn vị làm mịn đường bao được tạo cấu hình để xác định nhiều đường bao được điều chỉnh bằng cách áp dụng nhiều độ lợi đường bao vào nhiều đường bao được nội suy một cách lần lượt. Đơn vị làm phẳng ngược có thể được tạo cấu hình để xác định nhiều khối hệ số biến đổi được tái tạo bằng cách cung cấp dạng phổ cho nhiều khối tương ứng của các hệ số biến đổi đã làm phẳng được tái tạo, sử dụng nhiều đường bao được điều chỉnh tương ứng một cách lần lượt.

Bộ giải mã có thể được tạo cấu hình để xác định tín hiệu tiếng nói được tái tạo dựa trên nhiều khối hệ số biến đổi được tái tạo.

Theo khía cạnh khác, bộ mã hóa tiếng nói dựa trên biến đổi được tạo cấu hình để mã hóa tín hiệu tiếng nói thành dòng bit được mô tả. Bộ mã hóa có thể bao gồm bất kỳ đặc điểm và/hoặc các bộ phận liên quan đến bộ mã hóa được nêu trong bản mô tả này. Cụ thể là, bộ mã hóa có thể bao gồm đơn vị tạo khung được tạo cấu hình để nhận nhiều khối liên tiếp gồm các hệ số biến đổi. Nhiều khối liên tiếp bao gồm khối hiện tại và một hoặc nhiều khối trước đó. Như nêu trên, nhiều khối liên tiếp chỉ báo các mẫu của tín hiệu tiếng nói.

Hơn nữa, bộ mã hóa có thể bao gồm đơn vị làm phẳng được tạo cấu hình để xác định khối hiện tại và một hoặc nhiều khối trước đó của các hệ số biến đổi được làm phẳng bằng cách làm phẳng khối hiện tại tương ứng và một hoặc nhiều khối trước đó của các hệ số biến đổi nhờ sử dụng đường bao khối hiện tại tương ứng và một hoặc nhiều đường bao khối trước đó tương ứng một cách lần lượt. Các đường bao khối có thể tương ứng với các đường bao được điều chỉnh nêu trên.

Ngoài ra, bộ mã hóa bao gồm bộ dự đoán được tạo cấu hình để xác định khối hiện tại của các hệ số biến đổi được làm phẳng được ước lượng dựa trên một hoặc nhiều khối trước đó của các hệ số biến đổi được tái tạo và dựa trên một hoặc nhiều tham số dự đoán. Một hoặc nhiều khối trước đó của các hệ số biến đổi được tái tạo có thể được lấy từ một hoặc nhiều khối trước đó của các hệ số biến đổi được làm phẳng, một cách lần lượt (ví dụ, sử dụng bộ dự đoán).

Bộ dự đoán có thể bao gồm bộ tách được tạo cấu hình để xác định khối hiện tại của các hệ số biến đổi được ước lượng dựa trên một hoặc nhiều khối trước đó của các hệ số biến đổi được tái tạo và dựa trên một hoặc nhiều tham số dự đoán. Như thế, bộ tách có thể hoạt động trong miền không phẳng (tức là bộ tách có thể hoạt động trên các khối hệ số biến đổi có dạng phổ). Việc này có thể có lợi đối với mô hình tín hiệu được sử dụng bởi bộ tách để xác định khối hiện tại của các hệ số biến đổi được ước lượng.

Hơn nữa, bộ dự đoán có thể bao gồm bộ tạo hình phổ được tạo cấu hình để xác định khối hiện tại của các hệ số biến đổi được làm phẳng được ước lượng dựa trên khối hiện tại của các hệ số biến đổi được ước lượng, dựa trên ít nhất là một trong số một hoặc nhiều đường bao khối trước đó và dựa trên ít nhất là một trong số một hoặc nhiều tham số dự đoán. Như thế, bộ tạo hình phổ có thể được tạo cấu hình để chuyển đổi khối hiện tại của các hệ số biến đổi được ước lượng thành miền được làm phẳng để cung cấp khối hiện tại của các hệ số biến đổi được làm phẳng được ước lượng. Như đã được nêu trong bối cảnh của bộ giải mã tương ứng, bộ tạo hình phổ có thể sử dụng nhiều đường bao được điều chỉnh (hoặc nhiều đường bao khối) để thực hiện mục đích này.

Như nêu trên, bộ dự đoán (cụ thể là, bộ tách) có thể bao gồm bộ dự đoán dựa trên mô hình sử dụng mô hình tín hiệu. Mô hình tín hiệu có thể bao gồm một hoặc nhiều tham số mô hình, và một hoặc nhiều tham số dự đoán có thể chỉ báo một hoặc nhiều tham số mô hình. Việc sử dụng bộ dự đoán dựa trên mô hình có thể có lợi đối với việc cung cấp phương thức hiệu quả về tốc độ bit để mô tả các hệ số dự đoán được sử dụng bởi bộ dự đoán dải phụ (hoặc bin tần số). Cụ thể là, có thể xác định tập hợp hoàn chỉnh của các hệ số dự đoán bằng cách chỉ sử dụng một vài tham số mô hình, mà có thể được truyền đi dưới dạng dữ liệu bộ dự đoán đến bộ giải mã tương ứng theo cách hiệu quả về tốc độ bit.

Như thế, bộ dự đoán dựa trên mô hình có thể được tạo cấu hình để xác định một hoặc nhiều tham số mô hình của mô hình tín hiệu (ví dụ, nhờ sử dụng giải thuật Durbin-Levinson). Hơn nữa, bộ dự đoán dựa trên mô hình có thể được tạo cấu hình để xác định hệ số dự đoán sẽ được áp dụng vào hệ số biến đổi được tái tạo thứ nhất trong bin tần số thứ nhất của khối trước đó của các hệ số biến đổi được tái tạo, dựa trên mô hình tín hiệu và dựa trên một hoặc nhiều tham số mô hình. Cụ thể là, nhiều hệ số dự đoán đối với nhiều hệ số biến đổi được tái tạo có thể được xác định. Bằng cách này, giá trị ước tính của hệ số biến đổi được ước lượng thứ nhất trong bin tần số thứ nhất của khối hiện tại của các hệ số biến đổi được ước lượng có thể được xác định bằng cách áp dụng hệ số dự đoán vào hệ số biến đổi được tái tạo thứ nhất. Cụ thể là, bằng cách này, các hệ số biến đổi được ước lượng của khối hiện tại của các hệ số biến đổi được ước lượng có thể được xác định.

Ví dụ, mô hình tín hiệu có thể bao gồm một hoặc nhiều mô hình hình sin các bộ phận và một hoặc nhiều tham số mô hình có thể chỉ báo tần số của một hoặc nhiều mô hình hình sin các bộ phận. Cụ thể là, một hoặc nhiều tham số mô hình có thể chỉ báo tần số cơ sở của mô hình tín hiệu đa hình sin. Tần số cơ sở như thế có thể tương ứng với độ trễ trong miền thời gian.

Bộ dự đoán có thể được tạo cấu hình để xác định một hoặc nhiều tham số dự đoán sao cho giá trị quân phương của các hệ số sai số dự đoán của khối hiện tại của các hệ số sai số dự đoán được làm giảm đi (ví dụ, được làm giảm đến mức tối thiểu). Có thể đạt được điều này nhờ sử dụng, ví dụ, giải thuật Durbin-Levinson. Bộ dự đoán có thể được tạo cấu hình để chèn dữ liệu bộ dự đoán chỉ báo một hoặc nhiều tham số dự đoán vào dòng bit. Kết quả là, bộ giải mã tương ứng được cho phép xác định khối hiện tại của các hệ số biến đổi được làm phẳng được ước lượng theo cùng một cách như bộ mã hóa.

Hơn nữa, bộ mã hóa có thể bao gồm đơn vị vi sai được tạo cấu hình để xác định khối hiện tại của các hệ số sai số dự đoán dựa trên khối hiện tại của các hệ số biến đổi được làm phẳng và dựa trên khối hiện tại của các hệ số biến đổi được làm phẳng được ước lượng. Dòng bit có thể được xác định dựa trên khối hiện tại của các hệ số sai số dự đoán. Cụ thể là, dữ liệu hệ số của dòng bit có thể chỉ báo khối hiện tại của các hệ số sai số dự đoán.

Theo một khía cạnh khác nữa, bộ giải mã tiếng nói dựa trên biến đổi được tạo cấu hình để giải mã dòng bit nhằm cung cấp tín hiệu tiếng nói được tái tạo được mô tả. Bộ giải mã có thể bao gồm bất kỳ đặc điểm và/hoặc các bộ phận liên quan đến bộ giải mã được nêu trong bản mô tả này. Cụ thể là, bộ giải mã có thể bao gồm bộ dự đoán được tạo cấu hình để xác định khối hiện tại của các hệ số biến đổi được làm phẳng được ước lượng dựa trên một hoặc nhiều khối trước đó của các hệ số biến đổi được tái tạo và dựa trên một hoặc nhiều tham số dự đoán được lấy từ (bộ dự đoán dữ liệu của) dòng bit. Như đã được nêu trong bối cảnh của bộ mã hóa tương ứng, bộ dự đoán có thể bao gồm bộ tách được tạo cấu hình để xác định khối hiện tại của các hệ số biến đổi được ước lượng dựa trên ít nhất là một trong số một hoặc nhiều khối trước đó của các hệ số biến đổi được tái tạo và dựa trên ít nhất là một trong số một hoặc nhiều tham số dự đoán. Hơn nữa, bộ dự đoán có thể bao gồm bộ tạo hình phổ được tạo cấu hình để xác định khối hiện tại của các hệ số biến đổi được làm phẳng được ước lượng dựa trên khối hiện tại của các hệ số biến đổi được ước lượng, dựa trên một hoặc nhiều đường bao khối trước đó (ví dụ, các đường bao được điều chỉnh trước đó) và dựa trên một hoặc nhiều tham số dự đoán.

Một hoặc nhiều tham số dự đoán có thể bao gồm tham số độ trễ khối T . Tham số độ trễ khối có thể chỉ báo một số khối đi trước khối hiện tại của các hệ số biến đổi được làm phẳng được ước lượng. Cụ thể là, tham số độ trễ khối T có thể chỉ báo tính chu kỳ của tín hiệu tiếng nói. Như thế, tham số độ trễ khối T có thể chỉ báo một hoặc nhiều khối trước đó nào của các hệ số biến đổi được tái tạo là giống (nhất) với khối hiện tại của các hệ số biến đổi, và do đó có thể được sử dụng để dự đoán khối hiện tại của các hệ số biến đổi, tức là có thể được sử dụng để xác định khối hiện tại của các hệ số biến đổi được ước lượng.

Bộ tạo hình phổ có thể được tạo cấu hình để làm phẳng khối hiện tại của các hệ số biến đổi được ước lượng nhờ sử dụng đường bao được ước lượng hiện tại. Hơn nữa, bộ tạo hình phổ có thể được tạo cấu hình để xác định đường bao được ước lượng hiện tại dựa trên ít nhất là một trong số một hoặc nhiều đường bao khối trước đó và dựa trên tham số độ trễ khối. Cụ thể là, bộ tạo hình phổ có thể được tạo cấu hình để xác định trị số độ trễ nguyên T_n dựa trên tham số độ trễ khối T . Trị số độ trễ nguyên T_n có thể được xác định bằng cách làm tròn tham số độ trễ khối T đến số nguyên gần nhất. Hơn nữa, bộ tạo hình phổ có thể được tạo cấu hình để xác định đường bao được

ước lượng hiện tại là đường bao khối trước đó (ví dụ, đường bao được điều chỉnh trước đó) của khối trước đó của các hệ số biến đổi được tái tạo ở trước khối hiện tại của các hệ số biến đổi được làm phẳng được ước lượng bởi một số khối tương ứng với trị số độ trễ nguyên. Cần lưu ý rằng các đặc điểm được mô tả đối với bộ tạo hình phổ của bộ giải mã cũng có thể áp dụng được cho bộ tạo hình phổ của bộ mã hóa.

Bộ tách có thể được tạo cấu hình để xác định khối hiện tại của các hệ số biến đổi được ước lượng dựa trên ít nhất là một trong số một hoặc nhiều khối trước đó của các hệ số biến đổi được tái tạo và dựa trên tham số độ trễ khối T . Để thực hiện mục đích này, bộ tách có thể sử dụng bộ dự đoán dựa trên mô hình, như đã được nêu trong bối cảnh của bộ mã hóa tương ứng. Trong bối cảnh này, tham số độ trễ khối T có thể chỉ báo tần số cơ sở của mô hình đa tín hiệu hình sin.

Hơn nữa, bộ giải mã tiếng nói có thể bao gồm bộ giải mã phổ được tạo cấu hình để xác định khối hiện tại của các hệ số sai số dự đoán được lượng tử hóa dựa trên dữ liệu hệ số được bao gồm trong dòng bit. Để thực hiện mục đích này, bộ giải mã phổ có thể sử dụng các bộ đảo ngược lượng tử hóa như được nêu trong bản mô tả này. Ngoài ra, bộ giải mã tiếng nói có thể bao gồm đơn vị cộng được tạo cấu hình để xác định khối hiện tại của các hệ số biến đổi đã làm phẳng được tái tạo dựa trên khối hiện tại của các hệ số biến đổi được làm phẳng được ước lượng và dựa trên khối hiện tại của các hệ số sai số dự đoán được lượng tử hóa. Ngoài ra, bộ giải mã tiếng nói có thể bao gồm đơn vị làm phẳng ngược được tạo cấu hình để xác định khối hiện tại của các hệ số biến đổi được tái tạo bằng cách cung cấp khối hiện tại của các hệ số biến đổi đã làm phẳng được tái tạo với dạng phổ, sử dụng đường bao khối hiện tại. Hơn nữa, đơn vị làm phẳng có thể được tạo cấu hình để xác định một hoặc nhiều khối trước đó của các hệ số biến đổi được tái tạo bằng cách cung cấp một hoặc nhiều khối trước đó của các hệ số biến đổi đã làm phẳng được tái tạo với dạng phổ, sử dụng một hoặc nhiều đường bao khối trước đó (ví dụ, các đường bao được điều chỉnh trước đó) một cách lần lượt. Bộ giải mã tiếng nói có thể được tạo cấu hình để xác định tín hiệu tiếng nói được tái tạo dựa trên khối hiện tại và dựa trên một hoặc nhiều khối trước đó của các hệ số biến đổi được tái tạo.

Bộ giải mã tiếng nói dựa trên biến đổi có thể bao gồm bộ đệm đường bao được tạo cấu hình để lưu trữ một hoặc nhiều đường bao khối trước đó. Bộ tạo hình phổ có thể được tạo cấu hình để xác định trị số độ trễ nguyên T_0 bằng cách giới hạn trị số độ

trễ nguyên T_0 vào một số các đường bao khối trước đó được lưu trữ trong bộ đệm đường bao. Số lượng của các đường bao khối trước đó mà được lưu trữ trong bộ đệm đường bao có thể thay đổi (ví dụ, ở đầu khung I). Bộ tạo hình phổ có thể được tạo cấu hình để xác định số lượng các đường bao trước đó mà được lưu trữ trong bộ đệm đường bao và định giới hạn trị số độ trễ nguyên T_0 theo đó. Bằng cách này, có thể tránh các vòng lặp đường bao lỗi.

Bộ tạo hình phổ có thể được tạo cấu hình để làm phẳng khối hiện tại của các hệ số biến đổi được ước lượng, sao cho, trước khi áp dụng một hoặc nhiều tham số dự đoán (đặc biệt là trước khi áp dụng độ lợi bộ dự đoán), khối hiện tại của các hệ số biến đổi được ước lượng được làm phẳng thể hiện phương sai đơn vị (ví dụ, trong một số hoặc tất cả các dải tần số). Để thực hiện mục đích này, dòng bit có thể bao gồm tham số độ lợi phương sai và bộ tạo hình phổ có thể được tạo cấu hình để áp dụng tham số độ lợi phương sai vào khối hiện tại của các hệ số biến đổi được ước lượng. Việc này có thể có lợi đối với chất lượng của việc dự đoán.

Theo một khía cạnh khác nữa, bộ mã hóa tiếng nói dựa trên biến đổi được tạo cấu hình để mã hóa tín hiệu tiếng nói thành dòng bit được mô tả. Như đã nêu trên, bộ mã hóa có thể bao gồm bất kỳ đặc điểm và/hoặc các bộ phận liên quan đến bộ mã hóa được nêu trong bản mô tả này. Cụ thể là, bộ mã hóa có thể bao gồm đơn vị tạo khung được tạo cấu hình để nhận nhiều khối liên tiếp gồm các hệ số biến đổi. Nhiều khối liên tiếp bao gồm khối hiện tại và một hoặc nhiều khối trước đó. Hơn nữa, nhiều khối liên tiếp chỉ báo các mẫu của tín hiệu tiếng nói.

Ngoài ra, bộ mã hóa tiếng nói có thể bao gồm đơn vị làm phẳng được tạo cấu hình để xác định khối hiện tại của các hệ số biến đổi được làm phẳng bằng cách làm phẳng khối hiện tại tương ứng của các hệ số biến đổi nhờ sử dụng đường bao khối hiện tại tương ứng (ví dụ, đường bao được điều chỉnh tương ứng). Hơn nữa, bộ mã hóa tiếng nói có thể bao gồm bộ dự đoán được tạo cấu hình để xác định khối hiện tại của các hệ số biến đổi được làm phẳng được ước lượng dựa trên một hoặc nhiều khối trước đó của các hệ số biến đổi được tái tạo và dựa trên một hoặc nhiều tham số dự đoán (bao gồm, ví dụ, độ lợi bộ dự đoán). Như đã nêu trên, một hoặc nhiều khối trước đó của các hệ số biến đổi được tái tạo có thể được lấy từ một hoặc nhiều khối trước đó của các hệ số biến đổi. Ngoài ra, bộ mã hóa tiếng nói có thể bao gồm đơn vị vi sai được tạo cấu hình để xác định khối hiện tại của các hệ số sai số dự đoán dựa trên khối

hiện tại của các hệ số biến đổi được làm phẳng và dựa trên khối hiện tại của các hệ số biến đổi được làm phẳng được ước lượng.

Bộ dự đoán có thể được tạo cấu hình để xác định khối hiện tại của các hệ số biến đổi được làm phẳng được ước lượng nhờ sử dụng tiêu chí sai số quân phương được gia trọng (ví dụ, bằng cách tối thiểu hóa tiêu chí sai số quân phương được gia trọng). Tiêu chí sai số quân phương được gia trọng có thể xem xét đường bao khối hiện tại hoặc một số hàm định trước của đường bao khối hiện tại làm các gia trọng. Trong bản mô tả này, nhiều khác nhau để xác định độ lợi bộ dự đoán sử dụng tiêu chí sai số quân phương được gia trọng được mô tả.

Hơn nữa, bộ mã hóa tiếng nói có thể bao gồm đơn vị lượng tử hóa hệ số được tạo cấu hình để lượng tử hóa các hệ số được lấy từ khối hiện tại của các hệ số sai số dự đoán, sử dụng tập hợp của các bộ lượng tử hóa được định trước. Đơn vị lượng tử hóa hệ số có thể được tạo cấu hình để xác định tập hợp của các bộ lượng tử hóa được định trước phụ thuộc vào ít nhất là một trong số một hoặc nhiều tham số dự đoán. Điều này có nghĩa là hoạt động của bộ dự đoán có thể có tác động lên các bộ lượng tử hóa được sử dụng bởi đơn vị lượng tử hóa hệ số. Đơn vị lượng tử hóa hệ số có thể được tạo cấu hình để xác định dữ liệu hệ số cho dòng bit dựa trên các hệ số được lượng tử hóa. Như thế, dữ liệu hệ số có thể chỉ báo phiên bản được lượng tử hóa của khối hiện tại của các hệ số sai số dự đoán.

Bộ mã hóa tiếng nói dựa trên biến đổi còn có thể bao gồm đơn vị lập thang tỷ lệ được tạo cấu hình để xác định khối hiện tại của các hệ số sai số được lập thang tỷ lệ dựa trên khối hiện tại của các hệ số sai số dự đoán nhờ sử dụng một hoặc nhiều nguyên tắc lập thang tỷ lệ. Khối hiện tại của hệ số sai số được lập thang tỷ lệ có thể được xác định như thế và/hoặc một hoặc nhiều nguyên tắc lập thang tỷ lệ có thể là sao cho trung bình thì phương sai của các hệ số sai số được lập thang tỷ lệ của khối hiện tại của các hệ số sai số được lập thang tỷ lệ cao hơn phương sai của các hệ số sai số dự đoán của khối hiện tại của các hệ số sai số dự đoán. Cụ thể là, một hoặc nhiều nguyên tắc lập thang tỷ lệ có thể là sao cho phương sai của các hệ số sai số dự đoán gần thống nhất với tất cả các bin tần số hoặc dải tần số. Đơn vị lượng tử hóa hệ số có thể được tạo cấu hình để lượng tử hóa các hệ số sai số được lập thang tỷ lệ của khối hiện tại của các hệ số sai số được lập thang tỷ lệ, để cung cấp dữ liệu hệ số.

Khối hiện tại của các hệ số sai số dự đoán thường bao gồm nhiều hệ số sai số dự đoán đối với nhiều bin tần số tương ứng. Các độ lợi lập thang tỷ lệ mà được áp dụng bởi đơn vị lập thang tỷ lệ vào các hệ số sai số dự đoán theo nguyên tắc lập thang tỷ lệ có thể phụ thuộc vào các bin tần số của các hệ số sai số dự đoán tương ứng. Hơn nữa, nguyên tắc lập thang tỷ lệ có thể phụ thuộc vào một hoặc nhiều tham số dự đoán, ví dụ phụ thuộc vào độ lợi bộ dự đoán. Lựa chọn khác hoặc thêm vào đó, nguyên tắc lập thang tỷ lệ có thể phụ thuộc vào đường bao khối hiện tại. Trong bản mô tả này, nhiều khác nhau để xác định nguyên tắc lập thang tỷ lệ phụ thuộc vào bin tần số được mô tả.

Bộ mã hóa tiếng nói dựa trên biến đổi còn có thể bao gồm đơn vị phân bố bit được tạo cấu hình để xác định vector phân bố dựa trên đường bao khối hiện tại. Vector phân bố có thể chỉ báo bộ lượng tử hóa thứ nhất từ tập hợp của các bộ lượng tử hóa được định trước sẽ được sử dụng để lượng tử hóa hệ số thứ nhất được lấy từ khối hiện tại của các hệ số sai số dự đoán. Cụ thể là, vector phân bố có thể chỉ báo các bộ lượng tử hóa sẽ được sử dụng để lượng tử hóa tất cả các hệ số được lấy từ khối hiện tại của các hệ số sai số dự đoán một cách lần lượt. Ví dụ, vector phân bố có thể chỉ báo bộ lượng tử hóa khác sẽ được sử dụng cho mỗi dải tần số.

Đơn vị phân bố bit có thể được tạo cấu hình để xác định vector phân bố sao cho dữ liệu hệ số đối với khối hiện tại của các hệ số sai số dự đoán không vượt quá số lượng bit định trước. Hơn nữa, đơn vị phân bố bit có thể được tạo cấu hình để xác định trị số dịch vị chỉ báo độ dịch vị sẽ được áp dụng vào đường bao phân bố được lấy từ đường bao khối hiện tại (ví dụ, được lấy từ đường bao được điều chỉnh hiện tại). Trị số dịch vị có thể được đưa vào dòng bit để làm cho bộ giải mã tương ứng xác nhận các bộ lượng tử hóa mà đã được sử dụng để xác định dữ liệu hệ số. Theo khía cạnh khác, bộ giải mã tiếng nói dựa trên biến đổi được tạo cấu hình để giải mã dòng bit nhằm cung cấp tín hiệu tiếng nói được tái tạo được mô tả. Bộ giải mã tiếng nói có thể bao gồm bất kỳ đặc điểm và/hoặc các bộ phận được nêu trong bản mô tả này. Cụ thể là, bộ giải mã có thể bao gồm bộ dự đoán được tạo cấu hình để xác định khối hiện tại của các hệ số biến đổi được làm phẳng được ước lượng dựa trên một hoặc nhiều khối trước đó của các hệ số biến đổi được tái tạo và dựa trên một hoặc nhiều tham số dự đoán được lấy từ dòng bit. Hơn nữa, bộ giải mã tiếng nói có thể bao gồm bộ giải mã phổ được tạo cấu hình để xác định khối hiện tại của các hệ số sai số dự đoán được

lượng tử hóa (hoặc phiên bản được lập thang tỷ lệ của nó) dựa trên dữ liệu hệ số được bao gồm trong dòng bit, sử dụng tập hợp của các bộ lượng tử hóa được định trước. Cụ thể là, bộ giải mã phổ có thể sử dụng tập hợp của các bộ đảo ngược lượng tử hóa được định trước tương ứng với tập hợp được định trước của các bộ lượng tử hóa được sử dụng bởi bộ mã hóa tiếng nói tương ứng.

