



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

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

(11)



1-0030175

(51)⁷ G10L 19/028; G10L 21/038

(13) B

(21) 1-2017-00545

(22) 24/07/2015

(86) PCT/EP2015/067062 24/07/2015

(87) WO 2016/016146 04/02/2016

(30) 14178777.0 28/07/2014 EP

(45) 25/11/2021 404

(43) 26/06/2017 351A

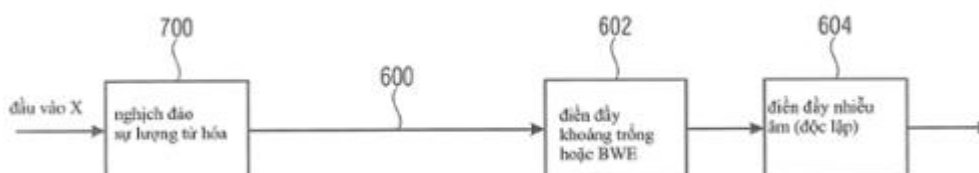
(73) Fraunhofer-Gesellschaft zur Foerderung der angewandten Forschung e.V. (DE)
Hansastrasse 27c, 80686 Muenchen, Germany

(72) DISCH, Sascha (DE); GEIGER, Ralf (DE); NIEDERMEIER, Andreas (DE);
NEUSINGER, Matthias (DE); SCHMIDT, Konstantin (DE); WILDE, Stephan (DE);
SCHUBERT, Benjamin (DE); NEUKAM, Christian (DE).

(74) CÔNG TY LUẬT TRÁCH NHIỆM HỮU HẠN AMBYS HÀ NỘI (AMBYS
HANOI)

(54) THIẾT BỊ VÀ PHƯƠNG PHÁP TẠO RA TÍN HIỆU ÂM THANH NÂNG CAO TỪ
TÍN HIỆU ĐẦU VÀO, HỆ THỐNG VÀ PHƯƠNG PHÁP XỬ LÝ TÍN HIỆU ÂM
THANH

(57) Sáng chế đề cập đến thiết bị và phương pháp tạo ra tín hiệu âm thanh nâng cao từ tín hiệu đầu vào, hệ thống và phương pháp xử lý tín hiệu âm thanh. Thiết bị tạo ra tín hiệu nâng cao từ tín hiệu đầu vào (600), trong đó tín hiệu nâng cao có các giá trị phổ cho vùng phổ nâng cao, các giá trị phổ cho các vùng phổ nâng cao không nằm trong tín hiệu đầu vào (600), bao gồm bộ ánh xạ (602) để ánh xạ vùng phổ nguồn của tín hiệu đầu vào đến vùng đích trong vùng phổ nâng cao, vùng phổ nguồn bao gồm vùng điện đầy nhiều âm (302); và bộ điện đầy nhiều âm (604) được tạo cấu hình để tạo ra các giá trị nhiều âm thứ nhất cho vùng điện đầy nhiều âm (302) trong vùng phổ nguồn của tín hiệu đầu vào và để tạo ra các giá trị nhiều âm thứ hai cho vùng nhiều âm trong vùng đích, trong đó các giá trị nhiều âm thứ hai được giải tương quan từ các giá trị nhiều âm thứ nhất hoặc để tạo ra các giá trị nhiều âm thứ hai cho vùng nhiều âm trong vùng đích, trong đó các giá trị nhiều âm thứ hai được giải tương quan từ các giá trị nhiều âm thứ nhất trong vùng nguồn.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến việc xử lý tín hiệu, và cụ thể, đến việc xử lý tín hiệu âm thanh.

Tình trạng kỹ thuật của sáng chế

Việc mã hóa theo cảm quan các tín hiệu âm thanh cho mục đích giảm dữ liệu để lưu trữ hoặc truyền dẫn hiệu quả các tín hiệu này là thực tiễn được sử dụng rộng rãi. Cụ thể, khi tốc độ bit thấp nhất đạt được, việc mã hóa được sử dụng dẫn đến sự giảm chất lượng âm thanh mà thường được gây ra bởi sự giới hạn tại phía bộ mã hóa của băng thông tín hiệu âm thanh sẽ được truyền. Trong các bộ mã hóa-giải mã hiện hành, các phương pháp nổi tiếng tồn tại cho việc khôi phục tín hiệu phía bộ giải mã thông qua sự mở rộng độ rộng băng tín hiệu âm thanh (Band Width Extension - BWE), ví dụ, sự sao chép băng phổ (Spectral Band Replication - SBR).

Trong việc mã hóa tốc độ bit thấp, việc thường cũng được gọi là điền đầy nhiễu âm được sử dụng. Các vùng phổ nổi bật mà đã được lượng tử hóa về 0 do sự hạn chế tốc độ bit nghiêm ngặt được điền đầy với nhiễu âm tổng hợp trong bộ giải mã.