Bộ giải mã phổ có thể được tạo cấu hình để xác định tập hợp của các bộ lượng tử hóa được định trước (và/hoặc tập hợp tương ứng của các bộ đảo ngược lượng tử hóa được định trước) phụ thuộc vào một hoặc nhiều tham số dự đoán. Cụ thể là, bộ giải mã phổ có thể thực hiện cùng một quá trình chọn lọc đối với tập hợp của các bộ lượng tử hóa được định trước như đơn vị lượng tử hóa hệ số của bộ mã hóa tiếng nói tương ứng. Bằng cách khiến cho tập hợp của các bộ lượng tử hóa được định trước phụ thuộc vào một hoặc nhiều tham số dự đoán, chất lượng cảm quan của tín hiệu tiếng nói được tái tạo có thể được cải thiện.

Tập hợp của các bộ lượng tử hóa được định trước có thể bao gồm các bộ lượng tử hóa khác nhau với các tỷ số tín hiệu trên nhiễu khác nhau (và các tốc độ bit kết hợp khác nhau). Hơn nữa, tập hợp của các bộ lượng tử hóa được định trước có thể bao gồm ít nhất một bộ lượng tử hóa hòa sắc. Một hoặc nhiều tham số dự đoán có thể bao gồm độ lợi bộ dự đoán g . Độ lợi bộ dự đoán g có thể chỉ báo độ chính xác của một hoặc nhiều khối trước đó của các hệ số biến đổi được tái tạo đối với khối hiện tại của các hệ số biến đổi được tái tạo. Như thế, độ lợi bộ dự đoán g có thể cung cấp chỉ số của lượng thông tin được bao gồm trong khối hiện tại của các hệ số sai số dự đoán. Độ lợi bộ dự đoán tương đối cao g có thể chỉ báo lượng thông tin tương đối thấp, và ngược lại. Một số bộ lượng tử hóa hòa sắc được bao gồm trong tập hợp của các bộ lượng tử hóa được định trước có thể phụ thuộc vào độ lợi bộ dự đoán. Cụ thể là, số lượng của các bộ lượng tử hóa hòa sắc được bao gồm trong tập hợp của các bộ lượng tử hóa được định trước có thể giảm đi cùng với sự tăng độ lợi bộ dự đoán.

Bộ giải mã phổ có thể truy cập vào tập hợp thứ nhất và thứ hai của các bộ lượng tử hóa được định trước. Tập hợp thứ hai có thể bao gồm lượng ít hơn của các bộ lượng tử hóa hòa sắc so với tập hợp thứ nhất của các bộ lượng tử hóa. Bộ giải mã phổ có thể được tạo cấu hình để xác định tiêu chí tập hợp rfu dựa trên độ lợi bộ dự đoán g . Bộ giải mã phổ có thể được tạo cấu hình để sử dụng tập hợp thứ nhất của các bộ lượng tử hóa được định trước nếu tiêu chí tập hợp rfu nhỏ hơn ngưỡng định trước. Hơn nữa,

bộ giải mã phổ có thể được tạo cấu hình để sử dụng tập hợp thứ hai của các bộ lượng tử hóa được định trước nếu tiêu chí tập hợp rfu lớn hơn hoặc bằng ngưỡng định trước. Tiêu chí tập hợp có thể là $rfu = \min(1, \max(g, 0))$, trong đó độ lợi bộ dự đoán là g . Tiêu chí tập hợp rfu này chấp nhận các trị số lớn hơn hoặc bằng không và nhỏ hơn hoặc bằng một. Ngưỡng định trước có thể là 0,75.

Như nêu trên, tiêu chí tập hợp có thể phụ thuộc vào tham số điều chỉnh được định trước, rfu. Trong ví dụ thay thế, tham số điều chỉnh rfu có thể được xác định nhờ sử dụng các điều kiện sau: $rfu = 1,0$ khi $g < -1,0$; $rfu = -g$ khi $-1,0 \leq g < 0,0$; $rfu = g$ for $0,0 \leq g < 1,0$; $rfu = 2,0 - g$ khi $1,0 \leq g < 2,0$; và/hoặc $rfu = 0,0$ khi $g \geq 2,0$.

Ngoài ra, bộ giải mã tiếng nói có thể bao gồm đơn vị cộng được tạo cấu hình để xác định khối hiện tại của các hệ số biến đổi đã làm phẳng được tái tạo dựa trên khối hiện tại của các hệ số biến đổi được làm phẳng được ước lượng và dựa trên khối hiện tại của các hệ số sai số dự đoán được lượng tử hóa. Hơn nữa, bộ giải mã tiếng nói có thể bao gồm đơn vị làm phẳng ngược được tạo cấu hình để xác định khối hiện tại của các hệ số biến đổi được tái tạo bằng cách cung cấp khối hiện tại của các hệ số biến đổi đã làm phẳng được tái tạo với dạng phổ, sử dụng đường bao khối hiện tại. Tín hiệu tiếng nói được tái tạo có thể được xác định dựa trên khối hiện tại của các hệ số biến đổi được tái tạo (ví dụ, nhờ sử dụng đơn vị biến đổi ngược).

Bộ giải mã tiếng nói dựa trên biến đổi có thể bao gồm đơn vị đảo ngược lập thang tỷ lệ được tạo cấu hình để lập thang tỷ lệ các hệ số sai số dự đoán được lượng tử hóa của khối hiện tại của các hệ số sai số dự đoán được lượng tử hóa nhờ sử dụng nguyên tắc đảo ngược lập thang tỷ lệ, để cung cấp khối hiện tại của các hệ số sai số dự đoán được tái lập thang tỷ lệ. Các độ lợi lập thang tỷ lệ mà được áp dụng bởi đơn vị đảo ngược lập thang tỷ lệ vào các hệ số sai số dự đoán được lượng tử hóa theo nguyên tắc đảo ngược lập thang tỷ lệ có thể phụ thuộc vào các bin tần số của các hệ số sai số dự đoán tương ứng được lượng tử hóa. Nói cách khác, nguyên tắc đảo ngược lập thang tỷ lệ có thể phụ thuộc vào tần số, tức là các độ lợi lập thang tỷ lệ phụ thuộc vào tần số. Nguyên tắc đảo ngược lập thang tỷ lệ có thể được tạo cấu hình để điều chỉnh phương sai của các hệ số sai số dự đoán được lượng tử hóa đối với các bin tần số khác nhau.

Nguyên tắc đảo ngược lập thang tỷ lệ is thường ngược lại so với nguyên tắc lập thang tỷ lệ được áp dụng bởi đơn vị lập thang tỷ lệ của bộ mã hóa tiếng nói tương ứng dựa trên biến đổi. Do đó, các khía cạnh mà được mô tả ở đây về việc xác định và các đặc tính của nguyên tắc lập thang tỷ lệ, cũng có thể áp dụng được (theo cách tương tự) cho nguyên tắc đảo ngược lập thang tỷ lệ.

Đơn vị cộng khi đó có thể được tạo cấu hình để xác định khối hiện tại của các hệ số biến đổi đã làm phẳng được tái tạo bằng cách cộng khối hiện tại của các hệ số sai số dự đoán được tái lập thang tỷ lệ vào khối hiện tại của các hệ số biến đổi được làm phẳng được ước lượng.

Một hoặc nhiều tham số điều chỉnh có thể bao gồm cờ bảo toàn phương sai. Cờ bảo toàn phương sai có thể chỉ báo cách phương sai của khối hiện tại của các hệ số sai số dự đoán được lượng tử hóa sẽ được tạo hình. Nói cách khác, cờ bảo toàn phương sai có thể chỉ báo cách xử lý mà sẽ được thực hiện bởi bộ giải mã, mà có ảnh hưởng đến phương sai của khối hiện tại của các hệ số sai số dự đoán được lượng tử hóa.

Ví dụ, tập hợp của các bộ lượng tử hóa được định trước có thể được xác định phụ thuộc vào cờ bảo toàn phương sai. Cụ thể là, tập hợp của các bộ lượng tử hóa được định trước có thể bao gồm bộ lượng tử hóa tổng hợp nhiều. Độ lợi nhiều của bộ lượng tử hóa tổng hợp nhiều có thể phụ thuộc vào cờ bảo toàn phương sai. Lựa chọn khác hoặc thêm vào đó, tập hợp của các bộ lượng tử hóa được định trước bao gồm một hoặc nhiều bộ lượng tử hóa hòa sắc bao phủ khoảng SNR. Khoảng SNR có thể được xác định phụ thuộc vào cờ bảo toàn phương sai. Ít nhất là một trong số một hoặc nhiều bộ lượng tử hóa hòa sắc có thể được tạo cấu hình để áp dụng độ lợi sau γ , khi xác định hệ số sai số dự đoán được lượng tử hóa. Độ lợi sau γ có thể phụ thuộc vào cờ bảo toàn phương sai.

Bộ giải mã tiếng nói dựa trên biến đổi có thể bao gồm đơn vị đảo ngược lập thang tỷ lệ được tạo cấu hình để lập thang tỷ lệ các hệ số sai số dự đoán được lượng tử hóa của khối hiện tại của các hệ số sai số dự đoán được lượng tử hóa, để cung cấp khối hiện tại của các hệ số sai số dự đoán được tái lập thang tỷ lệ. Đơn vị cộng có thể được tạo cấu hình để xác định khối hiện tại của các hệ số biến đổi đã làm phẳng được tái tạo bằng cách cộng khối hiện tại của các hệ số sai số dự đoán được tái lập thang tỷ lệ hoặc bằng cách cộng khối hiện tại của các hệ số sai số dự đoán được lượng tử hóa

vào khối hiện tại của các hệ số biến đổi được làm phẳng được ước lượng, tùy thuộc vào cờ bảo toàn phương sai.

Cờ bảo toàn phương sai có thể được sử dụng để điều chỉnh mức độ ồn của các bộ lượng tử hóa đến chất lượng của sự dự đoán. Kết quả của việc này là, chất lượng cảm quan của bộ mã hóa-giải mã có thể được cải thiện.

Theo khía cạnh khác, bộ mã hóa âm thanh dựa trên biến đổi được mô tả. Bộ mã hóa âm thanh được tạo cấu hình để mã hóa tín hiệu âm thanh bao gồm phân đoạn thứ nhất (ví dụ, đoạn tiếng nói) vào dòng bit. Cụ thể là, bộ mã hóa âm thanh có thể được tạo cấu hình để mã hóa một hoặc nhiều đoạn tiếng nói của tín hiệu âm thanh nhờ sử dụng bộ mã hóa tiếng nói dựa trên biến đổi. Hơn nữa, bộ mã hóa âm thanh có thể được tạo cấu hình để mã hóa một hoặc nhiều phân đoạn phi tiếng nói của tín hiệu âm thanh nhờ sử dụng bộ mã hóa âm thanh dựa trên biến đổi chung.

Bộ mã hóa âm thanh có thể bao gồm bộ phân loại tín hiệu được tạo cấu hình để nhận biết phân đoạn thứ nhất (ví dụ, đoạn tiếng nói) từ tín hiệu âm thanh. Khái quát hơn, bộ phân loại tín hiệu có thể được tạo cấu hình để xác định phân đoạn từ tín hiệu âm thanh mà sẽ được mã hóa bởi bộ mã hóa tiếng nói dựa trên biến đổi. Phân đoạn thứ nhất được xác định có thể được gọi là đoạn tiếng nói (mặc dù phân đoạn này không cần thiết phải chứa tiếng nói thực). Cụ thể là, bộ phân loại tín hiệu có thể được tạo cấu hình để phân loại các phân đoạn khác nhau (ví dụ, khung hay khối) của tín hiệu âm thanh thành tiếng nói hoặc phi tiếng nói. Như đã nêu trên, khối hệ số biến đổi có thể bao gồm nhiều hệ số biến đổi đối với nhiều bin tần số tương ứng. Hơn nữa, bộ mã hóa âm thanh có thể bao gồm đơn vị biến đổi được tạo cấu hình để xác định nhiều khối liên tiếp gồm các hệ số biến đổi dựa trên phân đoạn thứ nhất. Đơn vị biến đổi có thể được tạo cấu hình để biến đổi các đoạn tiếng nói và các phân đoạn phi tiếng nói.

Đơn vị biến đổi có thể được tạo cấu hình để xác định các khối dài bao gồm số thứ nhất của các hệ số biến đổi và các khối ngắn bao gồm số thứ hai của các hệ số biến đổi. Số thứ nhất của các mẫu có thể lớn hơn số thứ hai của các mẫu. Cụ thể là, số thứ nhất của các mẫu có thể là 1024 và số thứ hai của các mẫu có thể là 256. Các khối của nhiều khối liên tiếp có thể là các khối ngắn. Cụ thể là, bộ mã hóa âm thanh có thể được tạo cấu hình để biến đổi tất cả các phân đoạn của tín hiệu âm thanh, mà đã được phân loại là tiếng nói, thành các khối ngắn.

Hơn nữa, bộ mã hóa âm thanh có thể bao gồm bộ mã hóa tiếng nói dựa trên biến đổi (như được nêu trong bản mô tả này) được tạo cấu hình để mã hóa nhiều khối liên tiếp vào dòng bit. Ngoài ra, bộ mã hóa âm thanh có thể bao gồm bộ mã hóa âm thanh dựa trên biến đổi chung được tạo cấu hình để mã hóa phân đoạn của tín hiệu âm thanh hơn là phân đoạn thứ nhất (ví dụ, phân đoạn phi tiếng nói). Bộ mã hóa âm thanh dựa trên biến đổi chung có thể là bộ mã hóa AAC (Advanced Audio Coder - mã hóa âm thanh nâng cao) hoặc HE (High Efficiency – hiệu suất cao)-AAC. Như đã nêu trên, đơn vị biến đổi có thể được tạo cấu hình để thực hiện MDCT. Như thế, bộ mã hóa âm thanh có thể được tạo cấu hình để mã hóa tín hiệu âm thanh đầu vào hoàn thiện (bao gồm các đoạn tiếng nói và các phân đoạn phi tiếng nói) trong miền biến đổi (nhờ sử dụng một đơn vị biến đổi).

Theo khía cạnh khác, bộ giải mã âm thanh dựa trên biến đổi tương ứng được tạo cấu hình để giải mã dòng bit chỉ báo tín hiệu âm thanh bao gồm đoạn tiếng nói (tức là phân đoạn mà đã được mã hóa nhờ sử dụng bộ mã hóa tiếng nói dựa trên biến đổi) được mô tả. Bộ giải mã âm thanh có thể bao gồm bộ giải mã tiếng nói dựa trên biến đổi được tạo cấu hình để xác định nhiều khối liên tiếp gồm các hệ số biến đổi được tái tạo dựa trên dữ liệu (ví dụ, dữ liệu đường bao, dữ liệu độ lợi, bộ dự đoán dữ liệu và dữ liệu hệ số) được bao gồm trong dòng bit. Hơn nữa, dòng bit có thể chỉ báo rằng dữ liệu nhận được sẽ được giải mã nhờ sử dụng bộ giải mã tiếng nói.

Ngoài ra, bộ giải mã âm thanh có thể bao gồm đơn vị biến đổi ngược được tạo cấu hình để xác định đoạn tiếng nói được tái tạo dựa trên nhiều khối liên tiếp gồm các hệ số biến đổi được tái tạo. Khối hệ số biến đổi được tái tạo có thể bao gồm nhiều hệ số biến đổi được tái tạo đối với nhiều bin tần số tương ứng. Đơn vị biến đổi ngược có thể được tạo cấu hình để xử lý các khối dài bao gồm số thứ nhất của các hệ số biến đổi được tái tạo và các khối ngắn bao gồm số thứ hai của các hệ số biến đổi được tái tạo. Số thứ nhất của các mẫu có thể lớn hơn số thứ hai của các mẫu. Các khối của nhiều khối liên tiếp có thể là các khối ngắn.

Theo một khía cạnh khác nữa, phương pháp mã hóa tín hiệu tiếng nói thành dòng bit được mô tả. Phương pháp này có thể bao gồm việc nhận tập hợp khối. Tập hợp các khối có thể bao gồm nhiều khối liên tiếp gồm các hệ số biến đổi. Nhiều khối liên tiếp có thể chỉ báo các mẫu của tín hiệu tiếng nói. Hơn nữa, khối hệ số biến đổi có thể bao gồm nhiều hệ số biến đổi đối với nhiều bin tần số tương ứng. Phương pháp có

thể tiếp tục xác định đường bao hiện thời dựa trên nhiều khối liên tiếp gồm các hệ số biến đổi. Đường bao hiện thời có thể chỉ báo nhiều trị số năng lượng phổ đối với nhiều bin tần số tương ứng. Hơn nữa, phương pháp này có thể bao gồm việc xác định nhiều đường bao được nội suy đối với nhiều khối hệ số biến đổi một cách lần lượt, dựa trên đường bao hiện thời. Ngoài ra, phương pháp này có thể bao gồm việc xác định nhiều khối hệ số biến đổi được làm phẳng bằng cách làm phẳng nhiều khối tương ứng của các hệ số biến đổi nhờ sử dụng nhiều đường bao được nội suy tương ứng một cách lần lượt. Dòng bit có thể được xác định dựa trên nhiều khối hệ số biến đổi được làm phẳng.

Theo khía cạnh khác, phương pháp giải mã dòng bit nhằm cung cấp tín hiệu tiếng nói được tái tạo được mô tả. Phương pháp này có thể bao gồm việc xác định đường bao hiện thời được lượng tử hóa từ dữ liệu đường bao được bao gồm trong dòng bit. Đường bao hiện thời được lượng tử hóa có thể chỉ báo nhiều trị số năng lượng phổ đối với nhiều bin tần số tương ứng. Dòng bit có thể bao gồm dữ liệu (ví dụ, dữ liệu hệ số và/hoặc dữ liệu bộ dự đoán) chỉ báo nhiều khối liên tiếp của các hệ số biến đổi đã làm phẳng được tái tạo. Khối hệ số biến đổi đã làm phẳng được tái tạo có thể bao gồm nhiều hệ số biến đổi đã làm phẳng được tái tạo đối với nhiều bin tần số tương ứng. Hơn nữa, phương pháp này có thể bao gồm việc xác định nhiều đường bao được nội suy đối với nhiều khối hệ số biến đổi đã làm phẳng được tái tạo một cách lần lượt, dựa trên đường bao hiện thời được lượng tử hóa. Phương pháp có thể tiếp tục xác định nhiều khối hệ số biến đổi được tái tạo bằng cách cung cấp nhiều khối tương ứng của các hệ số biến đổi đã làm phẳng được tái tạo với dạng phổ, sử dụng nhiều đường bao được nội suy tương ứng một cách lần lượt. Tín hiệu tiếng nói được tái tạo có thể dựa trên nhiều khối hệ số biến đổi được tái tạo.

Theo khía cạnh khác, phương pháp mã hóa tín hiệu tiếng nói thành dòng bit được mô tả. Phương pháp này có thể bao gồm việc nhận nhiều khối liên tiếp gồm các hệ số biến đổi bao gồm khối hiện tại và một hoặc nhiều khối trước đó. Nhiều khối liên tiếp có thể chỉ báo các mẫu của tín hiệu tiếng nói. Phương pháp có thể tiếp tục xác định khối hiện tại và một hoặc nhiều khối trước đó của các hệ số biến đổi được làm phẳng bằng cách làm phẳng khối hiện tại tương ứng và một hoặc nhiều khối trước đó tương ứng của các hệ số biến đổi nhờ sử dụng đường bao khối hiện tại tương ứng và một hoặc nhiều đường bao khối trước đó tương ứng một cách lần lượt.

Hơn nữa, phương pháp này có thể bao gồm việc xác định khối hiện tại của các hệ số biến đổi được làm phẳng được ước lượng dựa trên một hoặc nhiều khối trước đó của các hệ số biến đổi được tái tạo và dựa trên tham số bộ dự đoán. Có thể đạt được điều này nhờ sử dụng các kỹ thuật dự đoán. Một hoặc nhiều khối trước đó của các hệ số biến đổi được tái tạo có thể được lấy từ một hoặc nhiều khối trước đó của các hệ số biến đổi được làm phẳng một cách lần lượt. Bước xác định khối hiện tại của các hệ số biến đổi được làm phẳng được ước lượng có thể bao gồm việc xác định khối hiện tại của các hệ số biến đổi được ước lượng dựa trên một hoặc nhiều khối trước đó của các hệ số biến đổi được tái tạo và dựa trên tham số bộ dự đoán, và xác định khối hiện tại của các hệ số biến đổi được làm phẳng được ước lượng dựa trên khối hiện tại của các hệ số biến đổi được ước lượng, dựa trên một hoặc nhiều đường bao khối trước đó và dựa trên tham số bộ dự đoán.

Hơn nữa, phương pháp này có thể bao gồm việc xác định khối hiện tại của các hệ số sai số dự đoán dựa trên khối hiện tại của các hệ số biến đổi được làm phẳng và dựa trên khối hiện tại của các hệ số biến đổi được làm phẳng được ước lượng. Dòng bit có thể được xác định dựa trên khối hiện tại của các hệ số sai số dự đoán.

Theo một khía cạnh khác nữa, phương pháp giải mã dòng bit nhằm cung cấp tín hiệu tiếng nói được tái tạo được mô tả. Phương pháp này có thể bao gồm việc xác định khối hiện tại của các hệ số biến đổi được làm phẳng được ước lượng dựa trên một hoặc nhiều khối trước đó của các hệ số biến đổi được tái tạo và dựa trên tham số bộ dự đoán được lấy từ dòng bit. Bước xác định khối hiện tại của các hệ số biến đổi được làm phẳng được ước lượng có thể bao gồm việc xác định khối hiện tại của các hệ số biến đổi được ước lượng dựa trên một hoặc nhiều khối trước đó của các hệ số biến đổi được tái tạo và dựa trên tham số bộ dự đoán; và xác định khối hiện tại của các hệ số biến đổi được làm phẳng được ước lượng dựa trên khối hiện tại của các hệ số biến đổi được ước lượng, dựa trên một hoặc nhiều đường bao khối trước đó và dựa trên tham số bộ dự đoán.

Ngoài ra, phương pháp này có thể bao gồm việc xác định khối hiện tại của các hệ số sai số dự đoán được lượng tử hóa dựa trên dữ liệu hệ số được bao gồm trong dòng bit. Phương pháp có thể tiếp tục xác định khối hiện tại của các hệ số biến đổi đã làm phẳng được tái tạo dựa trên khối hiện tại của các hệ số biến đổi được làm phẳng được ước lượng và dựa trên khối hiện tại của các hệ số sai số dự đoán được lượng tử

hóa. Khối hiện tại của các hệ số biến đổi được tái tạo có thể được xác định bằng cách cung cấp khối hiện tại của các hệ số biến đổi đã làm phẳng được tái tạo với dạng phổ, sử dụng đường bao khối hiện tại (ví dụ, đường bao được điều chỉnh hiện tại). Hơn nữa, một hoặc nhiều khối trước đó của các hệ số biến đổi được tái tạo có thể được xác định bằng cách cung cấp một hoặc nhiều khối trước đó của các hệ số biến đổi đã làm phẳng được tái tạo với dạng phổ, sử dụng một hoặc nhiều đường bao khối trước đó (ví dụ, một hoặc nhiều đường bao được điều chỉnh trước đó) một cách lần lượt. Ngoài ra, phương pháp này có thể bao gồm việc xác định tín hiệu tiếng nói được tái tạo dựa trên khối hiện tại và một hoặc nhiều khối trước đó của các hệ số biến đổi được tái tạo.

Theo một khía cạnh khác nữa, phương pháp mã hóa tín hiệu tiếng nói thành dòng bit được mô tả. Phương pháp này có thể bao gồm việc nhận nhiều khối liên tiếp gồm các hệ số biến đổi bao gồm khối hiện tại và một hoặc nhiều khối trước đó. Nhiều khối liên tiếp có thể chỉ báo các mẫu của tín hiệu tiếng nói. Hơn nữa, phương pháp này có thể bao gồm việc xác định khối hiện tại của các hệ số biến đổi được ước lượng dựa trên một hoặc nhiều khối trước đó của các hệ số biến đổi được tái tạo và dựa trên tham số bộ dự đoán. Một hoặc nhiều khối trước đó của các hệ số biến đổi được tái tạo có thể được lấy từ một hoặc nhiều khối trước đó của các hệ số biến đổi. Phương pháp có thể tiếp tục xác định khối hiện tại của các hệ số sai số dự đoán dựa trên khối hiện tại của các hệ số biến đổi và dựa trên khối hiện tại của các hệ số biến đổi được ước lượng. Hơn nữa, phương pháp này có thể bao gồm các hệ số lượng tử hóa được lấy từ khối hiện tại của các hệ số sai số dự đoán, sử dụng tập hợp của các bộ lượng tử hóa được định trước. Tập hợp của các bộ lượng tử hóa được định trước có thể phụ thuộc vào tham số bộ dự đoán. Hơn nữa, phương pháp này có thể bao gồm việc xác định dữ liệu hệ số đối với dòng bit dựa trên các hệ số được lượng tử hóa.