Thường thường, cả hai kỹ thuật được tổ hợp trong sự áp dụng mã hóa tốc độ bit thấp. Hơn nữa, các giải pháp tích hợp như điền đầy khoảng trống thông minh (Intelligent Gap Filling - IGF) tồn tại mà tổ hợp việc mã hóa âm thanh, điền đầy nhiễu âm và điền đầy khoảng trống phổ.

Tuy nhiên, tất cả các phương pháp này có điểm chung là bước thứ nhất, băng cơ bản hoặc tín hiệu âm thanh lỗi được khôi phục sử dụng việc giải mã dạng sóng và điền đầy nhiễu âm, và trong bước thứ hai, sự xử lý BWE hoặc IGF được thực hiện sử dụng tín hiệu sẵn sàng được khôi phục. Điều này dẫn đến thực tế rằng các giá trị nhiễu âm giống nhau mà đã được điền đầy trong băng cơ bản bằng sự điền đầy nhiễu âm trong đó sự khôi phục được sử dụng để tái tạo các phần bị mất trong băng cao (trong BWE) hoặc để điền đầy các khoảng trống phổ còn lại (trong IGF). Việc sử dụng nhiễu âm

tương quan cao để khôi phục nhiều vùng phổ trong BWE hoặc IGF có thể dẫn đến sự suy yếu cảm quan.

Các chủ đề tương tự trong tình trạng kỹ thuật bao gồm

- SBR như bộ xử lý sau để giải mã dạng sóng [1-3]
- AAC PNS [4]
- sự điền đầy nhiễu âm USAC MPEG-D [5]
- G.719 và G.722.1C [6]
- MPEG-H 3D IGF [8]

Các tín hiệu âm thanh được xử lý với các phương pháp này phải chịu các thành phần lạ như hỗn âm, sự méo biến điệu và âm sắc được nhận biết là khó chịu, cụ thể, tại tốc độ bit thấp và do đó băng thông thấp và/hoặc sự xuất hiện của các lỗ phổ trong phạm vi LF. Lý do cho điều này là, như được giải thích ở trên, trước hết thực tế là các thành phần được khôi phục của phổ được mở rộng hoặc được điền đầy khoảng trống được dựa trên một hoặc nhiều bản sao trực tiếp chứa nhiễu âm từ băng cơ bản. Sự biến điệu theo thời gian là kết quả của sự tương quan không mong muốn đã nêu trong nhiễu âm được khôi phục là có thể nghe được theo cách nhiễu loạn như hỗn âm cảm quan hoặc độ méo khó chịu. Tất cả các phương pháp hiện có như mp3+SBR, AAC+SBR, USAC, G.719 và G.722.1C, và tương tự MPEG-H 3D IGF đầu tiên thực hiện sự giải mã lỗi hoàn chỉnh gồm có điền đầy nhiễu âm trước khi điền đầy các khoảng trống phổ hoặc băng cao với dữ liệu phổ được sao chép hoặc được phản chiếu từ lỗi.

Tài liệu tham khảo

Các tài liệu và các đơn sáng chế sau mô tả các phương pháp mà được xem là liên quan đến đơn sáng chế này:

[1] M. Dietz, L. Liljeryd, K. Kjörling and O. Kunz, “Spectral Band Replication, a novel approach in audio coding,” in 112th AES Convention, Munich, Germany, 2002.

[2] S. Meltzer, R. Böhm and F. Henn, “SBR enhanced audio codecs for digital broadcasting such as “Digital Radio Mondiale” (DRM),” in 112th AES

Convention, Munich, Germany, 2002.

[3] T. Ziegler, A. Ehret, P. Ekstrand and M. Lutzky, “Enhancing mp3 with SBR: Features and Capabilities of the new mp3PRO Algorithm,” in 112th AES Convention, Munich, Germany, 2002.

[4] J. Herre, D. Schulz, Extending the MPEG-4 AAC Codec by Perceptual Noise Substitution, Audio Engineering Society 104th Convention, Preprint 4720, Amsterdam, Netherlands, 1998

[5] European Patent application EP2304720 USAC noise-filling

[6] ITU-T Recommendations G.719 and G.221C

[7] EP 2704142

[8] EP 13177350

Bản chất kỹ thuật của sáng chế

Mục đích của sáng chế là đề xuất khái niệm được cải thiện để tạo ra tín hiệu nâng cao.

Mục tiêu này đạt được bằng thiết bị để tạo ra tín hiệu âm thanh nâng cao theo điểm 1, phương pháp tạo ra tín hiệu âm thanh nâng cao theo điểm 10, hệ thống xử lý tín hiệu âm thanh theo điểm 11, phương pháp xử lý tín hiệu âm thanh theo điểm 12, hoặc vật ghi có thể đọc được bằng máy tính theo điểm 13.

Sáng chế được dựa trên việc tìm ra rằng sự cải thiện đáng kể chất lượng âm thanh của tín hiệu nâng cao được tạo ra bởi sự mở rộng băng thông hoặc điền đầy khoảng trống thông minh hoặc theo bất kỳ cách nào khác tạo ra tín hiệu nâng cao có các giá trị phổ cho vùng phổ nâng cao không được chứa trong tín hiệu đầu vào thu được bằng cách tạo ra các giá trị nhiễu âm thứ nhất cho vùng điền đầy nhiễu âm trong vùng phổ nguồn của tín hiệu đầu vào và bằng cách sau đó tạo ra các giá trị nhiễu âm độc lập thứ hai cho vùng nhiễu âm trong vùng đến hoặc vùng đích, tức là, trong vùng nâng cao mà bây giờ có các giá trị nhiễu âm, tức là, các giá trị nhiễu âm thứ hai mà độc lập từ các giá trị nhiễu âm thứ nhất.

Do đó, vấn đề tình trạng kỹ thuật có nhiễu âm phụ thuộc trong băng cơ bản và băng nâng cao do sự ánh xạ các giá trị phổ được loại bỏ và các vấn đề liên quan với

các thành phần lạ như hỗn âm, độ méo điều biến và âm sắc được nhận biết là khó chịu cụ thể tại các tốc độ bit thấp được loại bỏ.

Nói cách khác, sự điền đầy nhiễu âm của các giá trị nhiễu âm thứ hai là giải tương quan khỏi các giá trị nhiễu âm thứ nhất, tức là, các giá trị nhiễu âm mà ít nhất độc lập một phần khỏi các giá trị nhiễu âm thứ nhất đảm bảo rằng các thành phần lạ không xuất hiện nữa hoặc ít nhất được giảm so với tình trạng kỹ thuật. Do đó, sự xử lý tình trạng kỹ thuật của các giá trị phổ điền đầy nhiễu âm trong băng cơ bản bởi sự mở rộng băng thông hoặc thuật toán điền đầy khoảng trống thông minh đơn giản không giải tương quan nhiễu âm từ băng cơ bản, nhưng chỉ thay đổi mức độ, ví dụ là như vậy. Tuy nhiên, việc đưa các giá trị nhiễu âm được giải tương quan một mặt trong băng nguồn và mặt khác trong băng đích, ưu tiên là được suy ra từ quy trình nhiễu âm tách biệt cung cấp các kết quả tốt nhất. Tuy nhiên, thậm chí sự giới thiệu các giá trị nhiễu âm không hoàn toàn được giải tương quan hoặc không hoàn toàn độc lập, nhưng là được giải tương quan ít nhất một phần như giá trị giải tương quan là 0,5 hoặc ít hơn khi giá trị giải tương quan là 0 biểu thị sự giải tương quan hoàn toàn, cải thiện vấn đề tương quan đầy đủ của tình trạng kỹ thuật.

Sau đây, các phương án liên quan đến sự tổ hợp của giải mã dạng sóng, sự mở rộng băng thông hoặc điền đầy khoảng trống và điền đầy nhiễu âm trong bộ giải mã cảm quan.

Các lợi ích khác là, ngược lại với các khái niệm hiện có, sự xuất hiện các sự méo tín hiệu và các thành phần lạ hỗn âm cảm quan, mà thường là dạng đặc trưng để tính toán sự mở rộng băng thông hoặc sự điền đầy khoảng trống tiếp theo sự giải mã dạng sóng và sự điền đầy nhiễu âm được tránh.

Điều này là do, trong một số phương án, sự thay đổi thứ tự của các bước xử lý được đề cập. Tốt hơn là thực hiện sự mở rộng băng thông hoặc điền đầy nhiễu âm trực tiếp sau sự giải mã dạng sóng và tốt hơn nữa là tính toán sự điền đầy nhiễu âm tiếp theo trên tín hiệu đã được khôi phục sử dụng nhiễu âm giải tương quan.

Trong các phương án khác, sự giải mã dạng sóng và điền đầy nhiễu âm có thể được thực hiện theo thứ tự truyền thống và còn ngược dòng trong sự xử lý, các giá trị nhiễu âm có thể được thay thế bằng nhiễu âm giải tương quan được định tỉ lệ phù hợp.

Do đó, sáng chế nêu ra các vấn đề mà xảy ra do thuật toán sao chép hoặc thuật toán phản chiếu trên phổ được điền đầy nhiễu âm bằng cách dịch chuyển bước điền đầy nhiễu âm đến đầu cuối cùng của chuỗi xử lý và sử dụng nhiễu âm không tương quan cho việc ráp nối hoặc điền đầy khoảng trống.