Theo khía cạnh khác, phương pháp giải mã dòng bit nhằm cung cấp tín hiệu tiếng nói được tái tạo được mô tả. Phương pháp này có thể bao gồm việc xác định khối hiện tại của các hệ số biến đổi được ước lượng dựa trên một hoặc nhiều khối trước đó của các hệ số biến đổi được tái tạo và dựa trên tham số bộ dự đoán được lấy từ dòng bit. Hơn nữa, phương pháp này có thể bao gồm việc xác định khối hiện tại của các hệ số sai số dự đoán được lượng tử hóa dựa trên dữ liệu hệ số được bao gồm trong dòng bit, sử dụng tập hợp của các bộ lượng tử hóa được định trước. Tập hợp của các bộ lượng tử hóa được định trước có thể là hàm số của tham số bộ dự đoán. Phương

pháp có thể tiếp tục xác định khối hiện tại của các hệ số biến đổi được tái tạo dựa trên khối hiện tại của các hệ số biến đổi được ước lượng và dựa trên khối hiện tại của các hệ số sai số dự đoán được lượng tử hóa. Tín hiệu tiếng nói được tái tạo có thể được xác định dựa trên khối hiện tại của các hệ số biến đổi được tái tạo.

Theo khía cạnh khác, phương pháp mã hóa tín hiệu âm thanh bao gồm đoạn tiếng nói vào dòng bit được mô tả. Phương pháp này có thể bao gồm việc nhận biết đoạn tiếng nói từ tín hiệu âm thanh. Hơn nữa, phương pháp này có thể bao gồm việc xác định nhiều khối liên tiếp gồm các hệ số biến đổi dựa trên đoạn tiếng nói, nhờ sử dụng đơn vị biến đổi. Đơn vị biến đổi có thể được tạo cấu hình để xác định các khối dài bao gồm số thứ nhất của các hệ số biến đổi và các khối ngắn bao gồm số thứ hai của các hệ số biến đổi. Số thứ nhất có thể lớn hơn số thứ hai. Các khối của nhiều khối liên tiếp có thể là các khối ngắn. Ngoài ra, phương pháp này có thể bao gồm việc mã hóa nhiều khối liên tiếp vào dòng bit.

Theo khía cạnh khác, phương pháp giải mã dòng bit chỉ báo tín hiệu âm thanh bao gồm đoạn tiếng nói được mô tả. Phương pháp này có thể bao gồm việc xác định nhiều khối liên tiếp gồm các hệ số biến đổi được tái tạo dựa trên dữ liệu được bao gồm trong dòng bit. Hơn nữa, phương pháp này có thể bao gồm việc xác định đoạn tiếng nói được tái tạo dựa trên nhiều khối liên tiếp gồm các hệ số biến đổi được tái tạo, sử dụng đơn vị biến đổi ngược. Đơn vị biến đổi ngược có thể được tạo cấu hình để xử lý các khối dài bao gồm số thứ nhất của các hệ số biến đổi được tái tạo và các khối ngắn bao gồm số thứ hai của các hệ số biến đổi được tái tạo. Số thứ nhất có thể lớn hơn số thứ hai. Các khối của nhiều khối liên tiếp có thể là các khối ngắn.

Theo một khía cạnh khác nữa, chương trình phần mềm được mô tả. Chương trình phần mềm có thể được điều chỉnh để việc thực thi trên bộ xử lý và để thực hiện các bước phương pháp được nêu trong bản mô tả này khi được thực hiện trên bộ xử lý.

Theo khía cạnh khác, phương tiện lưu trữ được mô tả. Phương tiện lưu trữ có thể bao gồm chương trình phần mềm được điều chỉnh để thực thi trên bộ xử lý và để thực hiện các bước phương pháp được nêu trong bản mô tả này khi được thực hiện trên bộ xử lý.

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

Cần lưu ý rằng các phương pháp và hệ thống bao gồm các phương án ưu tiên như được nêu trong bản mô tả sáng chế này có thể được sử dụng độc lập hoặc kết hợp với các phương pháp và hệ thống khác được bộc lộ trong phần mô tả này. Hơn nữa, tất cả các khía cạnh của các phương pháp và hệ thống được nêu trong bản mô tả sáng chế này có thể được kết hợp theo nhiều khác nhau. Cụ thể là, các đặc điểm của các yêu cầu bảo hộ có thể được kết hợp với nhau theo cách tùy ý.

Mô tả vắn tắt các hình vẽ

Sáng chế sẽ được giải thích dưới đây làm mẫu có tham chiếu đến các hình vẽ kèm theo, trong đó

Fig.1a thể hiện sơ đồ khối của bộ mã hóa âm thanh làm ví dụ mà cung cấp dòng bit tại tốc độ bit không đổi;

Fig.1b thể hiện sơ đồ khối của bộ mã hóa âm thanh làm ví dụ mà cung cấp dòng bit tại tốc độ bit biến đổi;

Fig.2 minh họa việc tạo ra đường bao làm ví dụ dựa trên nhiều khối hệ số biến đổi;

Fig.3a minh họa các đường bao làm ví dụ của các khối hệ số biến đổi;

Fig.3b minh họa việc xác định đường bao được nội suy làm ví dụ;

Fig.4 minh họa các tập hợp làm ví dụ của các bộ lượng tử hóa;

Fig.5a thể hiện sơ đồ khối của bộ giải mã âm thanh làm ví dụ;

Fig.5b thể hiện sơ đồ khối của bộ giải mã đường bao làm ví dụ của bộ giải mã âm thanh trên Fig.5a;

Fig.5c thể hiện sơ đồ khối của bộ dự đoán dải phụ làm ví dụ của bộ giải mã âm thanh trên Fig.5a; và

Fig.5d thể hiện sơ đồ khối của bộ giải mã phổ làm ví dụ của bộ giải mã âm thanh trên Fig.5a.

Mô tả chi tiết sáng chế

Như được nêu trong phần bản chất kỹ thuật của sáng chế, cần có bộ mã hóa-giải mã âm thanh dựa trên biến đổi mà thể hiện các độ lợi mã hóa tương đối cao đối với tiếng nói hoặc các tín hiệu thoại. Bộ mã hóa-giải mã âm thanh dựa trên biến đổi này có thể được gọi là bộ mã hóa-giải mã tiếng nói dựa trên biến đổi hoặc bộ mã hóa-giải mã giọng nói dựa trên biến đổi. Bộ mã hóa-giải mã tiếng nói dựa trên biến đổi có thể được kết hợp một cách thuận lợi với bộ mã hóa-giải mã âm thanh dựa trên biến đổi chung, như AAC hoặc HE-AAC, do nó cũng hoạt động trong miền biến đổi. Hơn nữa, việc phân loại phân đoạn (ví dụ, khung) của tín hiệu âm thanh đầu vào thành tiếng nói hoặc phi tiếng nói, và trình tự chuyển đổi giữa bộ mã hóa-giải mã âm thanh chung và bộ mã hóa-giải mã âm thanh cụ thể có thể được đơn giản hóa, do cả hai bộ mã hóa-giải mã đều hoạt động trong miền biến đổi.

Fig.1a thể hiện sơ đồ khối của bộ mã hóa tiếng nói dựa trên biến đổi làm ví dụ 100. Bộ mã hóa 100 nhận làm khối đầu vào 131 của các hệ số biến đổi (còn được gọi là đơn vị mã hóa). Khối 131 của hệ số biến đổi có thể được thu bởi đơn vị biến đổi được tạo cấu hình để biến đổi trình tự của các mẫu của tín hiệu âm thanh đầu vào từ miền thời gian thành miền biến đổi. Đơn vị biến đổi có thể được tạo cấu hình để thực hiện MDCT. Đơn vị biến đổi có thể là một bộ phận của bộ mã hóa-giải mã âm thanh chung như AAC hoặc HE-AAC. Bộ mã hóa-giải mã âm thanh chung này có thể sử dụng các kích thước khối khác nhau, ví dụ khối dài và khối ngắn. Các kích thước khối làm ví dụ là 1024 mẫu đối với khối dài và 256 mẫu đối với khối ngắn. Giả sử tốc độ lấy mẫu là 44,1kHz và chồng lấp 50%, khối dài bao phủ gần khoảng 20ms của tín hiệu âm thanh đầu vào và khối ngắn bao phủ gần khoảng 5ms của tín hiệu âm thanh đầu vào. Các khối dài thường được sử dụng cho các phân đoạn cố định của tín hiệu âm thanh đầu vào và các khối ngắn thường được sử dụng cho các phân đoạn tạm thời của tín hiệu âm thanh đầu vào.

Các tín hiệu tiếng nói có thể được xem là ổn định trong các phân đoạn thời gian khoảng 20ms. Cụ thể là, đường bao phổ của tín hiệu tiếng nói có thể được xem là ổn định trong các phân đoạn thời gian khoảng 20ms. Để có thể lấy các số liệu có ý nghĩa trong miền biến đổi đối với các phân đoạn 20ms như thế, có thể hữu dụng khi trang bị cho bộ mã hóa tiếng nói dựa trên biến đổi 100 các khối ngắn 131 của các hệ số biến đổi (có độ dài là, ví dụ, 5ms). Bằng cách này, nhiều khối ngắn 131 có thể được sử dụng để lấy số liệu về phân đoạn thời gian là, ví dụ, 20ms (ví dụ, phân đoạn thời gian

của khối dài hoặc khung). Hơn nữa, điều này có lợi thế là cung cấp độ phân giải thời gian thích hợp đối với các tín hiệu tiếng nói.

Do đó, đơn vị biến đổi có thể được tạo cấu hình để cung cấp các khối ngắn 131 của các hệ số biến đổi, nếu phân đoạn hiện tại của tín hiệu âm thanh đầu vào được phân loại là tiếng nói. Bộ mã hóa 100 có thể bao gồm đơn vị tạo khung 101 được tạo cấu hình để tách nhiều khối 131 của các hệ số biến đổi, còn gọi là tập hợp 132 của các khối 131. Tập hợp 132 của các khối có thể còn được gọi là khung. Ví dụ, tập hợp 132 của các khối 131 có thể bao gồm bốn khối ngắn của 256 hệ số biến đổi, qua đó bao phủ gần khoảng phân đoạn 20ms của tín hiệu âm thanh đầu vào.

Bộ mã hóa tiếng nói dựa trên biến đổi 100 có thể được tạo cấu hình để hoạt động trong nhiều chế độ khác nhau, ví dụ trong chế độ bước ngắn và trong chế độ bước dài. Khi hoạt động trong chế độ bước ngắn, bộ mã hóa tiếng nói dựa trên biến đổi 100 có thể được tạo cấu hình để chia nhỏ phân đoạn hoặc khung của tín hiệu âm thanh (ví dụ, tín hiệu tiếng nói) thành tập hợp 132 của các khối ngắn 131 (như đã nêu trên). Mặt khác, khi hoạt động trong chế độ bước dài, bộ mã hóa tiếng nói dựa trên biến đổi 100 có thể được tạo cấu hình để trực tiếp xử lý phân đoạn hoặc khung của tín hiệu âm thanh.

Ví dụ, khi hoạt động trong chế độ bước ngắn, bộ mã hóa 100 có thể được tạo cấu hình để xử lý bốn khối 131 mỗi khung. Các khung của bộ mã hóa 100 có thể tương đối ngắn trong thời gian vật lý đối với các thiết lập nhất định của hoạt động đồng bộ khung video. Đây là trường hợp cụ thể khi tần số khung video tăng lên (ví dụ, 100Hz so với 50Hz), dẫn đến giảm độ dài thời gian của phân đoạn hoặc khung của tín hiệu tiếng nói. Trong các trường hợp như thế, việc chia nhỏ khung thành nhiều nhiều khối 131 (ngắn) có thể không có lợi, do làm giảm độ phân giải trong miền biến đổi. Do đó, chế độ bước dài có thể được sử dụng để yêu cầu việc sử dụng chỉ một khối 131 mỗi khung. Việc sử dụng một khối 131 mỗi khung có thể là có ích để mã hóa các tín hiệu âm thanh bao gồm âm nhạc (ngay cả đối với các khung tương đối dài). Các lợi ích này có thể là do độ phân giải tăng lên trong miền biến đổi, khi sử dụng chỉ một khối 131 mỗi khung hoặc khi sử dụng số lượng khối 131 giảm đi mỗi khung.

Sau đây, hoạt động của bộ mã hóa 100 trong chế độ bước ngắn được mô tả chi tiết hơn. Tập hợp 132 của các khối có thể được đưa vào đơn vị ước lượng đường bao 102. Đơn vị ước lượng đường bao 102 có thể được tạo cấu hình để xác định đường

bao 133 dựa trên tập hợp 132 của các khối. Đường bao 133 có thể dựa trên trị số quân phương (root means squared - RMS) của các hệ số biến đổi tương ứng của nhiều khối 131 được bao gồm trong tập hợp 132 của các khối. Khối 131 thường cung cấp nhiều hệ số biến đổi (ví dụ, 256 hệ số biến đổi) trong nhiều bin tần số tương ứng 301 (xem Fig.3a). Nhiều bin tần số 301 có thể được nhóm thành nhiều dải tần số 302. Nhiều dải tần số 302 có thể được lựa chọn dựa trên các xem xét về mặt âm thính học. Ví dụ, các bin tần số 301 có thể được nhóm vào thành các dải tần số 302 theo thang lôga hoặc thang Bark. Đường bao 134 mà được xác định dựa trên tập hợp hiện tại 132 của các khối có thể bao gồm nhiều trị số năng lượng đối với nhiều dải tần số 302 một cách lần lượt. Trị số năng lượng cụ thể đối với dải tần số cụ thể 302 có thể được xác định dựa trên các hệ số biến đổi của các khối 131 của tập hợp 132, tương ứng với các bin tần số 301 nằm trong dải tần số cụ thể 302. Trị số năng lượng cụ thể có thể được xác định dựa trên trị số RMS của các hệ số biến đổi này. Như thế, đường bao 133 đối với tập hợp hiện tại 132 của các khối (còn gọi là đường bao hiện tại 133) có thể chỉ báo đường bao trung bình của các khối 131 của các hệ số biến đổi được bao gồm trong tập hợp hiện tại 132 của các khối, hoặc có thể chỉ báo đường bao trung bình của các khối 132 của các hệ số biến đổi được sử dụng để xác định đường bao 133.

Cần lưu ý rằng đường bao hiện tại 133 có thể được xác định dựa trên một hoặc nhiều khối 131 khác của các hệ số biến đổi gần kề tập hợp hiện tại 132 của các khối. Điều này được minh họa trên Fig.2, trong đó đường bao hiện tại 133 (được chỉ báo bởi đường bao hiện tại được lượng tử hóa 134) được xác định dựa trên các khối 131 của tập hợp hiện tại 132 của các khối và dựa trên khối 201 từ tập hợp khối ở trước tập hợp hiện tại 132 của các khối. Trong ví dụ minh họa, đường bao hiện tại 133 được xác định dựa trên năm khối 131. Bằng cách xem xét các khối gần kề khi xác định đường bao hiện tại 133, tính liên tục của các đường bao của các tập hợp gần kề 132 của các khối có thể được đảm bảo.

Khi xác định đường bao hiện tại 133, các hệ số biến đổi của các khối 131 khác nhau có thể được gia trọng. Cụ thể là, các khối ngoài cùng 201, 202 mà được xem xét để xác định đường bao hiện tại 133 có thể có gia trọng thấp hơn các khối còn lại 131. Ví dụ, các hệ số biến đổi của các khối ngoài cùng 201, 202 có thể được gia trọng với 0,5, trong đó các hệ số biến đổi của các khối 131 khác có thể được gia trọng với 1.

Cần lưu ý rằng giống với cách xem xét các khối 201 của tập hợp đứng trước 132 của các khối, một hoặc nhiều khối (chính là các khối đứng phía trước) của khối trực tiếp đứng sau 132 của các khối có thể được xem xét để xác định đường bao hiện thời 133.

Các trị số năng lượng của đường bao hiện thời 133 có thể được biểu diễn trên thang lôga (ví dụ, trên thang dB). Đường bao hiện thời 133 có thể được đưa vào đơn vị lượng tử hóa đường bao 103 mà được tạo cấu hình để lượng tử hóa các trị số năng lượng của đường bao hiện thời 133. Đơn vị lượng tử hóa đường bao 103 có thể cung cấp độ phân giải bộ lượng tử hóa được định trước, ví dụ độ phân giải là 3dB. Các chỉ số lượng tử hóa của đường bao 133 có thể được cung cấp dưới dạng dữ liệu đường bao 161 trong dòng bit được tạo ra bởi bộ mã hóa 100. Hơn nữa, đường bao được lượng tử hóa 134, tức là đường bao bao gồm các trị số năng lượng được lượng tử hóa của đường bao 133, có thể được đưa vào đơn vị nội suy 104.

Đơn vị nội suy 104 được tạo cấu hình để xác định đường bao đối với mỗi khối 131 của tập hợp hiện tại 132 của các khối dựa trên đường bao hiện thời được lượng tử hóa 134 và dựa trên đường bao trước đó được lượng tử hóa 135 (mà được xác định đối với tập hợp 132 của các khối liền trước tập hợp hiện tại 132 của các khối). Hoạt động của đơn vị nội suy 104 được minh họa trên các hình vẽ Fig.2, Fig.3a và Fig.3b. Fig.2 thể hiện trình tự của các khối 131 của các hệ số biến đổi. Trình tự của các khối 131 được nhóm thành các tập hợp kế tiếp 132 của các khối, trong đó mỗi tập hợp 132 của các khối được sử dụng để xác định đường bao được lượng tử hóa, ví dụ đường bao hiện thời được lượng tử hóa 134 và đường bao trước đó được lượng tử hóa 135. Fig.3a thể hiện các ví dụ của đường bao trước đó được lượng tử hóa 135 và của đường bao hiện thời được lượng tử hóa 134. Như nêu trên, các đường bao có thể chỉ báo năng lượng phổ 303 (ví dụ, trên thang dB). Các trị số năng lượng tương ứng 303 của đường bao trước đó được lượng tử hóa 135 và của đường bao hiện thời được lượng tử hóa 134 đối với cùng dải tần số 302 có thể được nội suy (ví dụ, bằng cách sử dụng nội suy tuyến tính) để xác định đường bao được nội suy 136. Nói cách khác, các trị số năng lượng 303 của dải tần số cụ thể 302 có thể được nội suy để cung cấp trị số năng lượng 303 của đường bao được nội suy 136 trong dải tần số cụ thể 302.

Cần lưu ý rằng tập hợp khối mà các đường bao được nội suy 136 được xác định và áp dụng cho các khối này có thể khác với tập hợp hiện tại 132 của các khối, dựa

trên việc đường bao hiện thời được lượng tử hóa 134 được xác định. Điều này được minh họa trên Fig.2, hình vẽ này thể hiện tập hợp chuyển đổi 332 của các khối, mà được chuyển đổi so với tập hợp hiện tại 132 của các khối và bao gồm các khối 3 và 4 của tập hợp trước đó 132 của các khối (được chỉ báo bởi các số chỉ dẫn 203 và 201, một cách lần lượt) và các khối 1 và 2 của tập hợp hiện tại 132 của các khối (được chỉ báo bởi các số chỉ dẫn 204 và 205, một cách lần lượt). Thực tế là, các đường bao được nội suy 136 được xác định dựa trên đường bao hiện thời được lượng tử hóa 134 và dựa trên đường bao trước đó được lượng tử hóa 135 có thể thích hợp hơn đối với các khối của tập hợp chuyển đổi 332 của các khối, so với độ thích hợp đối với các khối của tập hợp hiện tại 132 của các khối.

Do đó, các đường bao được nội suy 136 được thể hiện trên Fig.3b có thể được sử dụng để làm phẳng các khối 131 của tập hợp chuyển đổi 332 của các khối. Điều này được thể hiện bằng Fig.3b kết hợp với Fig.2. Có thể thấy rằng đường bao được nội suy 341 trên Fig.3b có thể được áp dụng vào khối 203 trên Fig.2, đường bao được nội suy 342 trên Fig.3b có thể được áp dụng vào khối 201 trên Fig.2, đường bao được nội suy 343 trên Fig.3b có thể được áp dụng vào khối 204 trên Fig.2, và đường bao được nội suy 344 trên Fig.3b (đường bao này trong ví dụ minh họa tương ứng với đường bao hiện thời được lượng tử hóa 136) có thể được áp dụng vào khối 205 trên Fig.2. Như thế, tập hợp 132 của các khối để xác định đường bao hiện thời được lượng tử hóa 134 có thể khác với tập hợp chuyển đổi 332 của các khối mà các đường bao được nội suy 136 được xác định và các đường bao được nội suy 136 được áp dụng vào đó (nhằm mục đích làm phẳng). Cụ thể là, đường bao hiện thời được lượng tử hóa 134 có thể được xác định bằng cách sử dụng các khối dừng phía trước 203, 201, 204, 205 của tập hợp chuyển đổi 332 của các khối, mà sẽ được làm phẳng bằng cách sử dụng đường bao hiện thời được lượng tử hóa 134. Điều này là có lợi từ quan điểm về tính liên tục.

Việc nội suy các trị số năng lượng 303 để xác định các đường bao được nội suy 136 được minh họa trên Fig.3b. Có thể thấy rằng bằng phép nội suy giữa trị số năng lượng của đường bao trước đó được lượng tử hóa 135 đến trị số năng lượng tương ứng của đường bao hiện thời được lượng tử hóa 134 các trị số năng lượng của các đường bao được nội suy 136 có thể được xác định đối với các khối 131 của tập hợp chuyển đổi 332 của các khối. Cụ thể là, đối với mỗi khối 131 của tập hợp chuyển đổi 332

đường bao được nội suy 136 có thể được xác định, qua đó cung cấp nhiều đường bao được nội suy 136 đối với nhiều khối 203, 201, 204, 205 của tập hợp chuyển đổi 332 của các khối. Đường bao được nội suy 136 của khối 131 của hệ số biến đổi (ví dụ, khối bất kỳ trong số các khối 203, 201, 204, 205 của tập hợp chuyển đổi 332 của các khối) có thể được sử dụng để mã hóa khối 131 của các hệ số biến đổi. Cần lưu ý rằng các chỉ số lượng tử hóa 161 của đường bao hiện thời 133 được đưa vào bộ giải mã tương ứng trong dòng bit. Do đó, bộ giải mã tương ứng có thể được tạo cấu hình để xác định nhiều đường bao được nội suy 136 theo cách tương tự như đơn vị nội suy 104 của bộ mã hóa 100.

Đơn vị tạo khung 101, đơn vị ước lượng đường bao 102, đơn vị lượng tử hóa đường bao 103, và đơn vị nội suy 104 hoạt động trên tập hợp khối (tức là tập hợp hiện tại 132 của các khối và/hoặc tập hợp chuyển đổi 332 của các khối). Mặt khác, việc mã hóa thực của hệ số biến đổi có thể được thực hiện trên cơ sở khối-cạnh-khối. Sau đây sẽ tham chiếu đến việc mã hóa của khối hiện tại 131 của các hệ số biến đổi, mà có thể là khối bất kỳ trong số nhiều khối 131 của tập hợp chuyển đổi 332 của các khối (hoặc có thể là tập hợp hiện tại 132 của các khối theo các phương án thực hiện khác của bộ mã hóa tiếng nói dựa trên biến đổi 100).

Hơn nữa, cần lưu ý rằng bộ mã hóa 100 có thể được vận hành trong chế độ bước dài. Trong chế độ này, khung của phân đoạn của tín hiệu âm thanh không bị chia nhỏ và được xử lý như một khối. Do đó, chỉ một khối 131 của các hệ số biến đổi được xác định mỗi khung. Khi vận hành trong chế độ bước dài, đơn vị tạo khung 101 có thể được tạo cấu hình để tách một khối hiện tại 131 của các hệ số biến đổi đối với phân đoạn hoặc khung của tín hiệu âm thanh. Đơn vị ước lượng đường bao 102 có thể được tạo cấu hình để xác định đường bao hiện thời 133 đối với khối hiện tại 131 và đơn vị lượng tử hóa đường bao 103 có thể được tạo cấu hình để lượng tử hóa một đường bao hiện thời 133 để xác định đường bao hiện thời được lượng tử hóa 134 (và để xác định dữ liệu đường bao 161 đối với khối hiện tại 131). Khi ở trong chế độ bước dài, phép nội suy đường bao thường quá hạn. Do đó, đường bao được nội suy 136 đối với khối hiện tại 131 thường tương ứng với đường bao hiện thời được lượng tử hóa 134 (khi bộ mã hóa 100 được hoạt động trong chế độ bước dài).