Mô tả vắn tắt các hình vẽ

Theo đó, các phương án được ưu tiên của sáng chế được thảo luận với các hình vẽ đi kèm, trong đó:

Fig.1a minh họa thiết bị mã hóa tín hiệu âm thanh;

Fig.1b minh họa bộ giải mã để giải mã tín hiệu âm thanh đã mã hóa phù hợp với bộ mã hóa của Fig.1a;

Fig.2a minh họa việc thực hiện được ưu tiên của bộ giải mã;

Fig.2b minh họa việc thực hiện được ưu tiên của bộ mã hóa;

Fig.3a minh họa sự biểu diễn dưới dạng giản đồ của phổ như được tạo ra bởi bộ giải mã miền phổ của Fig.1b;

Fig.3b minh họa bảng biểu thị mối tương quan giữa các hệ số tỉ lệ cho các băng hệ số tỉ lệ và năng lượng cho các băng khôi phục và thông tin điền đầy nhiễu âm cho băng điền đầy nhiễu âm;

Fig.4a minh họa chức năng của bộ mã hóa miền phổ để áp dụng sự lựa chọn các phần phổ vào trong các tập hợp thứ nhất và tập hợp thứ hai gồm các phần phổ;

Fig.4b minh họa sự thực hiện các chức năng của Fig.4a;

Fig.5a minh họa chức năng của bộ mã hóa MDCT;

Fig.5b minh họa chức năng của bộ giải mã với công nghệ MDCT;

Fig.5c minh họa sự thực hiện của bộ tái tạo tần số;

Fig.6 minh họa sơ đồ khối của thiết bị để tạo ra tín hiệu nâng cao phù hợp với sáng chế;

Fig.7 minh họa dòng tín hiệu của sự điền đầy nhiễu âm độc lập được chỉnh hướng bởi thông tin lựa chọn trong bộ giải mã phù hợp với phương án của sáng chế;

Fig.8 minh họa dòng tín hiệu của việc điền đầy nhiễu âm độc lập được thực hiện qua thứ tự trao đổi của sự điền đầy nhiễu âm hoặc sự mở rộng băng thông và sự điền đầy nhiễu âm trong bộ giải mã;

Fig.9 minh họa lưu đồ của quy trình phù hợp với phương án khác của sáng chế;

Fig.10 minh họa lưu đồ của quy trình phù hợp với phương án khác của sáng chế;

Fig.11 minh họa lưu đồ để giải thích sự định tỉ lệ của các giá trị ngẫu nhiên;

Fig.12 minh họa lưu đồ minh họa sự nhúng của sáng chế vào sự mở rộng băng thông thông thường hoặc quy trình điền đầy khoảng trống;

Fig.13a minh họa bộ mã hóa với sự tính toán tham số mở rộng băng thông; và

Fig.13b minh họa bộ giải mã với sự mở rộng băng thông được thực hiện như bộ xử lý sau hơn là quy trình tích hợp như trong Fig.1a hoặc Fig.1b.

Mô tả chi tiết sáng chế

Fig.6 minh họa thiết bị để tạo ra tín hiệu được cải thiện như tín hiệu âm thanh từ tín hiệu đầu vào mà cũng có thể là tín hiệu âm thanh. Tín hiệu nâng cao có các giá trị phổ cho vùng phổ nâng cao, trong đó các giá trị phổ cho vùng phổ nâng cao không được chứa trong tín hiệu đầu vào ban đầu tại đầu vào tín hiệu đầu vào 600. Thiết bị bao gồm bộ ánh xạ 602 để ánh xạ vùng phổ nguồn của tín hiệu đầu vào đến vùng đích trong vùng phổ nâng cao, trong đó vùng phổ nguồn bao gồm vùng điền đầy nhiễu âm.

Hơn nữa, thiết bị bao gồm bộ điền đầy nhiễu âm 604 được tạo cấu hình để tạo ra các giá trị nhiễu âm thứ nhất cho vùng điền đầy nhiễu âm trong vùng phổ nguồn của tín hiệu đầu vào và để tạo ra các giá trị nhiễu âm thứ hai cho vùng nhiễu âm trong vùng đích, trong đó các giá trị nhiễu âm thứ hai, tức là, các giá trị nhiễu âm trong vùng đích là độc lập hoặc không tương quan hoặc được giải tương quan khỏi các giá trị nhiễu âm thứ nhất trong vùng điền đầy nhiễu âm.

Một phương án liên quan đến tình trạng, trong đó sự điền đầy nhiễu âm thực tế được thực hiện trong băng cơ bản, tức là, trong đó các giá trị nhiễu âm trong vùng nguồn đã được tạo ra bằng cách điền đầy nhiễu âm. Trong sự thay thế khác, giả thiết rằng sự điền đầy nhiễu âm trong vùng nguồn chưa được thực hiện. Tuy nhiên, vùng

nguồn có vùng nhiều âm thực sự được điền đầy với các giá trị phổ như nhiều âm được mã hóa một cách minh họa như các giá trị phổ bằng bộ mã hóa nguồn hoặc lỗi. Sự ánh xạ vùng nguồn như nhiều âm này đến vùng nâng cao cũng sẽ tạo ra nhiều âm phụ thuộc trong các vùng nguồn và đích. Để chỉ ra vấn đề này, bộ điền đầy nhiều âm chỉ điền đầy nhiều âm vào trong vùng đích của bộ ánh xạ, tức là, tạo ra các giá trị nhiều âm thứ hai cho vùng nhiều âm trong vùng đích, trong đó các giá trị nhiều âm thứ hai được giải tương quan từ các giá trị nhiều âm thứ nhất trong vùng nguồn. Sự thay thế hoặc điền đầy nhiều âm này có thể cũng diễn ra hoặc trong vùng đệm ô nguồn hoặc có thể diễn ra trong bản thân đích. Vùng nhiều âm có thể được xác định bằng bộ phân loại hoặc bằng cách phân tích vùng nguồn hoặc bằng cách phân tích vùng đích.