Đường bao được nội suy hiện tại 136 đối với khối hiện tại 131 có thể cung cấp trị số gần đúng của đường bao phổ của các hệ số biến đổi của khối hiện tại 131. Bộ mã

hóa 100 có thể bao gồm đơn vị trước làm phẳng 105 và đơn vị xác định độ lợi đường bao 106 được tạo cấu hình để xác định đường bao được điều chỉnh 139 đối với khối hiện tại 131, dựa trên đường bao được nội suy hiện tại 136 và dựa trên khối hiện tại 131. Cụ thể là, độ lợi đường bao đối với khối hiện tại 131 có thể được xác định sao cho phương sai của các hệ số biến đổi được làm phẳng của khối hiện tại 131 được điều chỉnh. $X(k)$, $k = 1, \dots, K$ có thể là các hệ số biến đổi của khối hiện tại 131 (với, ví dụ, $K = 256$), và $E(k)$, $k = 1, \dots, K$ có thể là các trị số năng lượng phổ trung bình 303 của đường bao được nội suy hiện tại 136 (với các trị số năng lượng $E(k)$ của cùng dải tần số 302 là bằng nhau). Độ lợi đường bao a có thể được xác định sao cho phương sai của các hệ số biến đổi được làm phẳng $\tilde{X}(k) = \frac{X(k)}{a \cdot \sqrt{E(k)}}$ được điều chỉnh.

Cụ thể là, độ lợi đường bao a có thể được xác định sao cho phương sai là một.

Cần lưu ý rằng độ lợi đường bao a có thể được xác định đối với khoảng nhỏ của khoảng tần số đầy đủ của khối hiện tại 131 của các hệ số biến đổi. Nói cách khác, độ lợi đường bao a có thể được xác định chỉ dựa trên tập hợp con của các bin tần số 301 và/hoặc chỉ dựa trên tập hợp con của các dải tần số 302. Ví dụ, độ lợi đường bao a có thể được xác định dựa trên các bin tần số 301 lớn hơn bin tần số khởi đầu 304 (bin tần số khởi đầu lớn hơn 0 hoặc 1). Kết quả là, đường bao được điều chỉnh 139 đối với khối hiện tại 131 có thể được xác định bằng cách áp dụng độ lợi đường bao a chỉ vào các trị số năng lượng phổ trung bình 303 của đường bao được nội suy hiện tại 136 mà liên kết với các bin tần số 301 nằm trên bin tần số khởi đầu 304. Do đó, đường bao được điều chỉnh 139 đối với khối hiện tại 131 có thể tương ứng với đường bao được nội suy hiện tại 136, đối với các bin tần số 301 tại và ở dưới bin tần số khởi đầu, và có thể tương ứng với đường bao được nội suy hiện tại 136 được bù bởi độ lợi đường bao a , đối với các bin tần số 301 nằm trên bin tần số khởi đầu. Điều này được minh họa trên Fig.3a bởi đường bao được điều chỉnh 339 (được thể hiện bằng nét đứt).

Việc áp dụng độ lợi đường bao a 137 (còn được gọi là độ lợi hiệu chỉnh mức) vào đường bao được nội suy hiện tại 136 tương ứng với độ điều chỉnh hoặc độ bù của đường bao được nội suy hiện tại 136, qua đó tạo ra đường bao được điều chỉnh 139, như được minh họa bởi Fig.3a. Độ lợi đường bao a 137 có thể được mã hóa dưới dạng dữ liệu độ lợi 162 vào dòng bit.

Bộ mã hóa 100 còn có thể bao gồm đơn vị làm mịn đường bao 107 mà được tạo cấu hình để xác định đường bao được điều chỉnh 139 dựa trên độ lợi đường bao a

137 và dựa trên đường bao được nội suy hiện tại 136. Đường bao được điều chỉnh 139 có thể được sử dụng cho việc xử lý tín hiệu của khối 131 của hệ số biến đổi. Độ lợi đường bao a 137 có thể được lượng tử hóa đến độ phân giải cao hơn (ví dụ, trong các bước 1dB) so với đường bao được nội suy hiện tại 136 (mà có thể được lượng tử hóa trong các bước 3dB). Như thế, đường bao được điều chỉnh 139 có thể được lượng tử hóa đến độ phân giải cao hơn của độ lợi đường bao a 137 (ví dụ, trong các bước 1dB).

Hơn nữa, đơn vị làm mịn đường bao 107 có thể được tạo cấu hình để xác định đường bao phân bố 138. Đường bao phân bố 138 có thể tương ứng với phiên bản được lượng tử hóa của đường bao được điều chỉnh 139 (ví dụ, được lượng tử hóa đến các mức lượng tử hóa 3dB). Đường bao phân bố 138 có thể được sử dụng cho các mục đích phân bố bit. Cụ thể là, đường bao phân bố 138 có thể được sử dụng để xác định – đối với hệ số biến đổi cụ thể của khối hiện tại 131 – bộ lượng tử hóa cụ thể từ tập hợp định trước của các bộ lượng tử hóa, trong đó bộ lượng tử hóa cụ thể sẽ được dùng để lượng tử hóa hệ số biến đổi cụ thể.

Bộ mã hóa 100 bao gồm đơn vị làm phẳng 108 được tạo cấu hình để làm phẳng khối hiện tại 131 bằng cách sử dụng đường bao được điều chỉnh 139, qua đó tạo ra khối 140 của các hệ số biến đổi được làm phẳng $\tilde{X}(k)$. Khối 140 của các hệ số biến đổi được làm phẳng $\tilde{X}(k)$ có thể được mã hóa bằng cách sử dụng vòng lặp dự đoán trong miền biến đổi. Như thế, khối 140 có thể được mã hóa bằng cách sử dụng bộ dự đoán dài phụ 117. Vòng lặp dự đoán bao gồm đơn vị vi sai 115 được tạo cấu hình để xác định khối 141 của các hệ số sai số dự đoán $\Delta(k)$, dựa trên khối 140 của các hệ số biến đổi được làm phẳng $\tilde{X}(k)$ và dựa trên khối 150 của các hệ số biến đổi được ước lượng $\hat{X}(k)$, ví dụ $\Delta(k) = \tilde{X}(k) - \hat{X}(k)$. Cần lưu ý rằng do khối 140 bao gồm các hệ số biến đổi được làm phẳng, tức là các hệ số biến đổi mà đã được chuẩn hóa hoặc làm phẳng bằng cách sử dụng các trị số năng lượng 303 của đường bao được điều chỉnh 139, khối 150 của các hệ số biến đổi được ước lượng còn bao gồm trị số ước tính của các hệ số biến đổi được làm phẳng. Nói cách khác, đơn vị vi sai 115 hoạt động trong chính miền được làm phẳng. Vì vậy, khối 141 của các hệ số sai số dự đoán $\Delta(k)$ được biểu diễn trong miền được làm phẳng.

Khối 141 của các hệ số sai số dự đoán $\Delta(k)$ có thể thể hiện phương sai khác một. Bộ mã hóa 100 có thể bao gồm đơn vị lập lại thang tỷ lệ 111 được tạo cấu hình để lập lại thang tỷ lệ các hệ số sai số dự đoán $\Delta(k)$ để tạo ra khối 142 của các hệ số sai

số được lập thang tỷ lệ. Đơn vị lập lại thang tỷ lệ 111 có thể sử dụng một hoặc nhiều nguyên tắc suy nghiệm định trước để thực hiện việc lập lại thang tỷ lệ. Kết quả là, khối 142 của các hệ số sai số được lập thang tỷ lệ thể hiện phương sai mà (trung bình) là gần với một hơn (so với khối 141 của các hệ số sai số dự đoán). Việc này có thể có lợi đối với việc lượng tử hóa và mã hóa sau đó.

Bộ mã hóa 100 bao gồm đơn vị lượng tử hóa hệ số 112 được tạo cấu hình để lượng tử hóa khối 141 của các hệ số sai số dự đoán hoặc khối 142 của các hệ số sai số được lập thang tỷ lệ. Đơn vị lượng tử hóa hệ số 112 có thể bao gồm hoặc có thể sử dụng tập hợp của các bộ lượng tử hóa được định trước. Tập hợp của các bộ lượng tử hóa được định trước có thể cung cấp các bộ lượng tử hóa với độ chính xác khác nhau và độ phân giải khác nhau. Điều này được minh họa trên Fig.4, trên đó các bộ lượng tử hóa khác nhau 321, 322, 323 được minh họa. Các bộ lượng tử hóa khác nhau có thể có độ chính xác khác nhau (được chỉ báo bởi các trị số dB khác nhau). Một bộ lượng tử hóa cụ thể trong số nhiều bộ lượng tử hóa 321, 322, 323 có thể tương ứng với trị số cụ thể của đường bao phân bố 138. Như thế, trị số năng lượng của đường bao phân bố 138 có thể hướng đến một bộ lượng tử hóa cụ thể trong nhiều bộ lượng tử hóa. Như thế, việc xác định đường bao phân bố 138 có thể đơn giản hóa quá trình lựa chọn ra bộ lượng tử hóa sẽ được dùng cho hệ số sai số cụ thể. Nói cách khác, đường bao phân bố 138 có thể đơn giản hóa quá trình phân bố bit.

Tập hợp của các bộ lượng tử hóa có thể bao gồm một hoặc nhiều bộ lượng tử hóa 322 mà sử dụng sự hòa sắc để ngẫu nhiên hóa sai số lượng tử hóa. Điều này được minh họa trên Fig.4 thể hiện tập hợp thứ nhất 326 của các bộ lượng tử hóa định trước mà bao gồm tập hợp con 324 của các bộ lượng tử hóa hòa sắc và tập hợp thứ hai 327 định trước các bộ lượng tử hóa mà bao gồm tập hợp con 325 của các bộ lượng tử hóa hòa sắc. Như thế, đơn vị lượng tử hóa hệ số 112 có thể sử dụng các tập hợp khác nhau 326, 327 của các bộ lượng tử hóa định trước, trong đó tập hợp của các bộ lượng tử hóa được định trước, mà sẽ được sử dụng bởi đơn vị lượng tử hóa hệ số 112 có thể phụ thuộc vào tham số điều chỉnh 146 được cung cấp bởi bộ dự đoán 117. Cụ thể là, đơn vị lượng tử hóa hệ số 112 có thể được tạo cấu hình để chọn tập hợp 326, 327 của các bộ lượng tử hóa định trước để lượng tử hóa khối 142 của hệ số sai số được lập thang tỷ lệ, dựa trên tham số điều chỉnh 146, trong đó tham số điều chỉnh 146 có thể phụ thuộc vào một hoặc nhiều tham số dự đoán được cung cấp bởi bộ dự đoán 117. Một

hoặc nhiều tham số dự đoán có thể chỉ báo chất lượng của khối 150 của các hệ số biến đổi được ước lượng được cung cấp bởi bộ dự đoán 117.

Các hệ số sai số được lượng tử hóa có thể được mã hóa entropi, bằng cách sử dụng, ví dụ, mã Huffman, qua đó tạo ra dữ liệu hệ số 163 sẽ được đưa vào dòng bit được tạo ra bởi bộ mã hóa 100.

Bộ mã hóa 100 có thể được tạo cấu hình để thực hiện quá trình phân bố bit. Để thực hiện mục đích này, bộ mã hóa 100 có thể bao gồm đơn vị phân bố các bit 109, 110. Đơn vị phân bố bit 109 có thể được tạo cấu hình để xác định tổng số các bit 143 có sẵn để mã hóa khối hiện tại 142 của các hệ số sai số được lập thang tỷ lệ. Tổng số các bit 143 có thể được xác định dựa trên đường bao phân bố 138. Đơn vị phân bố bit 110 có thể được tạo cấu hình để đưa sự phân bố bit tương đối đến các hệ số sai số được lập thang tỷ lệ khác nhau, tùy thuộc vào trị số năng lượng tương ứng trong đường bao phân bố 138.

Quá trình phân bố bit có thể sử dụng quy trình phân bố lặp. Trong quy trình phân bố, đường bao phân bố 138 có thể được bù bằng cách sử dụng tham số bù, qua đó lựa chọn các bộ lượng tử hóa có độ phân giải tăng / giảm. Như thế, tham số bù có thể được sử dụng để làm mịn hoặc làm thô sự lượng tử hóa toàn phần. Tham số bù có thể được xác định sao cho dữ liệu hệ số 163, dữ liệu này thu được bằng cách sử dụng các bộ lượng tử hóa được đưa ra bởi tham số bù và đường bao phân bố 138, bao gồm số lượng bit mà tương ứng với (hoặc không vượt quá) tổng số các bit 143 được ấn định cho khối hiện tại 131. Tham số bù mà đã được sử dụng bởi bộ mã hóa 100 để mã hóa khối hiện tại 131 được bao gồm dưới dạng dữ liệu hệ số 163 vào dòng bit. Kết quả là, bộ giải mã tương ứng được cho phép xác định các bộ lượng tử hóa mà đã được sử dụng bởi đơn vị lượng tử hóa hệ số 112 để lượng tử hóa khối 142 của các hệ số sai số được lập thang tỷ lệ.

Kết quả của việc lượng tử hóa các hệ số sai số được lập thang tỷ lệ là thu được khối 145 của các hệ số sai số được lượng tử hóa. Khối 145 của các hệ số sai số được lượng tử hóa tương ứng với khối hệ số sai số mà có sẵn tại bộ giải mã tương ứng. Do đó, khối 145 của các hệ số sai số được lượng tử hóa có thể được sử dụng để xác định khối 150 của các hệ số biến đổi được ước lượng. Bộ mã hóa 100 có thể bao gồm đơn vị đảo ngược lập thang tỷ lệ 113 được tạo cấu hình để thực hiện đảo ngược các thao tác lập lại thang tỷ lệ được thực hiện bởi đơn vị lập lại thang tỷ lệ 113, qua đó tạo ra

khối 147 của các hệ số sai số được lượng tử hóa được lập thang tỷ lệ. Đơn vị cộng thêm 116 có thể được sử dụng để xác định khối 148 của các hệ số được làm phẳng được tái tạo, bằng cách cộng khối 150 của các hệ số biến đổi được ước lượng vào khối 147 của các hệ số sai số được lượng tử hóa được lập thang tỷ lệ. Hơn nữa, đơn vị làm phẳng ngược 114 có thể được sử dụng để áp dụng đường bao được điều chỉnh 139 vào khối 148 của các hệ số được làm phẳng được tái tạo, qua đó tạo ra khối 149 của các hệ số được tái tạo. Khối 149 của các hệ số được tái tạo tương ứng với phiên bản của khối 131 của các hệ số biến đổi mà có sẵn tại bộ giải mã tương ứng. Vì vậy, khối 149 của các hệ số được tái tạo có thể được sử dụng trong bộ dự đoán 117 để xác định khối 150 của các hệ số được ước lượng.

Khối 149 của các hệ số được tái tạo được biểu diễn trong miền không phẳng, tức là khối 149 của các hệ số được tái tạo cũng là thể hiện của đường bao phổ của khối hiện tại 131. Như được nêu dưới đây, việc này có thể có lợi đối với hoạt động của bộ dự đoán 117.

Bộ dự đoán 117 có thể được tạo cấu hình để ước tính khối 150 của các hệ số biến đổi được ước lượng dựa trên một hoặc nhiều khối trước đó 149 của các hệ số được tái tạo. Cụ thể là, bộ dự đoán 117 có thể được tạo cấu hình để xác định một hoặc nhiều tham số dự đoán sao cho tiêu chí sai số dự đoán định trước được làm giảm đi (ví dụ, được làm giảm đến mức tối thiểu). Ví dụ, một hoặc nhiều tham số dự đoán có thể được xác định sao cho năng lượng, hoặc năng lượng được gia trọng cảm quan, của khối 141 của các hệ số sai số dự đoán được làm giảm đi (ví dụ, được làm giảm đến mức tối thiểu). Một hoặc nhiều tham số dự đoán có thể được bao gồm dưới dạng dữ liệu bộ dự đoán 164 vào dòng bit được tạo ra bởi bộ mã hóa 100.

Bộ dự đoán dữ liệu 164 có thể chỉ báo một hoặc nhiều tham số dự đoán. Như sẽ được nêu trong bản mô tả này, bộ dự đoán 117 có thể chỉ được sử dụng cho tập hợp con của các khung hoặc các khối 131 của tín hiệu âm thanh. Cụ thể là, bộ dự đoán 117 có thể không được sử dụng cho khối thứ nhất 131 của khung I (khung độc lập), khung này thường được mã hóa theo cách độc lập với khối trước đó. Ngoài ra, bộ dự đoán dữ liệu 164 có thể bao gồm một hoặc nhiều cờ mà chỉ báo sự có mặt của bộ dự đoán 117 đối với khối cụ thể 131. Đối với các khối, trong đó đóng góp của bộ dự đoán là gần như không đáng kể (ví dụ, khi độ lợi bộ dự đoán được lượng tử hóa đến không), có thể có lợi khi sử dụng cờ báo hiệu sự có mặt của bộ dự đoán để báo hiệu tình huống này,

tình huống này thường đòi hỏi số lượng bit giảm so với việc truyền độ lợi là không). Nói cách khác, bộ dự đoán dữ liệu 164 đối với khối 131 có thể bao gồm một hoặc nhiều các cờ báo hiệu sự có mặt của bộ dự đoán mà chỉ báo rằng một hoặc nhiều tham số dự đoán đã được xác định hay chưa (và được bao gồm trong bộ dự đoán dữ liệu 164). Việc sử dụng một hoặc nhiều các cờ báo hiệu sự có mặt của bộ dự đoán có thể được sử dụng để tiết kiệm bit, nên bộ dự đoán 117 không được sử dụng cho khối cụ thể 131. Do đó, tùy thuộc vào số lượng của các khối 131, các khối này được mã hóa mà không sử dụng bộ dự đoán 117, thì việc sử dụng một hoặc nhiều các cờ báo hiệu sự có mặt của bộ dự đoán có thể có hiệu quả hơn về tốc độ bit (trung bình) hơn là việc truyền các tham số dự đoán mặc định (ví dụ, có trị số bằng không).

Sự có mặt của bộ dự đoán 117 có thể được truyền rõ ràng trên cơ sở mỗi khối. Việc này cho phép tiết kiệm bit khi không sử dụng phép dự đoán. Ví dụ, đối với các khung I, chỉ ba cờ báo hiệu sự có mặt của bộ dự đoán có thể được sử dụng, vì khối thứ nhất của khung I không thể sử dụng phép dự đoán. Nói cách khác, nếu biết rằng khối cụ thể 131 là khối thứ nhất của khung I, thì khi đó cờ báo hiệu sự có mặt của bộ dự đoán có thể không cần được truyền đối với khối cụ thể 131 này (đã biết rằng đối với bộ giải mã tương ứng thì khối cụ thể 131 không sử dụng bộ dự đoán 117).

Bộ dự đoán 117 có thể sử dụng mô hình tín hiệu, như được mô tả trong đơn yêu cầu cấp bằng độc quyền sáng chế số US61750052 và các đơn yêu cầu cấp bằng độc quyền sáng chế làm đơn ưu tiên của nó, nội dung của chúng được đưa vào bằng cách viện dẫn. Một hoặc nhiều tham số dự đoán có thể tương ứng với một hoặc nhiều tham số mô hình của mô hình tín hiệu.

Fig.1b thể hiện sơ đồ khối của bộ mã hóa tiếng nói dựa trên biến đổi làm ví dụ khác 170. Bộ mã hóa tiếng nói dựa trên biến đổi 170 trên Fig.1b bao gồm nhiều bộ phận của bộ mã hóa 100 trên Fig.1a. Tuy nhiên, bộ mã hóa tiếng nói dựa trên biến đổi 170 trên Fig.1b được tạo cấu hình để tạo dòng bit có tốc độ bit biến thiên. Để thực hiện mục đích này, bộ mã hóa 170 bao gồm đơn vị báo cáo tốc độ bit trung bình (Average Bit Rate - ABR) 172 được tạo cấu hình để theo dõi tốc độ bit mà đã được sử dụng bởi dòng bit đối với các khối trước đó 131. Đơn vị phân bố bit 171 sử dụng thông tin này để xác định tổng số các bit 143 mà có sẵn để mã hóa khối hiện tại 131 của các hệ số biến đổi.

Tổng thể, các bộ mã hóa tiếng nói dựa trên biến đổi 100, 170 được tạo cấu hình để tạo dòng bit mà chỉ báo hoặc mà bao gồm

- dữ liệu đường bao 161 chỉ báo đường bao hiện thời được lượng tử hóa 134. Đường bao hiện thời được lượng tử hóa 134 được sử dụng để mô tả đường bao của các khối của tập hợp hiện tại 132 hoặc tập hợp chuyển đổi 332 của các khối hệ số biến đổi.
- dữ liệu độ lợi 162 chỉ báo độ lợi hiệu chỉnh mức α để điều chỉnh đường bao được nội suy 136 của khối hiện tại 131 của các hệ số biến đổi. Thông thường, độ lợi α khác được cung cấp cho mỗi khối 131 của tập hợp hiện tại 132 hoặc tập hợp chuyển đổi 332 của các khối.
- dữ liệu hệ số 163 chỉ báo khối 141 của các hệ số sai số dự đoán đối với khối hiện tại 131. Cụ thể là, dữ liệu hệ số 163 chỉ báo khối 145 của các hệ số sai số được lượng tử hóa. Hơn nữa, dữ liệu hệ số 163 có thể chỉ báo tham số bù mà có thể được sử dụng để xác định các bộ lượng tử hóa để thực hiện đảo ngược lượng tử hóa tại bộ giải mã.
- dữ liệu bộ dự đoán 164 chỉ báo một hoặc nhiều hệ số bộ dự đoán sẽ được sử dụng để xác định khối 150 của các hệ số được ước lượng từ các khối trước đó 149 của các hệ số được tái tạo.

Sau đây, bộ giải mã tiếng nói dựa trên biến đổi tương ứng 500 được mô tả trong bối cảnh của các hình vẽ từ Fig.5a đến Fig.5d. Fig.5a thể hiện sơ đồ khối của bộ giải mã tiếng nói dựa trên biến đổi làm ví dụ 500. Sơ đồ khối thể hiện giàn lọc tổng hợp 504 (còn được gọi là đơn vị biến đổi ngược) được sử dụng để chuyển đổi khối 149 của các hệ số được tái tạo từ miền biến đổi vào miền thời gian, qua đó tạo ra các mẫu của tín hiệu âm thanh được giải mã. Giàn lọc tổng hợp 504 có thể sử dụng MDCT nghịch đảo có bước được định trước (ví dụ, bước khoảng 5 ms hoặc 256 mẫu).

Vòng lặp chính của bộ giải mã 500 hoạt động trong các đơn vị theo bước này. Mỗi bước tạo ra vector miền biến đổi (còn được gọi là khối) có chiều dài hoặc kích thước tương ứng với thiết lập băng thông định trước của hệ thống. Nhờ vào việc đệm không theo kích thước biến đổi của giàn lọc tổng hợp 504, vector miền biến đổi sẽ được sử dụng để tổng hợp phiên bản cập nhật tín hiệu miền thời gian có độ dài định trước (ví dụ, 5ms) đến quá trình cộng/chồng lấp của giàn lọc tổng hợp 504.

Như nêu trên, bộ các mã hóa-giải mã âm thanh dựa trên biến đổi chung thường sử dụng các khung có trình tự là các khối ngắn trong khoảng 5 ms để thao tác tạm thời. Như thế, bộ các mã hóa-giải mã âm thanh dựa trên biến đổi chung cung cấp các biến đổi cần thiết và các công cụ chuyển đổi cửa sổ để các khối dài và các khối ngắn cùng tồn tại liền mạch. Do đó, đầu trước phổ lời nói được xác định bằng cách lược bỏ giàn lọc tổng hợp 504 trên Fig.5a có thể được tích hợp một cách thuận tiện vào bộ mã hóa-giải mã âm thanh dựa trên biến đổi đa dụng, mà không cần sử dụng các công cụ chuyển đổi bổ sung. Nói cách khác, bộ giải mã tiếng nói dựa trên biến đổi 500 trên Fig.5a có thể được kết hợp một cách thuận lợi với bộ giải mã âm thanh dựa trên biến đổi chung. Cụ thể là, bộ giải mã tiếng nói dựa trên biến đổi 500 trên Fig.5a có thể sử dụng giàn lọc tổng hợp 504 được cung cấp bởi bộ giải mã âm thanh dựa trên biến đổi chung (ví dụ, bộ giải mã AAC hoặc HE-AAC).

Từ dòng bit đầu vào (cụ thể là từ dữ liệu đường bao 161 và từ dữ liệu độ lợi 162 được bao gồm trong dòng bit), đường bao tín hiệu có thể được xác định bởi bộ giải mã đường bao 503. Cụ thể là, bộ giải mã đường bao 503 có thể được tạo cấu hình để xác định đường bao được điều chỉnh 139 dựa trên dữ liệu đường bao 161 và dữ liệu độ lợi 162. Như thế, bộ giải mã đường bao 503 có thể thực hiện các tác vụ giống với đơn vị nội suy 104 và đơn vị làm mịn đường bao 107 của bộ mã hóa 100, 170. Như đã nêu trên, đường bao được điều chỉnh 109 biểu diễn mô hình của phương sai tín hiệu trong tập hợp của các dải tần số được xác định trước 302.

Hơn nữa, bộ giải mã 500 bao gồm đơn vị làm phẳng ngược 114 được tạo cấu hình để áp dụng đường bao được điều chỉnh 139 vào vector miền được làm phẳng mà các mục vào của nó trên danh nghĩa có thể là của vector phương sai. Vector miền được làm phẳng tương ứng với khối 148 của các hệ số được làm phẳng được tái tạo được mô tả trong bối cảnh của bộ mã hóa 100, 170. Tại đầu ra của đơn vị làm phẳng ngược 114, thu được khối 149 của các hệ số được tái tạo. Khối 149 của các hệ số được tái tạo được đưa đến giàn lọc tổng hợp 504 (để tạo tín hiệu âm thanh được giải mã) và đến bộ giải mã dải phụ 517.

Bộ giải mã dải phụ 517 hoạt động theo cách giống với bộ dự đoán 117 của bộ mã hóa 100, 170. Cụ thể là, bộ giải mã dải phụ 517 được tạo cấu hình để xác định khối 150 của các hệ số biến đổi được ước lượng (trong miền được làm phẳng) dựa trên một hoặc nhiều khối trước đó 149 của các hệ số được tái tạo (sử dụng một hoặc

nhiều tham số dự đoán được lập tín hiệu dòng bit). Nói cách khác, bộ giải mã dải phụ 517 được tạo cấu hình để xuất ra vector miền được làm phẳng và được dự đoán từ bộ đệm của các vector đầu ra được giải mã trước đó và các đường bao tín hiệu, dựa trên các tham số bộ dự đoán như cờ bộ dự đoán và độ lợi bộ dự đoán. Bộ giải mã 500 bao gồm bộ giải mã bộ dự đoán 501 được tạo cấu hình để giải mã bộ dự đoán dữ liệu 164 để xác định một hoặc nhiều tham số dự đoán.

Bộ giải mã 500 còn bao gồm bộ giải mã phổ 502 được tạo cấu hình để hiệu chỉnh thêm cho vector miền được làm phẳng và được dự đoán, thường là dựa trên phần lớn nhất của dòng bit (tức là dựa trên dữ liệu hệ số 163). Quá trình giải mã phổ được kiểm soát chủ yếu bởi vector phân bố, vector này được lấy từ đường bao và tham số điều chỉnh phân bố được truyền (còn được gọi là tham số bù). Như được minh họa trên Fig.5a, có thể có sự phụ thuộc trực tiếp của bộ giải mã phổ 502 lên các tham số bộ dự đoán 520. Như thế, bộ giải mã phổ 502 có thể được tạo cấu hình để xác định khối 147 của các hệ số sai số được lượng tử hóa được lập thang tỷ lệ dựa trên dữ liệu nhận được hệ số 163. Như đã được nêu trong bối cảnh của bộ mã hóa 100, 170, các bộ lượng tử hóa 321, 322, 323 được sử dụng để lượng tử hóa khối 142 của các hệ số sai số được lập thang tỷ lệ thường phụ thuộc vào đường bao phân bố 138 (mà có thể được lấy từ đường bao được điều chỉnh 139) và phụ thuộc vào tham số bù. Hơn nữa, các bộ lượng tử hóa 321, 322, 323 có thể phụ thuộc vào tham số điều chỉnh 146 được cung cấp bởi bộ dự đoán 117. Tham số điều chỉnh 146 có thể được lấy bởi bộ giải mã 500 bằng cách sử dụng các tham số bộ dự đoán 520 (theo cách tương tự như bộ mã hóa 100, 170).

Như nêu trên, dòng bit nhận được bao gồm dữ liệu đường bao 161 và dữ liệu độ lợi 162 mà có thể được sử dụng để xác định đường bao được điều chỉnh 139. Cụ thể là, đơn vị 531 của bộ giải mã đường bao 503 có thể được tạo cấu hình để xác định đường bao hiện thời được lượng tử hóa 134 từ dữ liệu đường bao 161. Ví dụ, đường bao hiện thời được lượng tử hóa 134 có thể có độ phân giải 3 dB trong các dải tần số được xác định trước 302 (như được chỉ ra trên Fig.3a). Đường bao hiện thời được lượng tử hóa 134 có thể được cập nhật cho mỗi tập hợp 132, 332 của các khối (ví dụ, mỗi bốn đơn vị mã hóa, tức là các khối, hoặc mỗi 20ms), cụ thể là đối với mỗi tập hợp chuyển đổi 332 của các khối. Các dải tần số 302 của đường bao hiện thời được lượng

tử hóa 134 có thể bao gồm số lượng tăng tiến của các bin tần số 301 dưới dạng hàm số của tần số, nhằm thích nghi với các đặc tính của khả năng nghe của con người.

Đường bao hiện thời được lượng tử hóa 134 có thể được nội suy tuyến tính từ đường bao trước đó được lượng tử hóa 135 thành các đường bao được nội suy 136 đối với mỗi khối 131 của tập hợp chuyển đổi 332 của các khối (hoặc có thể là của tập hợp hiện tại 132 của các khối). Các đường bao được nội suy 136 có thể được xác định trong miền 3dB được lượng tử hóa. Điều này có nghĩa là các trị số năng lượng được nội suy 303 có thể được làm tròn đến mức 3dB gần nhất. Đường bao được nội suy làm ví dụ 136 được minh họa bởi đường nét chấm trên Fig.3a. Đối với mỗi đường bao hiện thời được lượng tử hóa 134, bốn độ lợi hiệu chỉnh mức α 137 (còn được gọi là các độ lợi đường bao) được cung cấp dưới dạng dữ liệu độ lợi 162. Đơn vị giải mã độ lợi 532 có thể được tạo cấu hình để xác định các độ lợi hiệu chỉnh mức α 137 từ dữ liệu độ lợi 162. Các độ lợi hiệu chỉnh mức có thể được lượng tử hóa trong các bước 1dB. Mỗi độ lợi hiệu chỉnh mức được áp dụng vào đường bao được nội suy tương ứng 136 nhằm đưa các đường bao được điều chỉnh 139 vào các khối khác nhau 131. Do độ phân giải của các độ lợi hiệu chỉnh mức 137 tăng lên, nên đường bao được điều chỉnh 139 có thể có độ phân giải tăng lên (ví dụ, độ phân giải 1dB).

Fig.3b thể hiện sự nội suy tuyến tính hoặc nội suy hình học giữa đường bao trước đó được lượng tử hóa 135 và đường bao hiện thời được lượng tử hóa 134. Các đường bao 135, 134 có thể được tách thành phần mức trung bình và phần hình của phổ lôga. Các phần này có thể được nội suy theo các cách độc lập như tuyến tính, hình học, hoặc điều hòa (điện trở song song). Như thế, các quy trình nội suy khác nhau có thể được sử dụng để xác định các đường bao được nội suy 136. Quy trình nội suy được sử dụng bởi bộ giải mã 500 thường tương ứng với quy trình nội suy được sử dụng bởi bộ mã hóa 100, 170.

Đơn vị làm mịn đường bao 107 của bộ giải mã đường bao 503 có thể được tạo cấu hình để xác định đường bao phân bố 138 từ đường bao được điều chỉnh 139 bằng cách lượng tử hóa đường bao được điều chỉnh 139 (ví dụ, thành các bước 3 dB). Đường bao phân bố 138 có thể được sử dụng kết hợp với tham số bù hoặc tham số điều chỉnh phân bố (được bao gồm trong dữ liệu hệ số 163) để tạo ra vector phân bố số nguyên trên danh nghĩa được sử dụng để kiểm soát việc giải mã phổ, tức là việc giải mã dữ liệu hệ số 163. Cụ thể là, vector phân bố số nguyên trên danh nghĩa có thể được

sử dụng để xác định bộ lượng tử hóa để đảo ngược lượng tử hóa các chỉ số lượng tử hóa được bao gồm trong dữ liệu hệ số 163. Đường bao phân bố 138 và vector phân bố số nguyên trên danh nghĩa có thể được xác định theo các tương tự trong bộ mã hóa 100, 170 và trong bộ giải mã 500.

Nhằm cho phép bộ giải mã 500 đồng bộ với dòng bit nhận được, có thể truyền các loại khung khác nhau. Khung có thể tương ứng với tập hợp 132, 332 của các khối, cụ thể là khối chuyển đổi 332 của các khối. Cụ thể là, chính các khung P có thể được truyền đi, các khung này được mã hóa theo cách tương đối với khung trước đó. Trong phần mô tả trên, giả sử rằng bộ giải mã 500 nhận biết được đường bao trước đó được lượng tử hóa 135. Đường bao trước đó được lượng tử hóa 135 có thể được bố trí trong khung trước đó, sao cho tập hợp hiện tại 132 hoặc tập hợp chuyển đổi tương ứng 332 có thể tương ứng với khung P. Tuy nhiên, khi khởi động, bộ giải mã 500 thường không nhận biết được đường bao trước đó được lượng tử hóa 135. Để thực hiện mục đích này, khung I có thể được truyền đi (ví dụ, trên cơ sở khởi động hoặc thông thường). Khung I có thể bao gồm hai đường bao, một trong số đó được sử dụng làm đường bao trước đó được lượng tử hóa 135 và đường bao còn lại được sử dụng làm đường bao hiện thời được lượng tử hóa 134. Các khung I có thể được sử dụng cho trong trường hợp khởi động đầu trước phổ lời nói (tức là của bộ giải mã tiếng nói dựa trên biến đổi 500), ví dụ khi theo sau khung sử dụng chế độ mã hóa âm thanh khác và/hoặc dùng làm công cụ để kích hoạt dứt khoát điểm nối của dòng bit âm thanh.

Hoạt động của bộ giải mã dải phụ 517 được minh họa trên Fig.5d. Trong ví dụ minh họa, các tham số bộ dự đoán 520 là tham số cờ và độ lợi tham số bộ dự đoán g . Các tham số bộ dự đoán 520 có thể được xác định từ bộ dự đoán dữ liệu 164 bằng cách sử dụng bảng định trước của các trị số có thể có đối với tham số cờ và độ lợi tham số bộ dự đoán. Điều này cho phép việc truyền các tham số dự đoán 520 đạt hiệu quả về tốc độ bit.

Một hoặc nhiều vector hệ số biến đổi được mã hóa trước đó (tức là một hoặc nhiều khối trước đó 149 của các hệ số được tái tạo) có thể được lưu trữ trong bộ đệm tín hiệu dải phụ (hoặc MDCT) 541. Bộ đệm 541 có thể được cập nhật theo bước (ví dụ, mỗi 5ms). Bộ tách bộ dự đoán 543 có thể được tạo cấu hình để hoạt động trên bộ đệm 541 tùy thuộc vào tham số cờ được chuẩn hóa T . Tham số cờ được chuẩn hóa T có thể được xác định bằng cách chuẩn hóa tham số cờ 520 thành các đơn vị bước (ví

dụ, thành các đơn vị bước MDCT). Nếu tham số T là số nguyên, bộ tách 543 có thể đưa một hoặc nhiều vector hệ số biến đổi được mã hóa trước đó T các đơn vị thời gian vào bộ đệm 541. Nói cách khác, tham số T có thể chỉ báo khối nào trong số một hoặc nhiều khối trước đó 149 của các hệ số được tái tạo sẽ được sử dụng để xác định khối 150 của các hệ số biến đổi được ước lượng. Phần mô tả chi tiết về các phương án có thể thực hiện của bộ tách 543 được mô tả trong đơn yêu cầu cấp bằng độc quyền sáng chế số US61750052 và các đơn yêu cầu cấp bằng độc quyền sáng chế là đơn ưu tiên của nó, nội dung của nó được đưa vào bằng cách viện dẫn.

Bộ tách 543 có thể hoạt động trên các vector (hoặc các khối) mang các đường bao tín hiệu đầy đủ. Mặt khác, khối 150 của các hệ số biến đổi được ước lượng (sẽ được cung cấp bởi bộ giải mã dài phụ 517) được biểu diễn trong miền được làm phẳng. Do đó, đầu ra của bộ tách 543 có thể được tạo hình thành vector miền được làm phẳng. Có thể đạt được điều này nhờ sử dụng bộ tạo hình 544 mà sử dụng các đường bao được điều chỉnh 139 của một hoặc nhiều khối trước đó 149 của các hệ số được tái tạo. Các đường bao được điều chỉnh 139 của một hoặc nhiều khối trước đó 149 của các hệ số được tái tạo có thể được lưu trữ trong bộ đệm đường bao 542. Đơn vị tạo hình 544 có thể được tạo cấu hình để đưa đường bao tín hiệu bị chậm mà sẽ được sử dụng trong việc làm phẳng từ các đơn vị thời gian T_0 vào bộ đệm đường bao 542, trong đó T_0 là số nguyên gần với T nhất. Sau đó, vector miền được làm phẳng có thể được lập thang tỷ lệ bởi tham số độ lợi g để tạo ra khối 150 của các hệ số biến đổi được ước lượng (trong miền được làm phẳng).

Đơn vị tạo hình 544 có thể được tạo cấu hình để xác định vector miền được làm phẳng sao cho các vector miền được làm phẳng tại đầu ra của đơn vị tạo hình 544 thể hiện phương sai đơn vị trong mỗi dải tần số. Đơn vị tạo hình 544 có thể căn cứ toàn bộ trên dữ liệu trong bộ đệm đường bao 542 để đạt được mục đích này. Ví dụ, đơn vị tạo hình 544 có thể được tạo cấu hình để lựa chọn đường bao tín hiệu bị chậm sao cho các vector miền được làm phẳng tại đầu ra của đơn vị tạo hình 544 thể hiện phương sai đơn vị trong mỗi dải tần số. Lựa chọn khác hoặc thêm vào đó, đơn vị tạo hình 544 có thể được tạo cấu hình để đo phương sai của các vector miền được làm phẳng tại đầu ra của đơn vị tạo hình 544 và để điều chỉnh phương sai của các vector theo đặc điểm phương sai đơn vị. Loại chuẩn hóa có thể sử dụng độ lợi băng rộng đơn (mỗi khe) mà làm chuẩn hóa các vector miền được làm phẳng thành vector phương sai đơn vị. Các độ

lợi có thể được truyền đi từ bộ mã hóa 100 đến bộ giải mã tương ứng 500 (ví dụ, dưới dạng được lượng tử hóa và mã hóa) trong dòng bit.

Thay vào đó, quá trình làm phẳng bị chậm được thực hiện bởi bộ tạo hình 544 có thể được lược bỏ bằng cách sử dụng bộ dự đoán dải phụ 517 mà hoạt động trong miền được làm phẳng, ví dụ bộ giải mã dải phụ 517 mà hoạt động trên các khối 148 của các hệ số được làm phẳng được tái tạo. Tuy nhiên, thấy rằng trình tự của các vector miền được làm phẳng (hoặc các khối) không ảnh xạ tốt đến các tín hiệu thời gian do các dấu hiệu thời gian lạ của việc biến đổi (ví dụ, biến đổi MDCT). Kết quả là, việc điều chỉnh cho thích nghi với mô hình tín hiệu cơ sở của bộ tách 543 được làm giảm xuống và mức nhiễu mã hóa cao hơn được tạo ra từ cấu trúc thay thế. Nói cách khác, thấy rằng các mô hình tín hiệu (ví dụ, các mô hình hình sin hoặc theo chu kỳ) được sử dụng bởi bộ giải mã dải phụ 517 làm tăng hiệu suất trong miền không phẳng (so với miền được làm phẳng).

Cần lưu ý rằng trong ví dụ thay thế, đầu ra của bộ dự đoán 517 (tức là khối 150 của các hệ số biến đổi được ước lượng) có thể được cộng tại đầu ra của đơn vị làm phẳng ngược 114 (tức là vào khối 149 của các hệ số được tái tạo) (xem Fig.5a). Đơn vị tạo hình 544 trên Fig.5c khi đó có thể được tạo cấu hình để thực hiện quá trình kết hợp giữa việc làm phẳng bị chậm và việc đảo ngược làm phẳng.

Các thành phần trong dòng bit nhận được có thể kiểm soát việc xáo trộn thi thoảng xảy ra của bộ đệm dải phụ 541 và của bộ đệm đường bao 542, ví dụ trong trường hợp đơn vị mã hóa thứ nhất (tức là khối thứ nhất) của khung I. Điều này cho phép giải mã khung I mà không cần biết đến dữ liệu trước đó. Đơn vị mã hóa thứ nhất sẽ thường không có khả năng sử dụng đóng góp mang tính dự đoán, nhưng dù sao vẫn có thể sử dụng số lượng bit tương đối ít để truyền thông tin bộ dự đoán 520. Sự hao hụt độ lợi dự đoán có thể được bù bằng cách phân bố nhiều bit hơn đến quá trình mã hóa sai số dự đoán của đơn vị mã hóa thứ nhất này. Thông thường, sự đóng góp bộ dự đoán cũng quan trọng đối với đơn vị mã hóa thứ hai (tức là khối thứ hai) của khung I. Do các khía cạnh này nên chất lượng có thể được duy trì cùng với tốc độ bit tăng tương đối ít, thậm chí là với việc sử dụng các khung I rất thường xuyên.

Nói cách khác, các tập hợp 132, 332 của các khối (còn được gọi là các khung) bao gồm nhiều khối 131 mà có thể được mã hóa bằng cách sử dụng mã hóa có tính dự đoán. Khi mã hóa khung I, duy nhất khối thứ nhất 203 của tập hợp 332 của các khối

không thể được mã hóa bằng cách sử dụng độ lợi mã hóa đạt được bởi bộ mã hóa có tính dự đoán. Khối trực tiếp theo sau 201 có thể đã sử dụng các lợi ích của mã hóa có tính dự đoán. Điều này có nghĩa là các mặt hạn chế của khung I về vấn đề hiệu suất mã hóa được hạn chế thành mã hóa khối thứ nhất 203 của các hệ số biến đổi của khung 332, và không áp dụng vào các khối khác 201, 204, 205 của khung 332. Do đó, quy trình mã hóa tiếng nói dựa trên biến đổi được nêu trong bản mô tả này cho phép sử dụng các khung I một cách khá thường xuyên mà không ảnh hưởng đáng kể đến hiệu suất mã hóa. Như thế, quy trình mã hóa tiếng nói dựa trên biến đổi được mô tả ở đây là đặc biệt phù hợp cho các ứng dụng yêu cầu sự đồng bộ hóa tương đối nhanh và/hoặc tương đối thường xuyên giữa bộ giải mã và bộ mã hóa.

Như nêu trên, trong quá trình khởi động của khung I, bộ đệm tín hiệu bộ dự đoán, tức là bộ đệm dải phụ 541, có thể được xáo trộn với các số không và bộ đệm đường bao 542 có thể được làm đầy với chỉ một khe thời của các trị số, tức là có thể được làm đầy với chỉ một đường bao được điều chỉnh 139 (tương ứng với khối thứ nhất 131 của khung I). Khối thứ nhất 131 của khung I thường không sử dụng phép dự đoán. Khối thứ hai 131 có thể truy cập vào duy nhất hai khe thời gian của bộ đệm đường bao 542 (tức là vào các đường bao 139 của các khối thứ nhất và thứ hai 131), khối thứ ba chỉ có thể truy cập vào ba khe thời gian (tức là vào các đường bao 139 của ba khối 131), và khối thứ tư 131 chỉ có thể truy cập vào bốn khe thời gian (tức là vào các đường bao 139 của bốn khối 131).

Nguyên tắc làm phẳng trễ của bộ tạo hình phổ 544 (để nhận biết đường bao để xác định khối 150 của các hệ số biến đổi được ước lượng (trong miền được làm phẳng)) dựa trên trị số độ trễ nguyên T_0 được xác định bằng cách làm tròn tham số bộ dự đoán T trong các đơn vị có kích thước khối K (trong đó đơn vị có kích thước khối có thể được gọi là khe thời gian hoặc khe) đến số nguyên gần nhất. Tuy nhiên, trong trường hợp của khung I, trị số độ trễ nguyên T_0 này có thể hướng đến các mục vào không có sẵn trong bộ đệm đường bao 542. Trên cơ sở này, bộ tạo hình phổ 544 có thể được tạo cấu hình để xác định trị số độ trễ nguyên T_0 sao cho trị số độ trễ nguyên T_0 được giới hạn ở số lượng của các đường bao 139 mà được lưu trữ trong bộ đệm đường bao 542, tức là sao cho trị số độ trễ nguyên T_0 không hướng đến các đường bao 139 không có sẵn trong bộ đệm đường bao 542. Để thực hiện mục đích này, trị số độ trễ nguyên T_0 có thể được giới hạn ở trị số mà là hàm số của chỉ số khối trong

khung hiện tại. Ví dụ, trị số độ trễ nguyên T_0 có thể được giới hạn tới trị số chỉ số của khối hiện tại 131 (mà sẽ được mã hóa) trong khung hiện tại (ví dụ, tới 1 đối với khối thứ nhất 131, tới 2 đối với khối thứ hai 131, tới 3 đối với khối thứ ba 131 và tới 4 đối với khối thứ tư 131 của khung). Bằng cách này, có thể tránh các trạng thái và/hoặc biến dạng không mong muốn do quá trình làm phẳng.

Fig.5d thể hiện sơ đồ khối của bộ giải mã phổ làm ví dụ 502. Bộ giải mã phổ 502 bao gồm bộ giải mã không hao hụt 551 được tạo cấu hình để giải mã dữ liệu hệ số được mã hóa entropi 163. Hơn nữa, bộ giải mã phổ 502 bao gồm bộ đảo ngược lượng tử hóa 552 được tạo cấu hình để ấn định các trị số hệ số cho các chỉ số lượng tử hóa được bao gồm trong dữ liệu hệ số 163. Như đã được nêu trong bối cảnh của bộ mã hóa 100, 170, các hệ số biến đổi khác nhau có thể được lượng tử hóa bằng cách sử dụng các bộ lượng tử hóa khác nhau được chọn từ tập hợp của các bộ lượng tử hóa được định trước, ví dụ tập hợp hữu hạn của các bộ lượng tử hóa vô hướng dựa trên mô hình. Như được thể hiện trên Fig.4, tập hợp của các bộ lượng tử hóa 321, 322, 323 có thể bao gồm các loại bộ lượng tử hóa khác nhau. Tập hợp của các bộ lượng tử hóa có thể bao gồm bộ lượng tử hóa 321 mà tổng hợp nhiều (trong trường hợp tốc độ bit bằng không), một hoặc nhiều bộ lượng tử hóa hòa sắc 322 (đối với các tỉ số tín hiệu trên nhiễu, SNR tương đối thấp, và đối với các tốc độ bit trung bình) và/hoặc một hoặc nhiều bộ lượng tử hóa đơn sắc 323 (đối với các SNR tương đối cao và đối với các tốc độ bit tương đối cao).

Đơn vị làm mịn đường bao 107 có thể được tạo cấu hình để cung cấp đường bao phân bố 138 mà có thể được kết hợp với tham số bù được bao gồm trong dữ liệu hệ số 163 để tạo ra vector phân bố. Vector phân bố chứa trị số nguyên đối với mỗi dải tần số 302. Trị số nguyên đối với dải tần số cụ thể 302 hướng đến điểm biến dạng tốc độ sẽ được sử dụng để đảo ngược lượng tử hóa các hệ số biến đổi của dải cụ thể 302. Nói cách khác, trị số nguyên đối với dải tần số cụ thể 302 hướng đến bộ lượng tử hóa sẽ được sử dụng để đảo ngược lượng tử hóa các hệ số biến đổi của dải cụ thể 302. Trị số nguyên tăng một tương ứng với SNR tăng 1,5 dB. Đối với các bộ lượng tử hóa hòa sắc 322 và các bộ lượng tử hóa đơn sắc 323, mô hình phân phối xác suất Laplacian có thể được sử dụng trong việc mã hóa không hao hụt, mà có thể sử dụng mã hóa số học. Một hoặc nhiều bộ lượng tử hóa hòa sắc 322 có thể được sử dụng để thu nhỏ khoảng cách theo cách liên mạch giữa các trường hợp tốc độ bit thấp và cao. Các bộ lượng tử

hóa hòa sắc 322 có thể có lợi trong việc tạo ra chất lượng âm thanh đầu ra đủ êm dịu đối với các tín hiệu giống như nhiễu cố định.

Nói cách khác, bộ đảo ngược lượng tử hóa 552 có thể được tạo cấu hình để nhận các chỉ số lượng tử hóa hệ số của khối hiện tại 131 của các hệ số biến đổi. Một hoặc nhiều chỉ số lượng tử hóa hệ số của dải tần số cụ thể 302 đã được xác định bằng cách sử dụng bộ lượng tử hóa tương ứng từ tập hợp định trước của các bộ lượng tử hóa. Trị số của vector phân bố (có thể được xác định bằng cách bù đường bao phân bố 138 với tham số bù) đối với dải tần số cụ thể 302 chỉ báo bộ lượng tử hóa mã đã được sử dụng để xác định một hoặc nhiều chỉ số lượng tử hóa hệ số của dải tần số cụ thể 302. Đã nhận biết bộ lượng tử hóa thì một hoặc nhiều chỉ số lượng tử hóa hệ số có thể được đảo ngược lượng tử hóa để tạo ra khối 145 của các hệ số sai số được lượng tử hóa.