Để đạt được mục đích này, sự viện dẫn được thực hiện đến Fig.3A. Fig.3A minh họa vùng điền đầy như bằng thừa số định tỉ lệ 301 trong tín hiệu đầu vào, và bộ điền đầy nhiều âm tạo ra các giá trị phổ nhiều âm thứ nhất trong băng điền đầy nhiều âm 301 này trong thuật toán giải mã của tín hiệu đầu vào.

Hơn nữa,, băng điền đầy nhiều âm 301 này được ánh xạ đến vùng đích, tức là, phù hợp với tình trạng kỹ thuật, các giá trị nhiều âm được tạo ra được ánh xạ đến vùng đích và, do đó, vùng đích sẽ có nhiều âm phụ thuộc hoặc tương quan với vùng nguồn.

Tuy nhiên, phù hợp với sáng chế, bộ điền đầy nhiều âm 604 của Fig.6 tạo ra các giá trị nhiều âm thứ hai cho vùng nhiều âm trong vùng đến hoặc vùng đích, trong đó các giá trị nhiều âm thứ hai được giải tương quan hoặc không tương quan hoặc độc lập với các giá trị nhiều âm thứ nhất trong băng điền đầy nhiều âm 301 của Fig.3A.

Thông thường, sự điền đầy nhiều âm và bộ ánh xạ để ánh xạ vùng phổ nguồn lên vùng đến có thể nằm trong bộ tái tạo tần số cao như được minh họa trong ngữ cảnh của các hình từ Figs.1A đến Fig.5C minh họa trong sự điền đầy khoảng trống được tích hợp hoặc có thể được thực hiện như bộ xử lý trước như được minh họa trong Fig.13B và bộ mã hóa tương ứng trong Fig.13A.

Thông thường, tín hiệu đầu vào phải chịu sự lượng tử hóa 700 hoặc bất kỳ sự xử lý bộ giải mã được định trước nào khác hoặc bổ sung 700 có nghĩa là tại đầu ra của khối 700, tín hiệu đầu vào của Fig.6 thu được, để được nhập vào khối điền đầy nhiều âm, bộ giải mã lỗi hoặc khối bộ điền đầy nhiều âm 704 là đầu vào 600 của Fig.6. Bộ

ánh xạ trong Fig.6 tương ứng với bộ điền đầy khoảng trống hoặc khối mở rộng băng thông 602 và khối điền đầy nhiễu âm độc lập 702 cũng được bao gồm trong bộ điền đầy nhiễu âm 604 của Fig.6. Do đó, các khối 704 và 702 đều nằm trong khối bộ điền đầy nhiễu âm 604 của Fig.6 và khối 704 tạo ra các giá trị được gọi là nhiễu âm thứ nhất cho vùng nhiễu âm trong vùng điền đầy nhiễu âm và khối 702 tạo ra các giá trị nhiễu âm thứ hai cho vùng nhiễu âm trong vùng đến hoặc vùng đích, mà được suy ra từ vùng điền đầy nhiễu âm trong băng cơ bản bằng sự mở rộng băng thông được thực hiện bởi bộ ánh xạ hoặc sự điền đầy khoảng trống hoặc khối mở rộng băng thông 602. Hơn nữa, như được thảo luận ở trên, thuật toán điền đầy nhiễu âm độc lập được thực hiện bởi khối 702 được điều khiển bởi vectơ điều khiển PHI được minh họa bởi đường điều khiển 706.

Bước: Nhận biết nhiễu âm

Trong bước thứ nhất, tất cả các vạch phổ mà biểu diễn nhiễu âm trong khung âm thanh được truyền được nhận biết. Quy trình nhận biết có thể được điều khiển bằng kiến thức hiện có, được truyền của các vị trí nhiễu âm được sử dụng bởi sự điền đầy nhiễu âm [4][5] hoặc có thể được nhận biết với bộ phân loại bổ sung. Kết quả của sự nhận biết vạch nhiễu âm là vectơ chứa các số 0 và các số 1 trong đó vị trí với 1 biểu thị vạch phổ mà biểu diễn nhiễu âm.

Trong thuật ngữ toán học, quy trình này có thể được mô tả như:

Để $\tilde{X} \in \mathbb{C}^N$ là phổ được truyền và được lượng tử hóa lại sau khi sự điền đầy nhiễu âm [4][5] của tín hiệu được mã hóa biến đổi, được tạo cửa sổ có chiều dài $N \in \mathbb{N}$. Để $m \in \mathbb{N}, 0 < m \leq N$, là vạch dừng của toàn bộ quy trình giải mã.