Hơn nữa, bộ giải mã phổ 502 có thể bao gồm đơn vị đảo ngược lặp lại thang tỷ lệ 113 để cung cấp khối 147 của các hệ số sai số được lượng tử hóa được lập thang tỷ lệ. Các công cụ bổ sung và các mối liên kết lẫn nhau xung quanh bộ giải mã không hao hụt 551 và bộ đảo ngược lượng tử hóa 552 trên Fig.5d có thể được sử dụng để làm cho việc giải mã phổ thích nghi với công dụng của nó trong bộ giải mã tổng 500 được thể hiện trên Fig.5a, trong đó đầu ra của bộ giải mã phổ 502 (tức là khối 145 của các hệ số sai số được lượng tử hóa) được sử dụng để hiệu chỉnh thêm cho vector miền được làm phẳng và được dự đoán (tức là khối 150 của các hệ số biến đổi được ước lượng). Cụ thể là, các công cụ bổ sung có thể đảm bảo rằng việc xử lý được thực hiện bởi bộ giải mã 500 tương ứng với việc xử lý được thực hiện bởi bộ mã hóa 100, 170.

Cụ thể là, bộ giải mã phổ 502 có thể bao gồm đơn vị lập thang tỷ lệ suy nghiệm 111. Như được thể hiện cùng với bộ mã hóa 100, 170, đơn vị lập thang tỷ lệ suy nghiệm 111 có thể có tác động lên sự phân bố bit. Trong bộ mã hóa 100, 170, các khối hiện tại 141 của các hệ số sai số dự đoán có thể được lập thang tỷ lệ theo phương sai đơn vị bằng nguyên tắc suy nghiệm. Kết quả là, việc phân bố mặc định có thể dẫn đến sự lượng tử hóa quá mịn của đầu ra được giảm tỷ lệ cuối cùng của đơn vị lập thang tỷ lệ suy nghiệm 111. Do đó sự phân bố cần được thay đổi theo cách giống với sự thay đổi của các hệ số sai số dự đoán.

Tuy nhiên, như được nêu dưới đây, có thể có lợi khi tránh làm giảm các nguồn mã hóa đối với một hoặc nhiều trong số các bin tần số thấp (hoặc các dải tần số thấp).

Cụ thể là, việc này có thể có lợi để chống lại thành phần lộn/hiều LF (tần số thấp) mà diễn ra nổi bật nhất trong các trường hợp hữu thanh (tức là đối với tín hiệu có tham số điều chỉnh tương đối lớn 146, rfu). Như thế, việc lựa chọn phân bố bit / bộ lượng tử hóa phụ thuộc vào tham số điều chỉnh 146, mà được mô tả dưới đây, có thể được xem là “tăng chất lượng LF thích nghi tiếng nói”.

Bộ giải mã phổ có thể phụ thuộc vào tham số điều chỉnh 146 tên là rfu, tham số này có thể là phiên bản được giới hạn của độ lợi bộ dự đoán g , ví dụ

$$rfu = \min(1, \max(g, 0)).$$

Các phương pháp thay thế để xác định tham số điều chỉnh 146, rfu, có thể được sử dụng. Cụ thể là, tham số điều chỉnh 146 có thể được xác định bằng cách sử dụng mã giả được cho trong Bảng 1.

```
f_gain = f_pred_gain;
if (f_gain < -1.0)
    f_rfu = 1.0;
else if (f_gain < 0.0)
    f_rfu = -f_gain;
else if (f_gain < 1.0)
    f_rfu = f_gain;
else if (f_gain < 2.0)
    f_rfu = 2.0 - f_gain;
else // f_gain >= 2.0
    f_rfu = 0.0.
```

Bảng 1

variable f_gain và f_pred_gain có thể được thiết lập bằng nhau. Cụ thể là, variable f_gain có thể tương ứng với độ lợi bộ dự đoán g . Tham số điều chỉnh 146, rfu, còn được gọi là f_rfu trong Bảng 1. gain f_gain có thể là số thực.

So với định nghĩa thứ nhất của tham số điều chỉnh 146, định nghĩa thứ hai (theo Bảng 1) làm giảm tham số điều chỉnh 146, rfu, đối với các độ lợi bộ dự đoán trên 1 và làm tăng tham số điều chỉnh 146, rfu, đối với các độ lợi bộ dự đoán âm.

Bằng cách sử dụng tham số điều chỉnh 146, tập hợp của các bộ lượng tử hóa được sử dụng trong đơn vị lượng tử hóa hệ số 112 của bộ mã hóa 100, 170 và được sử dụng trong bộ đảo ngược lượng tử hóa 552 có thể được điều chỉnh. Cụ thể là, tính nhiều của tập hợp của các bộ lượng tử hóa có thể được điều chỉnh dựa trên tham số điều chỉnh 146. Ví dụ, trị số của tham số điều chỉnh 146, r_{fu} , gần đến 1 có thể kích hoạt sự giới hạn khoảng của các mức phân bố bằng cách sử dụng các bộ lượng tử hóa hòa sắc và có thể kích hoạt sự giảm phương sai của mức tổng hợp nhiều. Trong một ví dụ, ngưỡng quyết định hòa sắc tại $r_{fu} = 0.75$ và độ lợi nhiều bằng $1 - r_{fu}$ có thể được thiết lập. Sự thích nghi hòa sắc có thể tác động đến cả giải mã không hao hụt và bộ đảo ngược lượng tử hóa, trong khi sự thích nghi độ lợi nhiều thường chỉ tác động đến bộ đảo ngược lượng tử hóa.

Có thể giả định rằng sự đóng góp bộ dự đoán là quan trọng đối với các trường hợp có lời nói/có âm. Như thế, độ lợi bộ dự đoán tương đối cao g (tức là tham số điều chỉnh tương đối cao 146) có thể chỉ báo tín hiệu tiếng nói có lời nói hoặc có âm. Trong các trường hợp này, việc thêm nhiều liên quan đến hòa sắc hoặc nhiều tường minh (trường hợp phân bố không) khi thử nghiệm đã thể hiện điều ngược lại so chất lượng cảm quan mong muốn của tín hiệu được mã hóa. Kết quả là, số lượng của các bộ lượng tử hóa hòa sắc 322 và/hoặc loại nhiều được sử dụng cho bộ lượng tử hóa tổng hợp nhiều 321 có thể được điều chỉnh dựa trên độ lợi bộ dự đoán g , qua đó nâng cao chất lượng cảm quan tín hiệu tiếng nói được mã hóa.

Như thế, tham số điều chỉnh 146 có thể được sử dụng để điều biến khoảng 324, 325 của các SNR mà các bộ lượng tử hóa hòa sắc 322 được sử dụng cho chúng. Ví dụ, nếu tham số điều chỉnh 146 $r_{fu} < 0,75$ thì khoảng 324 đối với các bộ lượng tử hóa hòa sắc có thể được sử dụng. Nói cách khác, nếu tham số điều chỉnh 146 dưới ngưỡng định trước, tập hợp thứ nhất 326 của các bộ lượng tử hóa có thể được sử dụng. Mặt khác, nếu tham số điều chỉnh 146 $r_{fu} \geq 0,75$ thì khoảng 325 đối với các bộ lượng tử hóa hòa sắc có thể được sử dụng. Nói cách khác, nếu tham số điều chỉnh 146 lớn hơn hoặc bằng ngưỡng định trước, tập hợp thứ hai 327 của các bộ lượng tử hóa có thể được sử dụng.

Hơn nữa, tham số điều chỉnh 146 có thể được sử dụng để điều biến phương sai và phân bố bit. Lý do là thường thì phép dự đoán thành công sẽ yêu cầu phải hiệu chỉnh một chút, đặc biệt là trong khoảng tần số thấp hơn từ 0-1 kHz. Có thể có lợi khi

khuyến bộ lượng tử hóa nhận ra độ lệch từ mô hình phương sai đơn vị nhằm dọn các nguồn mã đến các dải tần số cao hơn 302. Điều này được mô tả trong bối cảnh của Fig.17c ô iii trong công bố đơn quốc tế WO2009/086918, nội dung của nó được đưa vào bằng cách viện dẫn. Trong bộ giải mã 500, sự điều biến này có thể được thực hiện bằng cách điều biến vector phân bố trên danh nghĩa theo nguyên tắc lập thang tỷ lệ suy nghiệm (được áp dụng bằng cách sử dụng đơn vị lập thang tỷ lệ 111), và cùng lúc đó lập thang tỷ lệ đầu ra của bộ đảo ngược lượng tử hóa 552 theo nguyên tắc đảo ngược lập thang tỷ lệ suy nghiệm bằng cách sử dụng đơn vị đảo ngược lập thang tỷ lệ 113. Theo lý thuyết của công bố đơn quốc tế WO2009/086918, nguyên tắc lập thang tỷ lệ suy nghiệm và nguyên tắc đảo ngược lập thang tỷ lệ suy nghiệm phải khớp với nhau. Tuy nhiên, qua thử nghiệm thấy rằng có lời khi bỏ qua việc điều biến phân bố đối với một hoặc nhiều dải tần số thấp nhất 302, nhằm chống lại các vấn đề thi thoảng xảy ra với nhiễu LF (tần số thấp) đối với các bộ phận tín hiệu có lời nói. Việc bỏ qua điều biến phân bố có thể được thực hiện phụ thuộc vào trị số của độ lợi bộ dự đoán g và/hoặc tham số điều chỉnh 146. Cụ thể là, việc bỏ qua điều biến phân bố có thể được thực hiện chỉ khi tham số điều chỉnh 146 vượt quá ngưỡng quyết định hòa sắc.

Như đã nêu trên, bộ mã hóa 100, 170 và/hoặc bộ giải mã 500 có thể bao gồm đơn vị lập thang tỷ lệ 111 mà được tạo cấu hình để lập lại thang tỷ lệ các hệ số sai số dự đoán $\Delta(k)$ để tạo ra khối 142 của các hệ số sai số được lập thang tỷ lệ. Đơn vị lập lại thang tỷ lệ 111 có thể sử dụng một hoặc nhiều nguyên tắc suy nghiệm định trước để thực hiện việc lập lại thang tỷ lệ. Trong một ví dụ, đơn vị lập lại thang tỷ lệ 111 có thể sử dụng nguyên tắc lập thang tỷ lệ suy nghiệm mà bao gồm độ lợi $d(f)$, ví dụ

$$d(f) = 1 + \frac{7 \cdot rfu^2}{1 + \left(\frac{f}{f_0}\right)^3}$$

trong đó tần số đóng ngắt f_0 có thể được thiết lập là, ví dụ, 1000 Hz. Do đó, đơn vị lập lại thang tỷ lệ 111 có thể được tạo cấu hình để áp dụng độ lợi phụ thuộc vào tần số $d(f)$ vào các hệ số sai số dự đoán để tạo ra khối 142 của các hệ số sai số được lập thang tỷ lệ. Đơn vị đảo ngược lập thang tỷ lệ 113 có thể được tạo cấu hình để áp dụng trị số nghịch đảo của độ lợi phụ thuộc vào tần số $d(f)$. Độ lợi phụ thuộc vào tần số $d(f)$ có thể phụ thuộc vào tham số điều chỉnh rfu 146. Trong ví dụ nêu trên, độ lợi $d(f)$ thể hiện đặc điểm thông thấp, sao cho các hệ số sai số dự đoán bị suy giảm hơn tại các tần số cao hơn là tại các tần số thấp và/hoặc sao cho các hệ số sai số dự đoán được tăng cường hơn tại các tần số thấp hơn là tại các tần số cao. Độ lợi $d(f)$ nêu trên

luôn lớn hơn hoặc bằng một. Do đó, theo phương án ưu tiên, nguyên tắc lập thang tỷ lệ suy nghiệm là sao cho các hệ số sai số dự đoán được tăng cường bởi hệ số một hoặc hơn (tùy thuộc vào tần số).

Cần lưu ý rằng độ lợi phụ thuộc vào tần số có thể chỉ báo công suất hoặc phương sai. Trong các trường hợp như thế, nguyên tắc lập thang tỷ lệ và nguyên tắc đảo ngược lập thang tỷ lệ được lấy dựa trên căn bậc hai của độ lợi phụ thuộc vào tần số, ví dụ dựa trên $\sqrt{d(f)}$.

Mức độ tăng cường và/hoặc suy giảm có thể phụ thuộc vào chất lượng của sự dự đoán đạt được bởi bộ dự đoán 117. Độ lợi bộ dự đoán g và/hoặc tham số điều chỉnh rfu 146 có thể chỉ báo chất lượng của sự dự đoán. Cụ thể là, trị số tương đối thấp của tham số điều chỉnh rfu 146 (tương đối gần với không) có thể chỉ báo chất lượng dự đoán thấp. Trong các trường hợp như thế, giả sử các hệ số sai số dự đoán có các trị số (tuyệt đối) tương đối cao qua tất cả các tần số. Trị số tương đối cao của tham số điều chỉnh rfu 146 (tương đối gần với một) có thể chỉ báo chất lượng dự đoán cao. Trong các trường hợp như thế, giả sử các hệ số sai số dự đoán có các trị số (tuyệt đối) tương đối cao đối với các tần số cao (mà khó dự đoán hơn). Do đó, nhằm đạt được phương sai đơn vị tại đầu ra của đơn vị lập lại thang tỷ lệ 111, độ lợi $d(f)$ có thể là sao cho trong trường hợp chất lượng dự đoán tương đối thấp, độ lợi $d(f)$ về cơ bản là phẳng đối với tất cả tần số, trong khi trong trường hợp chất lượng dự đoán tương đối cao, độ lợi $d(f)$ có đặc điểm thông thấp, làm tăng hoặc khuếch đại phương sai tại các tần số thấp. Đây là trường hợp đối với độ lợi phụ thuộc vào rfu $d(f)$ nêu trên.

Như đã nêu trên, đơn vị phân bố bit 110 có thể được tạo cấu hình để đưa sự phân bố bit tương đối đến các hệ số sai số được lập thang tỷ lệ khác nhau, tùy thuộc vào trị số năng lượng tương ứng trong đường bao phân bố 138. Đơn vị phân bố bit 110 có thể được tạo cấu hình để xem xét nguyên tắc lập thang tỷ lệ suy nghiệm. Nguyên tắc lập thang tỷ lệ suy nghiệm có thể phụ thuộc vào chất lượng của sự dự đoán. Trong trường hợp chất lượng dự đoán tương đối cao, có thể có lợi khi ấn định số lượng bit tương đối tăng vào mã hóa các hệ số sai số dự đoán (hoặc khối 142 của các hệ số sai số được lập thang tỷ lệ) tại các tần số cao hơn là mã hóa các hệ số tại các tần số thấp. Điều này có thể là do trong trường hợp chất lượng dự đoán cao, các hệ số tần số thấp được dự đoán tốt rồi, trong khi các hệ số tần số cao thường được dự đoán

không tốt bằng. Mặt khác, trong trường hợp chất lượng dự đoán tương đối thấp, việc phân bố bit cần giữ không đổi.

Điều trên có thể được thực hiện bằng cách áp dụng đảo ngược các nguyên tắc suy nghiệm / độ lợi $d(f)$ vào đường bao được điều chỉnh hiện tại 139, nhằm xác định đường bao phân bố 138 mà xem xét chất lượng của phép dự đoán.

Đường bao được điều chỉnh 139, các hệ số sai số dự đoán và độ lợi $d(f)$ có thể được biểu diễn trong miền log hoặc dB. Trong trường hợp này, việc áp dụng độ lợi $d(f)$ vào các hệ số sai số dự đoán có thể tương ứng với thao tác “cộng” và việc áp dụng trị số nghịch đảo của độ lợi $d(f)$ vào đường bao được điều chỉnh 139 có thể tương ứng với thao tác “trừ”.

Cần lưu ý rằng có thể có nhiều biến thể của các nguyên tắc suy nghiệm / độ lợi $d(f)$. Cụ thể là, các đường cong phụ thuộc vào tần số cố định của đặc tính thông thấp $\left(1 + \left(\frac{f}{f_0}\right)^3\right)^{-1}$ có thể được thay thế bằng hàm số mà phụ thuộc vào dữ liệu đường bao (ví dụ, trên đường bao được điều chỉnh 139 đối với khối hiện tại 131). Các nguyên tắc suy nghiệm thay đổi có thể phụ thuộc vào cả tham số điều chỉnh rfu 146 và dữ liệu đường bao.

Sau đây các cách khác nhau để xác định độ lợi bộ dự đoán ρ , mà có thể tương ứng với độ lợi bộ dự đoán g , sẽ được mô tả. Độ lợi bộ dự đoán ρ có thể được sử dụng làm dấu hiệu của chất lượng của sự dự đoán. Vector dư dự đoán (tức là khối 141 của các hệ số sai số dự đoán z có thể được cho bởi: $z = x - \rho y$, trong đó x là vector đích (ví dụ, khối hiện tại 140 của các hệ số biến đổi được làm phẳng hoặc khối hiện tại 131 của các hệ số biến đổi), y là vector biểu diễn những gì có thể dùng cho phép dự đoán (ví dụ, các khối trước đó 149 của các hệ số được tái tạo), và ρ là độ lợi bộ dự đoán vô hướng.

$w \geq 0$ có thể là vector gia trọng được sử dụng để xác định độ lợi bộ dự đoán ρ . Theo một số phương án, vector gia trọng là hàm số của đường bao tín hiệu (ví dụ, hàm số của đường bao được điều chỉnh 139, mà có thể được ước lượng tại bộ mã hóa 100, 170 và sau đó được truyền đến bộ giải mã 500). Vector gia trọng thường cùng chiều với vector đích và vector ứng cử. Mục vào thứ i -th của vector x có thể được thể hiện bởi x_i (ví dụ, $i=1, \dots, K$).

Có các cách khác nhau để định nghĩa độ lợi bộ dự đoán ρ . Theo một phương án, độ lợi bộ dự đoán ρ là độ lợi MMSE (minimum mean square error – sai số bình phương nhỏ nhất) được định nghĩa theo tiêu chí sai số bình phương nhỏ nhất. Trong trường hợp này, độ lợi bộ dự đoán ρ có thể được tính bằng cách sử dụng công thức sau:

$$\rho = \frac{\sum_i x_i y_i}{\sum_i y_i^2}.$$

Độ lợi bộ dự đoán ρ này thường giảm thiểu sai số bình phương được định nghĩa là $D = \sum_i (x_i - \rho y_i)^2$.

Thường có lợi (về mặt cảm quan) khi đưa gia trọng vào định nghĩa của sai số bình phương D . Gia trọng có thể được sử dụng để tăng cường tính quan trọng của sự liên kết giữa x và y đối với các phân quan trọng về mặt cảm quan của phổ tín hiệu và giảm cường tính quan trọng của sự liên kết giữa x và y đối với các phần của phổ tín hiệu mà tương đối ít quan trọng hơn. Phương pháp này dẫn đến tiêu chí sai số sau đây: $D = \sum_i (x_i - \rho y_i)^2 w_i$, điều này dẫn đến tiêu chí sau đây của độ lợi bộ dự đoán tối ưu (xét về sai số bình phương được gia trọng):

$$\rho = \frac{\sum_i w_i x_i y_i}{\sum_i w_i y_i^2}.$$

Định nghĩa nêu trên về độ lợi bộ dự đoán thường dẫn đến độ lợi không giới hạn. Như nêu trên, các gia trọng w_i của vector gia trọng w có thể được xác định dựa trên đường bao được điều chỉnh 139. Ví dụ, vector gia trọng w có thể được xác định bằng cách sử dụng hàm định nghĩa trước của đường bao được điều chỉnh 139. Hàm định nghĩa trước có thể được nhận biết tại bộ mã hóa và tại bộ giải mã (đây cũng là trường hợp đối với đường bao được điều chỉnh 139). Do đó, vector gia trọng có thể được xác định theo cùng một cách tại bộ mã hóa và tại bộ giải mã.

Công thức khác để tính độ lợi bộ dự đoán có thể có được đưa ra bởi

$$\rho = \frac{2C}{E_x + E_y},$$

trong đó $C = \sum_i w_i x_i y_i$, $E_x = \sum_i w_i x_i^2$ và $E_y = \sum_i w_i y_i^2$. Định nghĩa này về độ lợi

bộ dự đoán tạo ra độ lợi luôn nằm trong khoảng $[-1, 1]$. Đặc điểm quan trọng của độ lợi bộ dự đoán được nêu trong công thức sau đó là độ lợi bộ dự đoán ρ tạo điều kiện cho mối liên hệ dễ vận dụng giữa năng lượng của tín hiệu đích x và năng lượng của tín hiệu dư z . Năng lượng dư LTP có thể được biểu diễn là: $\sum_i w_i z_i^2 = E_x(1 - \rho^2)$.

Tham số điều chỉnh rfu 146 có thể được xác định dựa trên độ lợi bộ dự đoán g bằng cách sử dụng các công thức nêu trên. Độ lợi bộ dự đoán g có thể bằng độ lợi bộ dự đoán ρ , được xác định bằng cách sử dụng công thức bất kỳ trong số các công thức nêu trên.

Như đã nêu trên, bộ mã hóa 100, 170 được tạo cấu hình để lượng tử hóa và mã hóa vector dư z (tức là khối 141 của các hệ số sai số dự đoán). Quá trình lượng tử hóa thường được dẫn hướng bởi đường bao tín hiệu (ví dụ, bởi đường bao phân bố 138) theo mô hình cảm quan cơ sở nhằm phân bố các bit có sẵn trong số các bộ phận thuộc phổ của tín hiệu theo cách có ý nghĩa về mặt cảm quan. Quá trình phân bố tốc độ được dẫn hướng bởi đường bao tín hiệu (ví dụ, bởi đường bao phân bố 138), đường bao tín hiệu này được lấy từ tín hiệu đầu vào (ví dụ, từ khối 131 của các hệ số biến đổi). Hoạt động của bộ dự đoán 117 thường làm thay đổi đường bao tín hiệu. Đơn vị lượng tử hóa 112 thường sử dụng các bộ lượng tử hóa mà được thiết kế có hoạt động giả thiết trên nguồn phương sai đơn vị. Đặc biệt là trong trường hợp dự đoán chất lượng cao (tức là khi bộ dự đoán 117 thành công), đặc tính phương sai đơn vị có thể không nằm trong trường hợp này nữa, tức là khối 141 của các hệ số sai số dự đoán có thể không thể hiện phương sai đơn vị.

Thường không hiệu quả khi ước lượng đường bao của khối 141 của các hệ số sai số dự đoán (tức là đối với tín hiệu dư z) và khi truyền đường bao này đến bộ giải mã (và khi làm phẳng lại khối 141 của các hệ số sai số dự đoán bằng cách sử dụng đường bao được ước lượng). Thay vào đó, bộ mã hóa 100 và bộ giải mã 500 có thể sử dụng nguyên tắc suy nghiệm để tái lập thang tỷ lệ khối 141 của các hệ số sai số dự đoán (như đã nêu trên). Nguyên tắc suy nghiệm có thể được sử dụng để lập lại thang tỷ lệ khối 141 của các hệ số sai số dự đoán, sao cho khối 142 của các hệ số được lập lại thang tỷ lệ đến gần phương sai đơn vị. Kết quả của việc này là, các kết quả lượng

tử hóa có thể được cải thiện (bằng cách sử dụng các bộ lượng tử hóa mà giả sử phương sai đơn vị).

Hơn nữa, như đã nêu trên, nguyên tắc suy nghiệm có thể được sử dụng để điều biến đường bao phân bố 138, mà được dùng cho quá trình phân bố bit. Việc điều biến đường bao phân bố 138 và lập lại thang tỷ lệ khối 141 của các hệ số sai số dự đoán thường được thực hiện bởi bộ mã hóa 100 và bộ giải mã 500 theo cùng một cách (bằng cách sử dụng cùng nguyên tắc suy nghiệm đó).

Nguyên tắc suy nghiệm có thể có $d(f)$ được mô tả như trên. Sau đây, phương pháp khác để xác định nguyên tắc suy nghiệm được mô tả. Trị số nghịch đảo của độ lợi dự đoán năng lượng miền được gia trọng có thể được cho bởi $p \in [0,1]$ sao cho $\|z\|_w^2 = p\|x\|_w^2$, trong đó $\|z\|_w^2$ chỉ báo năng lượng bình phương của vector dư (tức là khối 141 của các hệ số sai số dự đoán) trong miền được gia trọng và trong đó $\|x\|_w^2$ chỉ báo năng lượng bình phương của vector đích (tức là khối 140 của các hệ số biến đổi được làm phẳng) trong miền được gia trọng.

Có thể thực hiện các giả thiết sau

1. Các mục vào của vector đích x có phương sai đơn vị. Đây có thể là kết quả của việc làm phẳng được thực hiện bởi đơn vị làm phẳng 108. Giả thiết này được thỏa mãn tùy thuộc vào chất lượng của việc làm phẳng dựa trên đường bao được thực hiện bởi đơn vị làm phẳng 108.