Bộ phân loại C_0 xác định các vạch phổ trong đó sự điền đầy nhiễu âm [4][5] trong vùng lõi được sử dụng:

$$C_0: \mathbb{C}^N \rightarrow \{0,1\}^m,$$

$$\varphi[i] := C_0(\tilde{X})[i] := \begin{cases} 1 & \text{trên } \tilde{X}[i] \text{ điền đầy nhiễu âm được sử dụng} \\ 0 & \text{nếu không} \end{cases}, 0 \leq i < m \leq N,$$

và kết quả $\varphi \in \{0,1\}^m$ là vectơ có chiều dài m.

Bộ phân loại bổ sung C_1 có thể nhận biết các vạch khác trong $\tilde{X}$ mà biểu diễn nhiễu âm. Bộ phân loại có thể được mô tả như:

$$C_1: \mathbb{C}^N \times \{0,1\}^m \rightarrow \{0,1\}^m,$$

$$\varphi[i] := C_1(\tilde{X}, \varphi)[i] = \begin{cases} 1 & \text{nếu } \varphi[i] = 1 \vee \tilde{X}[i] \text{ được phân loại là nhiễu âm} \\ 0 & \text{nếu không} \end{cases}, 0 \leq i < m \leq N.$$

Sau quy trình nhận biết nhiễu âm, vectơ biểu thị nhiễu âm $\varphi \in \{0,1\}^m$ được xác định là:

$$\varphi[i] = \begin{cases} 1, & \text{vạch quang phổ } \tilde{X}[i] \text{ được nhận biết như vạch nhiễu âm} \\ 0, & \text{vạch quang phổ } \tilde{X}[i] \text{ không được nhận biết là vạch nhiễu âm} \end{cases}, 0 \leq i < m \leq N.$$

Bước: Nhiễu âm độc lập

Trong bước thứ hai, vùng cụ thể của phổ được truyền được lựa chọn và sao chép đến ô nguồn. Trong ô nguồn này, nhiễu âm được nhận biết được thay thế bằng nhiễu âm ngẫu nhiên. Năng lượng của nhiễu âm ngẫu nhiên được chèn vào được điều chỉnh đến năng lượng giống nhau của nhiễu âm ban đầu trong ô nguồn.

Trong thuật ngữ toán học, quy trình này có thể được mô tả như:

Để $n, n < m$, là vạch bắt đầu cho quy trình sao chép, được mô tả trong Bước 3. Để $\tilde{X}_{sT} \subset \tilde{X}$ là phần kè nhau của phổ được truyền $\tilde{X}$, biểu diễn ô nguồn có chiều dài $v < n$, mà chứa các vạch phổ $l_k, l_{k+1}, \dots, l_{k+v-1}$ của $\tilde{X}$, trong đó k là chỉ số của vạch phổ thứ nhất trong ô nguồn $\tilde{X}_{sT}$, để $\tilde{X}_{sT}[i] = l_{k+i}, 0 \leq i < v$. Hơn nữa,, để $\varphi' \subset \varphi$, để $\varphi'[i] = \varphi[k+i], 0 \leq i < v$.

Nhiễu âm được nhận biết bây giờ được thay thế bằng nhiễu âm tổng hợp được tạo ra ngẫu nhiên. Để giữ năng lượng phổ tại mức giống nhau, năng lượng E của nhiễu âm được biểu thị bởi φ ban đầu được tính toán:

$$E = \sum_{i=0}^{v-1} \varphi'[i] |\tilde{X}_{sT}[i]|^2.$$

Nếu $E = 0$ bỏ qua sự thay thế nhiễu âm độc lập cho ô nguồn $\tilde{X}_{sT}$, nếu không thay

thể nhiều âm được biểu thị bởi φ' :

$$\tilde{X}'_{sT}[i] = \begin{cases} r[i], & \varphi'[i] = 1 \\ \tilde{X}_{sT}[i], & \varphi'[i] = 0 \end{cases}, \quad 0 \leq i < v,$$

trong đó $r[i] \in \mathbb{C}$ là số ngẫu nhiên cho tất cả $0 \leq i < v$.

Sau đó tính toán năng lượng E' của các số ngẫu nhiên được chèn:

$$E' = \sum_{i=0}^{v-1} \varphi[i]' |\tilde{X}'_{sT}[i]|^2.$$

Nếu $E' > 0$ tính toán thừa số g , nếu không thiết lập $g = 0$:

$$g = \sqrt{\frac{E}{E'}}$$

Với g , định tỉ lệ lại nhiều âm được thay thế:

$$\tilde{X}''_{sT}[i] = \begin{cases} g\tilde{X}'_{sT}[i], & \varphi'[i] = 1 \\ \tilde{X}'_{sT}[i], & \varphi'[i] = 0 \end{cases}, \quad 0 \leq i < v.$$

Sau sự thay thế nhiều âm, ô nguồn $\tilde{X}''_{sT}[i]$ chứa các vạch nhiều âm mà độc lập với các vạch nhiều âm trong $\tilde{X}$.