2. Phương sai của các mục vào của vector dư dự đoán z ở dạng $E\{z^2(i)\} = \min\left\{\frac{t}{w(i)}, 1\right\}$ với $i = 1, \dots, K$ và với một số $t \geq 0$. Giả thiết này là dựa trên suy nghiệm mà tra cứu bộ dự đoán định hướng bình phương nhỏ nhất dẫn đến đóng góp sai số được phân bố đều nhau trong miền được gia trọng, sao cho vector dư $\sqrt{w}z$ phẳng hơn hoặc ít phẳng hơn. Hơn nữa, có thể cho rằng bộ dự đoán dự phòng là gần phẳng mà dẫn đến biên hợp lý $E\{z^2(i)\} \leq 1$. Cần lưu ý rằng nhiều cải biến của giả thiết thứ hai này có thể được sử dụng.

3. Nhằm ước tính tham số t , có thể chèn hai giả thiết nêu trên vào công thức sai số dự đoán (ví dụ, $D = \sum_i (x_i - \rho y_i)^2 w_i$) và qua đó đưa ra phương trình “kiểu mực nước” $\sum_i \min\{t, w(i)\} = p \sum_i w(i)$

Có thể thấy rằng có cách giải cho phương trình trên trong khoảng $t \in [0, \max(w(i))]$. Phương trình để tìm tham số t có thể được giải bằng cách sử dụng các tiện ích phân loại.

Sau đó, nguyên tắc suy nghiệm có thể được cho bởi $d(i) = \max\left\{\frac{w(i)}{t}, 1\right\}$, trong đó $i = 1, \dots, K$ nhận biết bin tần số. Nguyên tắc đảo ngược lập thang tỷ lệ suy nghiệm được cho bởi $\frac{1}{d(i)} = \min\left\{\frac{t}{w(i)}, 1\right\}$. Nguyên tắc đảo ngược lập thang tỷ lệ suy nghiệm được áp dụng bởi đơn vị đảo ngược lập thang tỷ lệ 113. Nguyên tắc lập thang tỷ lệ phụ thuộc vào tần số phụ thuộc vào các gia trọng $w(i) = w_i$. Như nêu trên, các gia trọng $w(i)$ có thể phụ thuộc vào hoặc có thể tương ứng với khối hiện tại 131 của các hệ số biến đổi (ví dụ, đường bao được điều chỉnh 139, hoặc một số hàm định trước của đường bao được điều chỉnh 139).

Có thể thấy rằng khi sử dụng công thức $\rho = \frac{2C}{E_x + E_y}$ để xác định độ lợi bộ dự

đoán, hệ thức sau áp dụng: $p = 1 - \rho^2$.

Do đó, nguyên tắc lập thang tỷ lệ suy nghiệm có thể được xác định theo nhiều cách khác nhau. Qua thử nghiệm, thấy rằng nguyên tắc lập thang tỷ lệ được xác định dựa trên hai giả thiết nêu trên (còn gọi là phương pháp lập thang tỷ lệ B) là có lợi hơn so với nguyên tắc lập thang tỷ lệ cố định $d(f)$. Cụ thể là, nguyên tắc lập thang tỷ lệ mà được xác định dựa trên hai giả thiết có thể xem xét hiệu quả của việc gia trọng được sử dụng khi tra cứu bộ dự đoán dự phòng. Phương pháp lập thang tỷ lệ B được kết hợp một cách thuận lợi với định nghĩa của độ lợi $\rho = \frac{2C}{E_x + E_y}$, do mối liên hệ dễ vận dụng về mặt phân tích giữa phương sai của số dư và phương sai của tín hiệu (mà tạo điều kiện cho việc dẫn xuất p như đã nêu trên).

Sau đây, khía cạnh khác nhằm nâng cao hiệu suất của bộ mã hóa-giải mã âm thanh dựa trên biến đổi được mô tả. Cụ thể là, việc sử dụng chính cờ bảo toàn phương sai được đề xuất. Cờ bảo toàn phương sai có thể được xác định và truyền trên cơ sở mỗi khối 131. Cờ bảo toàn phương sai có thể chỉ báo chất lượng của phép dự đoán. Theo một phương án, cờ bảo toàn phương sai tắt, trong trường hợp chất lượng dự đoán tương đối cao, và cờ bảo toàn phương sai bật, trong trường hợp chất lượng dự đoán tương đối thấp. Cờ bảo toàn phương sai có thể được xác định bởi bộ mã hóa 100, 170, ví dụ dựa trên độ lợi bộ dự đoán ρ và/hoặc dựa trên độ lợi bộ dự đoán g . Ví dụ,

cờ bảo toàn phương sai có thể được thiết lập là “bật” nếu độ lợi bộ dự đoán ρ hoặc g (hoặc tham số được lấy từ đó) dưới ngưỡng định trước (ví dụ, 2dB) và ngược lại. Như đã nêu trên, trị số nghịch đảo của độ lợi dự đoán năng lượng miền được gia trọng p thường phụ thuộc vào độ lợi bộ dự đoán, ví dụ $p = 1 - \rho^2$. Trị số nghịch đảo của tham số p có thể được sử dụng để xác định trị số của cờ bảo toàn phương sai. Ví dụ, $1/p$ (ví dụ, được thể hiện bằng dB) có thể so với ngưỡng định trước (ví dụ, 2dB), nhằm xác định trị số của cờ bảo toàn phương sai. Nếu $1/p$ lớn hơn ngưỡng định trước, cờ bảo toàn phương sai có thể được thiết lập là “tắt” (chỉ báo chất lượng dự đoán tương đối cao), và ngược lại.

Cờ bảo toàn phương sai có thể được sử dụng để kiểm soát nhiều thiết lập khác nhau của bộ mã hóa 100 và của bộ giải mã 500. Cụ thể là, cờ bảo toàn phương sai có thể được sử dụng để kiểm soát độ nhiễu của nhiều bộ lượng tử hóa 321, 322, 323. Cụ thể là, cờ bảo toàn phương sai có thể ảnh hưởng đến một hoặc nhiều thiết lập sau đây

- Độ lợi nhiễu thích nghi đối với phân bố bit không. Nói cách khác, độ lợi nhiễu của bộ lượng tử hóa tổng hợp nhiễu 321 có thể bị ảnh hưởng bởi cờ bảo toàn phương sai.
- Khoảng các bộ lượng tử hóa hòa sắc. Nói cách khác, khoảng 324, 325 của các SNR mà các bộ lượng tử hóa hòa sắc 322 được sử dụng có thể bị ảnh hưởng bởi cờ bảo toàn phương sai.
- Độ lợi sau của các bộ lượng tử hóa hòa sắc. Độ lợi sau có thể được áp dụng vào đầu ra của các bộ lượng tử hóa hòa sắc, nhằm ảnh hưởng đến hiệu suất sai số bình phương của các bộ lượng tử hóa hòa sắc. Độ lợi sau có thể phụ thuộc vào cờ bảo toàn phương sai.
- Áp dụng lập thang tỷ lệ suy nghiệm. Việc sử dụng lập thang tỷ lệ suy nghiệm (trong đơn vị lập lại thang tỷ lệ 111 và trong đơn vị đảo ngược lập thang tỷ lệ 113) có thể phụ thuộc vào cờ bảo toàn phương sai.

Ví dụ về cách cờ bảo toàn phương sai có thể thay đổi một hoặc nhiều thiết lập của bộ mã hóa 100 và/hoặc bộ giải mã 500 được đưa ra trên Bảng 2.

Loại thiết lập	Bảo toàn phương sai tắt	Bảo toàn phương sai bật
Độ lợi nhiễu	$g_N = (1 - \text{rfu})$	$g_N = \sqrt{(1 - \text{rfu}^2)}$
Khoảng của các bộ	Phụ thuộc vào tham số	Được cố định vào khoảng

lượng tử hóa hòa sắc	điều chỉnh rfu	tương đối rộng (ví dụ, vào khoảng có thể là rộng nhất)
Độ lợi sau của các bộ lượng tử hóa hòa sắc. $\gamma_o = \frac{\sigma_x^2}{\sigma_x^2 + \frac{\Delta^2}{12}}; \gamma_1 = \sqrt{\gamma_o}$	$\gamma = \gamma_o$	$\gamma = \max(\gamma_o, g_N \cdot \gamma_1)$
Nguyên tắc lập thang tỷ lệ suy nghiệm	Bật	Tắt

Bảng 2

Trong công thức đối với độ lợi sau, $\sigma_x^2 = E\{X^2\}$ là phương sai của một hoặc nhiều của các hệ số của khối 141 của các hệ số sai số dự đoán (mà sẽ được lượng tử hóa), và Δ là kích thước bước bộ lượng tử hóa của bộ lượng tử hóa vô hướng (612) của bộ lượng tử hóa hòa sắc mà độ lợi sau được áp dụng vào.

Như có thể thấy từ ví dụ trên bảng 2, độ lợi nhiễu g_N của bộ lượng tử hóa tổng hợp nhiễu 321 (tức là phương sai của bộ lượng tử hóa tổng hợp nhiễu 321) có thể phụ thuộc vào cờ bảo toàn phương sai. Như đã nêu trên, tham số điều chỉnh rfu 146 có thể nằm trong khoảng $[0, 1]$, trong đó trị số tương đối thấp của rfu chỉ báo chất lượng dự đoán tương đối thấp và trị số tương đối cao của rfu chỉ báo chất lượng dự đoán tương đối cao. Đối với các trị số rfu trong khoảng $[0, 1]$, công thức ở cột trái cung cấp các độ lợi nhiễu thấp hơn g_N so với công thức ở cột phải. Do đó, khi cờ bảo toàn phương sai bật (chỉ báo chất lượng dự đoán tương đối thấp), độ lợi nhiễu cao hơn được sử dụng hơn là khi cờ bảo toàn phương sai tắt (chỉ báo chất lượng dự đoán tương đối cao). Qua thử nghiệm, thấy rằng việc này làm cải thiện chất lượng cảm quan nói chung.

Như đã nêu trên, khoảng SNR của khoảng 324, 325 của các bộ lượng tử hóa hòa sắc 322 có thể thay đổi tùy thuộc vào tham số điều chỉnh rfu. Theo Bảng 2, khi cờ bảo toàn phương sai bật (chỉ báo chất lượng dự đoán tương đối thấp), khoảng lớn cố định của các bộ lượng tử hóa hòa sắc 322 được sử dụng (ví dụ, khoảng 324). Mặt

khác, khi cỡ bảo toàn phương sai tất (chỉ báo chất lượng dự đoán tương đối cao), các khoảng khác nhau 324, 325 được sử dụng, tùy thuộc vào tham số điều chỉnh rfu.

Việc xác định khối 145 của các hệ số sai số được lượng tử hóa có thể bao gồm việc áp dụng độ lợi sau γ vào các hệ số sai số được lượng tử hóa, mà đã được lượng tử hóa bằng cách sử dụng bộ lượng tử hóa hòa sắc 322. Độ lợi sau γ có thể được lấy ra để nâng cao hiệu suất MSE của bộ lượng tử hóa hòa sắc 322 (ví dụ, bộ lượng tử hóa có hòa sắc trừ). Độ lợi sau có thể được cho bởi:

$$\gamma = \frac{\sigma_x^2}{\sigma_x^2 + \frac{\Delta^2}{12}}.$$

Qua thử nghiệm, thấy rằng chất lượng mã hóa cảm quan có thể được cải thiện khi khiến độ lợi sau phụ thuộc vào cỡ bảo toàn phương sai. Độ lợi sau tối ưu MSE nêu trên được sử dụng khi cỡ bảo toàn phương sai tất (chỉ báo chất lượng dự đoán tương đối cao). Mặt khác, khi cỡ bảo toàn phương sai bất (chỉ báo chất lượng dự đoán tương đối thấp), có thể có lợi khi sử dụng độ lợi sau cao hơn (được xác định theo công thức nằm bên phải của Bảng 2).

Như đã nêu trên, lập thang tỷ lệ suy nghiệm có thể được sử dụng để đưa ra các khối 142 của các hệ số sai số được lập thang tỷ lệ mà gần hơn đối với đặc tính phương sai đơn vị so với các khối 141 của các hệ số sai số dự đoán. Các nguyên tắc lập thang tỷ lệ suy nghiệm có thể được làm cho phụ thuộc vào tham số điều chỉnh 146. Nói cách khác, các nguyên tắc lập thang tỷ lệ suy nghiệm có thể được làm cho phụ thuộc vào chất lượng của phép dự đoán. Lập thang tỷ lệ suy nghiệm có thể là đặc biệt có lợi trong trường hợp chất lượng dự đoán tương đối cao, trong khi đó các lợi ích này có thể được giới hạn trong trường hợp chất lượng dự đoán tương đối thấp. Trên cơ sở này, có thể có lợi nếu chỉ sử dụng lập thang tỷ lệ suy nghiệm khi cỡ bảo toàn phương sai tất (chỉ báo chất lượng dự đoán tương đối cao).

Trong bản mô tả này, bộ mã hóa tiếng nói dựa trên biến đổi 100, 170 và bộ giải mã tương ứng tiếng nói dựa trên biến đổi 500 đã được mô tả. Bộ mã hóa-giải mã tiếng nói dựa trên biến đổi có thể sử dụng nhiều dấu hiệu mà cho phép nâng cao chất lượng của các tín hiệu tiếng nói được mã hóa. Bộ mã hóa-giải mã tiếng nói có thể sử dụng các khối tương đối ngắn (còn được gọi là các đơn vị mã hóa), ví dụ trong khoảng 5 ms, qua đó đảm bảo độ phân giải thời gian phù hợp và thống kê có ý nghĩa đối với các

tín hiệu tiếng nói. Hơn nữa, bộ mã hóa-giải mã tiếng nói có thể cung cấp bản mô tả đầy đủ của đường bao phổ biến đổi theo thời gian của các đơn vị mã hóa. Ngoài ra, bộ mã hóa-giải mã tiếng nói có thể sử dụng phép dự đoán trong miền biến đổi, trong đó phép dự đoán có thể xem xét các đường bao phổ của các đơn vị mã hóa. Do đó, bộ mã hóa-giải mã tiếng nói có thể cung cấp các cập nhật có tính dự đoán nhận biết đường bao cho các đơn vị mã hóa. Hơn nữa, bộ mã hóa-giải mã tiếng nói có thể sử dụng các bộ lượng tử hóa định trước mà thích nghi với các kết quả của phép dự đoán. Nói cách khác, bộ mã hóa-giải mã tiếng nói có thể sử dụng các bộ lượng tử hóa vô hướng thích nghi với dự đoán.

Các phương pháp và hệ thống được nêu trong bản mô tả này có thể được thực hiện dưới dạng phần mềm, phần sụn và/hoặc phần cứng. Các bộ phận nhất định có thể, ví dụ, được thực hiện dưới dạng phần mềm chạy trên bộ xử lý hoặc bộ vi xử lý tín hiệu kỹ thuật số. Các bộ phận khác có thể, ví dụ, được thực hiện dưới dạng phần cứng và hoặc dưới dạng các mạch tích hợp chuyên dụng. Các tín hiệu có trong các phương pháp và hệ thống được mô tả có thể được lưu trữ trên các phương tiện như bộ nhớ truy cập ngẫu nhiên hoặc phương tiện lưu trữ quang học. Chúng có thể được chuyển qua các loại mạng, như mạng vô tuyến, mạng vệ tinh, mạng không dây hoặc mạng có dây, ví dụ mạng Internet. Các thiết bị tiêu biểu mà sử dụng các phương pháp và hệ thống được nêu trong bản mô tả này là các thiết bị điện tử xách tay hoặc thiết bị tiêu dùng khác mà được sử dụng để lưu trữ và/hoặc kết xuất các tín hiệu âm thanh.

YÊU CẦU BẢO HỘ

1. Bộ mã hóa tiếng nói dựa trên biến đổi được tạo cấu hình để mã hóa tín hiệu tiếng nói thành dòng bit, bộ mã hóa này bao gồm:

đơn vị tạo khung được tạo cấu hình để nhận tập hợp các khối; trong đó tập hợp các khối bao gồm nhiều khối liên tiếp gồm các hệ số biến đổi; trong đó nhiều khối chỉ báo các mẫu của tín hiệu tiếng nói; trong đó khối hệ số biến đổi bao gồm nhiều hệ số biến đổi cho nhiều bin tần số tương ứng;

đơn vị ước lượng đường bao được tạo cấu hình để xác định đường bao hiện thời dựa trên nhiều khối liên tiếp gồm các hệ số biến đổi; trong đó đường bao hiện thời chỉ báo nhiều trị số năng lượng phổ cho nhiều bin tần số tương ứng;

đơn vị lượng tử hóa đường bao được tạo cấu hình để xác định đường bao hiện thời được lượng tử hóa bằng cách lượng tử hóa đường bao hiện thời;

đơn vị nội suy đường bao được tạo cấu hình để xác định nhiều đường bao được nội suy cho nhiều khối hệ số biến đổi lần lượt, dựa trên đường bao hiện thời được lượng tử hóa và dựa trên đường bao trước đó được lượng tử hóa; và

đơn vị làm phẳng được tạo cấu hình để xác định nhiều khối hệ số biến đổi được làm phẳng bằng cách làm phẳng nhiều khối tương ứng gồm các hệ số biến đổi nhờ sử dụng nhiều đường bao được nội suy tương ứng, một cách lần lượt; trong đó dòng bit được xác định dựa trên nhiều khối hệ số biến đổi được làm phẳng và tạo ra tín hiệu có thể nghe được.

2. Bộ mã hóa tiếng nói dựa trên biến đổi theo điểm 1, trong đó:

bộ mã hóa tiếng nói dựa trên biến đổi còn bao gồm đơn vị xác định độ lợi đường bao được tạo cấu hình để xác định lần lượt các độ lợi đường bao cho các khối hệ số biến đổi;

bộ mã hóa tiếng nói dựa trên biến đổi còn bao gồm đơn vị làm mịn đường bao được tạo cấu hình để xác định nhiều đường bao được điều chỉnh bằng cách bù các trị số năng lượng phổ của các đường bao được nội suy theo nhiều độ lợi đường bao, một cách lần lượt;

đơn vị làm phẳng được tạo cấu hình để xác định nhiều khối hệ số biến đổi được làm phẳng bằng cách làm phẳng nhiều khối tương ứng gồm các hệ số biến đổi nhờ sử dụng nhiều đường bao được điều chỉnh tương ứng một cách lần lượt.

3. Bộ mã hóa tiếng nói dựa trên biến đổi theo điểm 2, trong đó đơn vị xác định độ lợi đường bao được tạo cấu hình để xác định độ lợi đường bao thứ nhất đối với khối thứ nhất gồm các hệ số biến đổi, sao cho phương sai của các hệ số biến đổi được làm phẳng của khối thứ nhất tương ứng gồm các hệ số biến đổi được làm phẳng thu được nhờ sử dụng đường bao đã điều chỉnh thứ nhất được điều chỉnh so với phương sai của các hệ số biến đổi được làm phẳng của khối thứ nhất tương ứng gồm các hệ số biến đổi được làm phẳng thu được nhờ sử dụng đường bao được nội suy thứ nhất.

4. Bộ mã hóa tiếng nói dựa trên biến đổi theo điểm 3, trong đó đơn vị xác định độ lợi đường bao được tạo cấu hình để xác định độ lợi đường bao thứ nhất cho khối thứ nhất gồm các hệ số biến đổi, sao cho phương sai của các hệ số biến đổi được làm phẳng của khối thứ nhất tương ứng gồm các hệ số biến đổi được làm phẳng thu được nhờ sử dụng đường bao được điều chỉnh thứ nhất là một.

5. Bộ mã hóa tiếng nói dựa trên biến đổi theo điểm 2, trong đó đơn vị xác định độ lợi đường bao được tạo cấu hình để chèn dữ liệu độ lợi chỉ báo các độ lợi đường bao vào dòng bit.

6. Bộ mã hóa tiếng nói dựa trên biến đổi theo điểm 1, trong đó:

đường bao hiện thời chỉ báo các trị số năng lượng phổ cho nhiều dải tần số tương ứng;

dải tần số bao gồm một hoặc nhiều bin tần số;

đơn vị ước lượng đường bao được tạo cấu hình để xác định trị số năng lượng phổ cho dải tần số cụ thể dựa trên các hệ số biến đổi của nhiều khối liên tiếp cho dải tần số cụ thể.

7. Bộ mã hóa tiếng nói dựa trên biến đổi theo điểm 1, trong đó đơn vị lượng tử hóa đường bao được tạo cấu hình để chèn dữ liệu đường bao vào dòng bit chỉ báo đường bao hiện thời được lượng tử hóa.

8. Bộ mã hóa tiếng nói dựa trên biến đổi theo điểm 1, trong đó:

khối hệ số biến đổi bao gồm các hệ số MDCT; và/hoặc

khối hệ số biến đổi bao gồm các hệ số biến đổi trong các bin tần số; và/hoặc

tập hợp các khối bao gồm bốn khối hệ số biến đổi hoặc nhiều hơn.

9. Bộ mã hóa tiếng nói dựa trên biến đổi theo điểm 1, trong đó:

bộ mã hóa tiếng nói dựa trên biến đổi được tạo cấu hình để hoạt động trong nhiều chế độ khác nhau bao gồm chế độ bước ngắn và chế độ bước dài;

đơn vị tạo khung, đơn vị ước lượng đường bao và đơn vị nội suy đường bao được tạo cấu hình để xử lý tập hợp các khối bao gồm nhiều khối liên tiếp gồm các hệ số biến đổi, khi bộ mã hóa tiếng nói dựa trên biến đổi được hoạt động trong chế độ bước ngắn; và

đơn vị tạo khung, đơn vị ước lượng đường bao và đơn vị nội suy đường bao được tạo cấu hình để xử lý tập hợp các khối bao gồm một khối hệ số biến đổi, khi bộ mã hóa tiếng nói dựa trên biến đổi được hoạt động trong chế độ bước dài.

10. Bộ mã hóa tiếng nói dựa trên biến đổi theo điểm 9, trong đó, trong chế độ bước dài,

đơn vị ước lượng đường bao được tạo cấu hình để xác định đường bao hiện thời của một khối hệ số biến đổi được bao gồm trong tập hợp các khối; và

đơn vị nội suy đường bao được tạo cấu hình để xác định đường bao được nội suy cho một khối hệ số biến đổi làm đường bao hiện thời của một khối hệ số biến đổi.

11. Bộ giải mã tiếng nói dựa trên biến đổi được tạo cấu hình để giải mã dòng bit nhằm cung cấp tín hiệu tiếng nói được tái tạo, bộ giải mã này bao gồm:

đơn vị giải mã đường bao được tạo cấu hình để xác định đường bao hiện thời được lượng tử hóa từ dữ liệu đường bao được bao gồm trong dòng bit; trong đó đường bao hiện thời được lượng tử hóa chỉ báo nhiều trị số năng lượng phổ cho nhiều bin tần

số tương ứng; trong đó dòng bit bao gồm dữ liệu chỉ báo nhiều khối liên tiếp gồm các hệ số biến đổi đã làm phẳng được tái tạo; trong đó khối hệ số biến đổi đã làm phẳng được tái tạo bao gồm nhiều hệ số biến đổi đã làm phẳng được tái tạo cho nhiều bin tần số tương ứng;

đơn vị nội suy đường bao được tạo cấu hình để xác định nhiều đường bao được nội suy cho nhiều khối hệ số biến đổi đã làm phẳng được tái tạo lần lượt, dựa trên đường bao hiện thời được lượng tử hóa và dựa trên đường bao trước đó được lượng tử hóa; và

đơn vị làm phẳng ngược được tạo cấu hình để xác định nhiều khối hệ số biến đổi được tái tạo bằng cách cung cấp hình dạng phổ cho nhiều khối tương ứng gồm các hệ số biến đổi đã làm phẳng được tái tạo, nhờ sử dụng nhiều đường bao được nội suy tương ứng, một cách lần lượt; trong đó tín hiệu tiếng nói được tái tạo được xác định dựa trên nhiều khối hệ số biến đổi được tái tạo và tạo ra tín hiệu có thể nghe được.

12. Bộ giải mã tiếng nói dựa trên biến đổi theo điểm 11, trong đó đường bao trước đó đã lượng tử hóa kết hợp với nhiều khối trước đó gồm các hệ số biến đổi được tái tạo, đứng ngay trước nhiều khối hệ số biến đổi được tái tạo.

13. Bộ giải mã tiếng nói dựa trên biến đổi theo điểm 11, trong đó:

nhiều khối liên tiếp gồm các hệ số biến đổi đã làm phẳng được tái tạo bao gồm khối thứ nhất gồm các hệ số biến đổi đã làm phẳng được tái tạo ở thời điểm trung gian thứ nhất;

đơn vị nội suy đường bao được tạo cấu hình để xác định trị số năng lượng phổ cho bin tần số cụ thể của đường bao được nội suy thứ nhất bằng cách nội suy các trị số năng lượng phổ cho bin tần số cụ thể của đường bao hiện thời được lượng tử hóa và của đường bao trước đó được lượng tử hóa tại thời điểm trung gian thứ nhất;

đường bao được nội suy thứ nhất được kết hợp với khối thứ nhất gồm các hệ số biến đổi đã làm phẳng được tái tạo.