Bước: Sao chép

Ô nguồn $\tilde{X}''_{sT}[i]$ được ánh xạ vào vùng đến trong $\tilde{X}$:

$$\tilde{X}[c+i] = \tilde{X}''_{sT}[i], \quad 0 \leq i < v, \quad c \geq n, \quad c+i < m < N,$$

hoặc, nếu sơ đồ IGF [8] được sử dụng:

$$\tilde{X}[c+i] = \begin{cases} \tilde{X}''_{sT}[i], & \tilde{X}[c+i] = 0 \\ \tilde{X}[c+i], & \tilde{X}[c+i] \neq 0 \end{cases}, \quad 0 \leq i < v, \quad c \geq n, \quad c+i < m < N.$$

Fig.8 minh họa phương án, trong đó, tiếp theo bất kỳ sự xử lý sau nào như sự giải mã miền phổ được minh họa trong khối 112 trên Fig.1B hoặc, phương án bộ xử lý sau được minh họa bởi khối 1326 trên Fig.13B, tín hiệu đầu vào phải chịu sự điền đầy khoảng trống hoặc sự mở rộng băng thông đầu tiên, tức là, phải chịu thuật toán ánh xạ đầu tiên và, sau đó, sự điền đầy nhiều âm độc lập được thực hiện sau đó, tức là, trong

phổ đầy đủ.

Quy trình được mô tả trong ngữ cảnh nêu trên của Fig.7 có thể được thực hiện như thuật toán ban đầu, sao cho vùng đệm trung gian $\hat{X}_{st}''$ không cần thiết. Do đó thứ tự của sự thực hiện được thích ứng.

Thực hiện bước thứ nhất như được mô tả trong ngữ cảnh của Fig.7, một lần nữa tập hợp các vạch phổ $k, k+1, \dots, k+v-1$ của $\hat{X}$ là vùng nguồn. Thực hiện;

Bước: Sao chép

$$\hat{X}[c+i] = \hat{X}[k+i], \quad 0 \leq i < v, \quad c \geq n, \quad 0 < k+i < n, \quad c+i < m < N,$$

hoặc, nếu sơ đồ IGF [8] được sử dụng:

$$\hat{X}[c+i] = \begin{cases} \hat{X}[k+i], & \hat{X}[c+i] = 0 \\ \hat{X}[c+i], & \hat{X}[c+i] < > 0 \end{cases}$$

$$0 \leq i < v, c \geq n, \quad 0 < k+i < n, \quad c+i < m < N.$$

Bước: Điền đầy nhiễu âm độc lập

Thực hiện việc điền đầy nhiễu âm để lại đến n và tính toán năng lượng của các vạch phổ nhiễu âm trong vùng nguồn $k, k+1, \dots, k+v-1$:

$$E := \sum_{i=0}^{v-1} \varphi[k+i] |\hat{X}[k+i]|^2.$$

Thực hiện sự điền đầy nhiễu âm độc lập trong sự điền đầy khoảng trống hoặc vùng phổ BWE;

$$\hat{X}[c+i] := \begin{cases} r[i], & \varphi[k+i] = 1 \\ \hat{X}[c+i], & \varphi[k+i] = 0 \end{cases}, \quad 0 \leq i < v,$$

trong đó $r[i], 0 \leq i < v$ lần nữa là tập hợp các số ngẫu nhiên.

Tính toán năng lượng E' của các số ngẫu nhiên được chèn:

$$E' := \sum_{i=0}^{v-1} \varphi[k+i] |\hat{X}[c+i]|^2.$$

Lần nữa, nếu $E' > 0$ tính toán thừa số g , nếu không thiết lập $g = 0$:

$$g := \sqrt{\frac{E}{E'}}$$

Với g , định tỉ lệ lại nhiễu âm được thay thế:

$$\hat{X}[c+i] := \begin{cases} g\hat{X}[c+i], & \varphi[k+i] = 1 \\ \hat{X}[c+i], & \varphi[k+i] = 0 \end{cases}, \quad 0 \leq i < v.$$

Sự điền đầy nhiễu âm độc lập mới cũng có thể được sử dụng trong môi trường cặp kênh âm lập thể. Do đó, bộ mã hóa tính toán sự biểu diễn cặp kênh phù hợp, L/R hoặc M/S, mỗi băng tần và các hệ số dự báo tùy chọn. Bộ mã hóa áp dụng sự điền đầy nhiễu âm độc lập như được mô tả như trên đến sự biểu diễn được chọn một cách phù hợp của các kênh trước sự tính toán tiếp theo của sự chuyển đổi cuối cùng của tất cả các băng tần trong sự biểu diễn L/R.