14. Bộ giải mã tiếng nói dựa trên biến đổi theo điểm 13, trong đó đơn vị nội suy đường bao được tạo cấu hình để xác định trị số năng lượng phổ cho bin tần số cụ thể của đường bao được nội suy thứ nhất bằng cách lượng tử hóa phép nội suy giữa các trị

số năng lượng phổ cho bin tần số cụ thể của đường bao hiện thời được lượng tử hóa và của đường bao trước đó được lượng tử hóa.

15. Bộ giải mã tiếng nói dựa trên biến đổi theo điểm 13, trong đó:

nhiều khối liên tiếp gồm các hệ số biến đổi đã làm phẳng được tái tạo bao gồm khối thứ hai gồm các hệ số biến đổi đã làm phẳng được tái tạo ở thời điểm trung gian thứ hai;

đơn vị nội suy đường bao được tạo cấu hình để xác định trị số năng lượng phổ cho bin tần số cụ thể của đường bao được nội suy thứ hai bằng cách nội suy các trị số năng lượng phổ cho bin tần số cụ thể của đường bao hiện thời được lượng tử hóa và của đường bao trước đó được lượng tử hóa tại thời điểm trung gian thứ hai;

đường bao được nội suy thứ hai được kết hợp với khối thứ hai gồm các hệ số biến đổi đã làm phẳng được tái tạo;

khối thứ hai gồm các hệ số biến đổi đã làm phẳng được tái tạo là theo sau khối thứ nhất gồm các hệ số biến đổi đã làm phẳng được tái tạo; và

thời điểm trung gian thứ hai là theo sau thời điểm trung gian thứ nhất, trong đó hiệu số giữa thời điểm trung gian thứ hai và thời điểm trung gian thứ nhất tương ứng với khoảng thời gian giữa khối thứ hai gồm các hệ số biến đổi đã làm phẳng được tái tạo và khối thứ nhất gồm các hệ số biến đổi đã làm phẳng được tái tạo.

16. Bộ giải mã tiếng nói dựa trên biến đổi theo điểm 11, trong đó:

dòng bit chỉ báo nhiều độ lợi đường bao cho nhiều khối hệ số biến đổi đã làm phẳng được tái tạo, một cách lần lượt;

bộ giải mã tiếng nói dựa trên biến đổi còn bao gồm đơn vị làm mịn đường bao được tạo cấu hình để xác định nhiều đường bao được điều chỉnh bằng cách áp dụng các độ lợi đường bao cho các đường bao được nội suy, một cách lần lượt;

đơn vị làm phẳng ngược được tạo cấu hình để xác định nhiều khối hệ số biến đổi được tái tạo bằng cách cung cấp hình dạng phổ cho nhiều khối tương ứng gồm các hệ số biến đổi đã làm phẳng được tái tạo, nhờ sử dụng nhiều đường bao được điều chỉnh tương ứng, một cách lần lượt.

17. Phương pháp mã hóa tín hiệu tiếng nói thành dòng bit, phương pháp này bao gồm các bước:

nhận tập hợp các khối; trong đó tập hợp các khối bao gồm nhiều khối liên tiếp gồm các hệ số biến đổi; trong đó nhiều khối liên tiếp chỉ báo các mẫu của tín hiệu tiếng nói; trong đó khối hệ số biến đổi bao gồm nhiều hệ số biến đổi cho nhiều bin tần số tương ứng;

xác định đường bao hiện thời dựa trên nhiều khối liên tiếp gồm các hệ số biến đổi; trong đó đường bao hiện thời chỉ báo nhiều trị số năng lượng phổ cho nhiều bin tần số tương ứng;

xác định đường bao hiện thời được lượng tử hóa bằng cách lượng tử hóa đường bao hiện thời;

xác định nhiều đường bao được nội suy cho nhiều khối hệ số biến đổi một cách lần lượt, dựa trên đường bao hiện thời được lượng tử hóa và dựa trên đường bao trước đó được lượng tử hóa;

xác định nhiều khối hệ số biến đổi được làm phẳng bằng cách làm phẳng nhiều khối tương ứng gồm các hệ số biến đổi nhờ sử dụng nhiều đường bao được nội suy tương ứng, một cách lần lượt; và

xác định dòng bit dựa trên nhiều khối hệ số biến đổi được làm phẳng và tạo ra tín hiệu có thể nghe được.

18. Phương pháp giải mã dòng bit nhằm cung cấp tín hiệu tiếng nói được tái tạo, phương pháp này bao gồm các bước:

xác định đường bao hiện thời được lượng tử hóa từ dữ liệu đường bao được bao gồm trong dòng bit; trong đó đường bao hiện thời được lượng tử hóa chỉ báo nhiều trị số năng lượng phổ cho nhiều bin tần số tương ứng; trong đó dòng bit bao gồm dữ liệu chỉ báo nhiều khối liên tiếp gồm các hệ số biến đổi đã làm phẳng được tái tạo; trong đó khối hệ số biến đổi đã làm phẳng được tái tạo bao gồm nhiều hệ số biến đổi đã làm phẳng được tái tạo cho nhiều bin tần số tương ứng;

xác định nhiều đường bao được nội suy cho nhiều khối hệ số biến đổi đã làm phẳng được tái tạo, một cách lần lượt, dựa trên đường bao hiện thời được lượng tử hóa và dựa trên đường bao trước đó được lượng tử hóa;

xác định nhiều khối hệ số biến đổi được tái tạo bằng cách cung cấp hình dạng phổ cho nhiều khối tương ứng gồm các hệ số biến đổi đã làm phẳng được tái tạo, nhờ sử dụng nhiều đường bao được nội suy tương ứng, một cách lần lượt; và

xác định tín hiệu tiếng nói được tái tạo dựa trên nhiều khối hệ số biến đổi được tái tạo và tạo ra tín hiệu có thể nghe được.

19. Phương pháp mã hóa tín hiệu âm thanh bao gồm đoạn tiếng nói thành dòng bit, trong đó phương pháp này bao gồm các bước:

nhận biết đoạn tiếng nói từ tín hiệu âm thanh;

xác định nhiều khối liên tiếp gồm các hệ số biến đổi dựa trên đoạn tiếng nói, nhờ sử dụng đơn vị biến đổi; trong đó khối hệ số biến đổi bao gồm nhiều hệ số biến đổi cho nhiều bin tần số tương ứng; trong đó đơn vị biến đổi được tạo cấu hình để xác định các khối dài bao gồm số lượng thứ nhất các hệ số biến đổi và các khối ngắn bao gồm số lượng thứ hai các hệ số biến đổi; trong đó số lượng thứ nhất lớn hơn số lượng thứ hai; trong đó các khối trong số nhiều khối liên tiếp là các khối ngắn; và

mã hóa nhiều khối liên tiếp vào dòng bit theo điểm 17.

20. Phương pháp giải mã dòng bit chỉ báo tín hiệu âm thanh bao gồm đoạn tiếng nói; phương pháp này bao gồm các bước:

xác định nhiều khối liên tiếp gồm các hệ số biến đổi được tái tạo dựa trên dữ liệu được bao gồm trong dòng bit theo điểm 18; và

xác định đoạn tiếng nói được tái tạo dựa trên nhiều khối liên tiếp gồm các hệ số biến đổi được tái tạo, nhờ sử dụng đơn vị biến đổi ngược; trong đó khối hệ số biến đổi được tái tạo bao gồm nhiều hệ số biến đổi được tái tạo cho nhiều bin tần số tương ứng; trong đó đơn vị biến đổi ngược được tạo cấu hình để xử lý các khối dài bao gồm số lượng thứ nhất các hệ số biến đổi được tái tạo và các khối ngắn bao gồm số lượng thứ hai các hệ số biến đổi được tái tạo; trong đó số lượng thứ nhất lớn hơn số lượng thứ hai; trong đó các khối trong số các khối liên tiếp là các khối ngắn và tạo ra tín hiệu có thể nghe được.

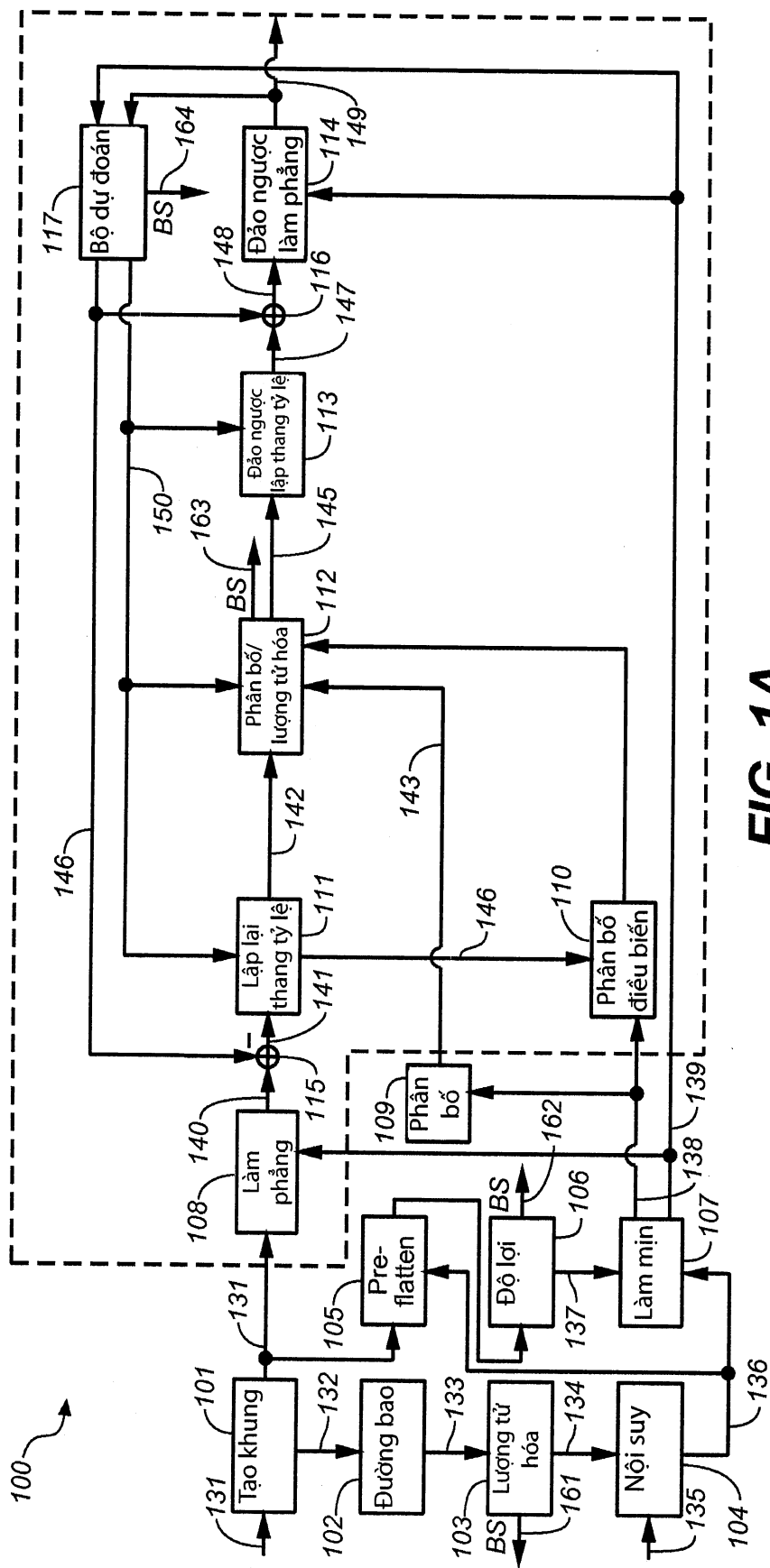
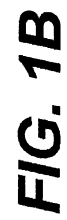
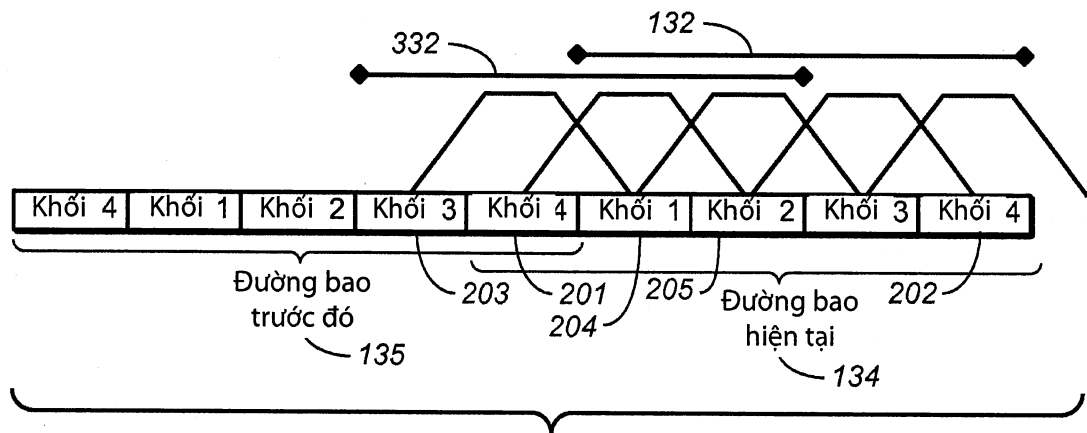
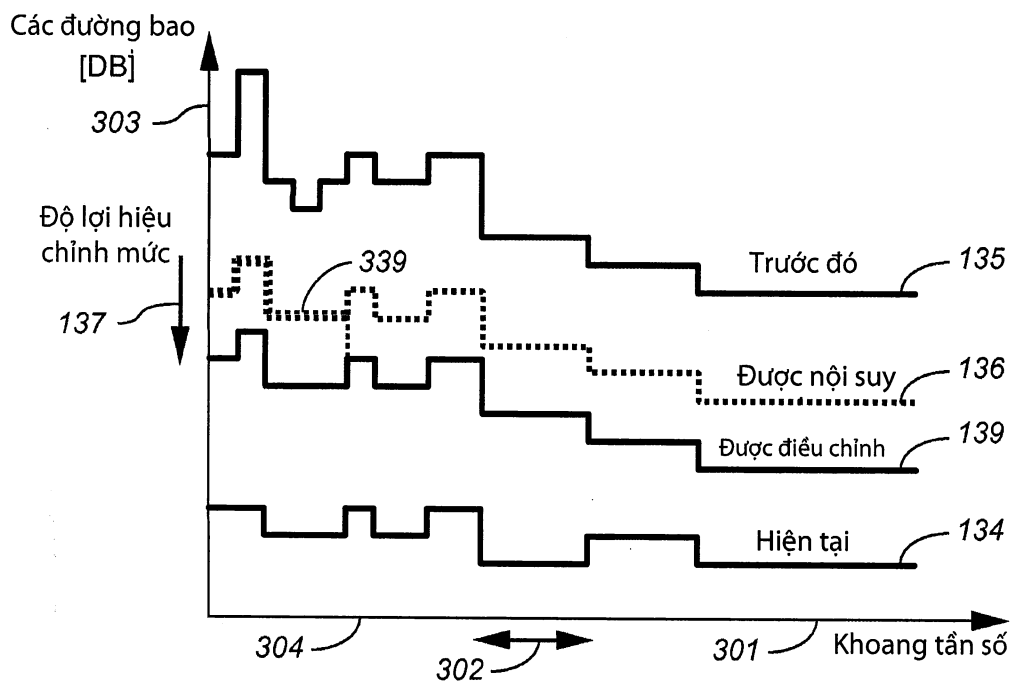
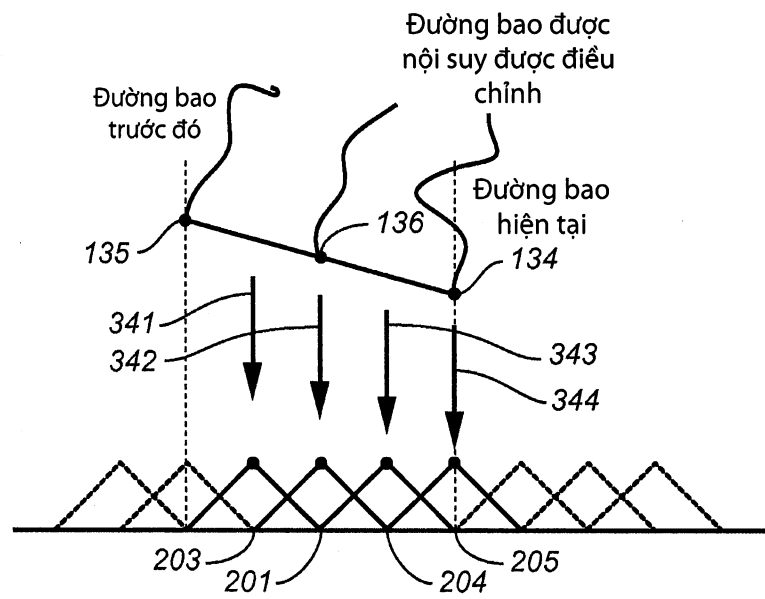
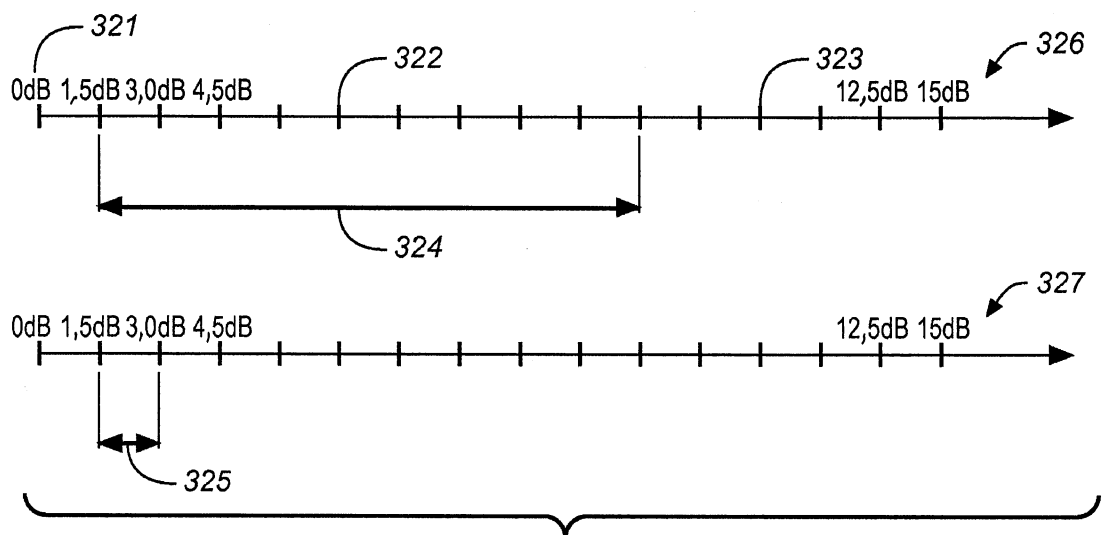


FIG. 1A



**FIG. 2****FIG. 3A**

**FIG. 3B****FIG. 4**

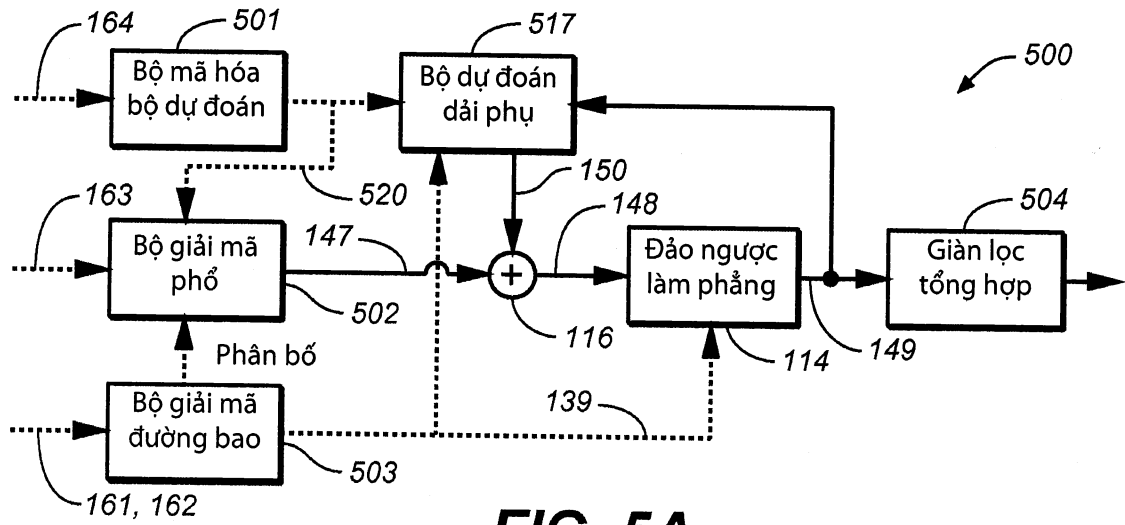


FIG. 5A

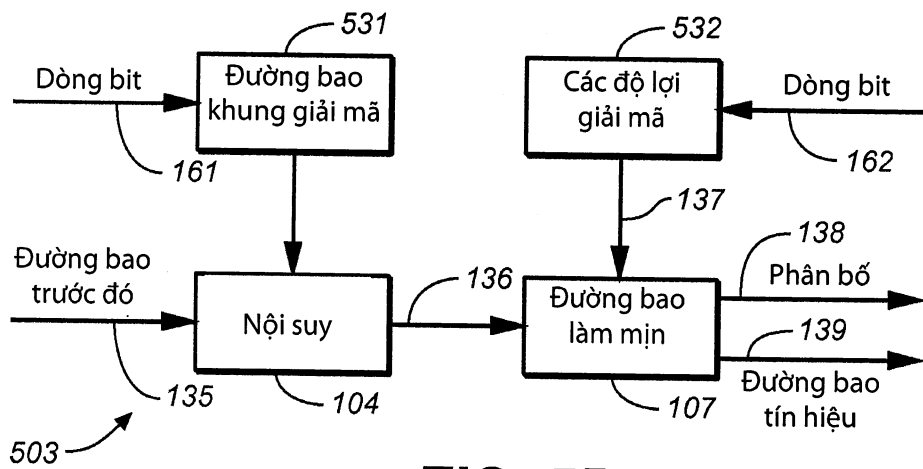
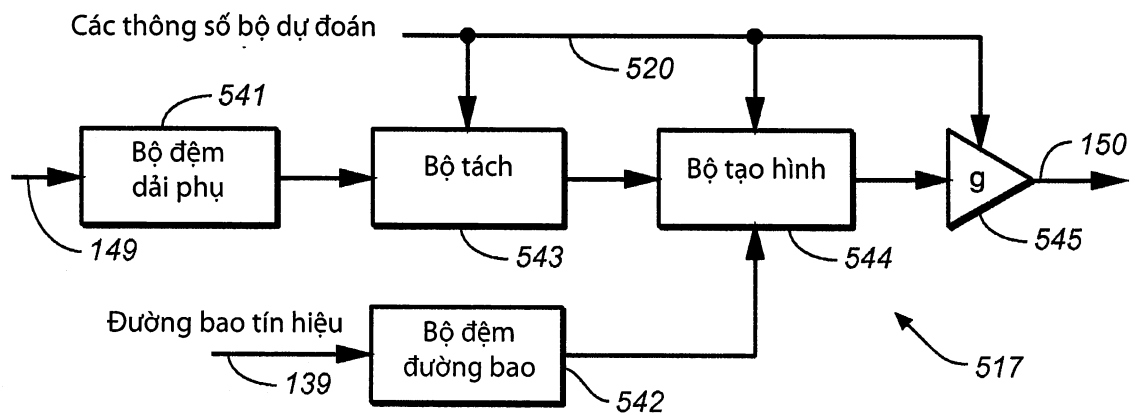
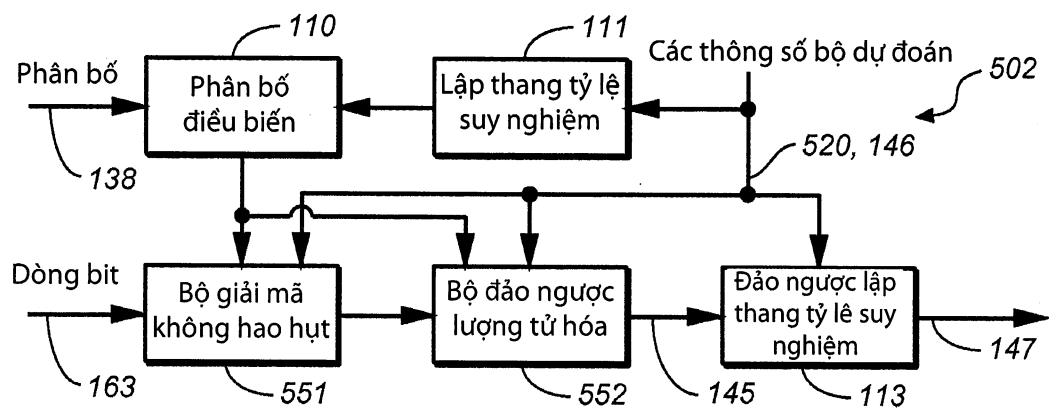


FIG. 5B

**FIG. 5C****FIG. 5D**