Sáng chế là có thể áp dụng hoặc thích hợp cho tất cả các sự ứng dụng âm thanh trong đó băng thông đầy đủ không có sẵn hoặc nó sử dụng việc điền đầy khoảng trống để điền đầy các lỗ hổng. Sáng chế có thể tìm kiếm việc sử dụng trong sự phân bố hoặc phát sóng nội dung âm thanh như, ví dụ, với vô tuyến kỹ thuật số, truyền theo dòng liên mạng và các ứng dụng truyền thông âm thanh.

Theo đó, các phương án của sáng chế được thảo luận với các hình vẽ từ Fig.9 đến Fig.12. Trong bước 900, các vùng nhiễu âm được nhận biết trong phạm vi nguồn. Quy trình này, mà đã được thảo luận trước đối với "sự nhận biết nhiễu âm" có thể dựa trên thông tin phụ điền đầy nhiễu âm được nhận đầy đủ từ phía bộ mã hóa hoặc có thể cũng được tạo cấu hình để ngoài ra hoặc cách khác dựa vào sự phân tích tín hiệu của tín hiệu đầu vào đã được tạo ra, nhưng không có các giá trị phổ cho vùng phổ nâng cao, tức là, không có các giá trị phổ cho vùng phổ của sự nâng cao này.

Sau đó, trong bước 902, phạm vi nguồn mà đã phải chịu sự điền đầy nhiễu âm dễ dàng như được biết trong tình trạng kỹ thuật, tức là, phạm vi nguồn hoàn chỉnh được sao chép cho vùng đệm ô nguồn.

Sau đó, trong bước 904, các giá trị nhiễu âm thứ nhất, tức là, các giá trị nhiễu âm dễ dàng được tạo ra trong vùng điền đầy nhiễu âm của tín hiệu đầu vào được thay thế trong vùng đệm ô nguồn bằng các giá trị ngẫu nhiên. Sau đó, trong bước 906, các giá

trị ngẫu nhiên này được định tỉ lệ trong vùng đệm ô nguồn để thu được các giá trị nhiễu âm thứ hai cho vùng đích. Sau đó, trong bước 908, thuật toán ánh xạ được thực hiện, tức là, nội dung của chúng của vùng đệm ô nguồn có sẵn tiếp theo các bước 904 và 906 được ánh xạ vào phạm vi đến. Do đó, bằng thuật toán thay thế 904, và tiếp theo thuật toán ánh xạ 908, thuật toán điền đầy nhiễu âm độc lập trong phạm vi nguồn và trong phạm vi đích đã thu được.

Fig.10 minh họa phương án khác của sáng chế. Lần nữa, trong bước 900, nhiễu âm trong phạm vi nguồn được nhận biết. Tuy nhiên, chức năng của bước 900 này khác với chức năng của bước 900 trên Fig.9, vì bước 900 trên Fig.9 có thể thực hiện trên phổ tín hiệu đầu vào mà đã được nhận các giá trị nhiễu âm, tức là, trong đó thuật toán điền đầy nhiễu âm đã được thực hiện.

Tuy nhiên, trên Fig.10, bất kỳ thuật toán điền đầy nhiễu âm nào cho tín hiệu đầu vào vẫn chưa được thực hiện và tín hiệu đầu vào chưa có bất kỳ giá trị nhiễu âm nào trong vùng điền đầy nhiễu âm tại đầu vào trong bước 902. Trong bước 902, phạm vi nguồn được ánh xạ đến phạm vi đến hoặc đích trong đó các giá trị điền đầy nhiễu âm không được nằm trong phạm vi nguồn.

Do đó, sự nhận biết nhiễu âm trong phạm vi nguồn trong bước 900 có thể, đối với vùng điền đầy nhiễu âm, được thực hiện bằng cách nhận biết các giá trị phổ bằng 0 trong tín hiệu và/hoặc bằng cách sử dụng thông tin phụ điền đầy nhiễu âm từ tín hiệu đầu vào, tức là, thông tin điền đầy nhiễu âm được tạo ra phía bộ mã hóa. Sau đó, trong bước 904, thông tin điền đầy nhiễu âm và, cụ thể, thông tin năng lượng nhận biết năng lượng sẽ được dẫn vào tín hiệu đầu vào phía bộ giải mã được đọc.

Khi đó, như được minh họa trong bước 1006, sự điền đầy nhiễu âm trong phạm vi nguồn được thực hiện và, tiếp theo hoặc đồng thời, bước 1008 được thực hiện, ví dụ, các giá trị ngẫu nhiên được chèn trong các vị trí trong phạm vi đến mà đã được nhận biết bằng bước 900 qua băng đầy đủ hoặc thông tin tín hiệu đầu vào cùng với thông tin ánh xạ, tức là, phạm vi nguồn nào (trong số nhiều phạm vi nguồn) được ánh xạ lên phạm vi đích nào (trong số nhiều phạm vi đích).

Cuối cùng, các giá trị ngẫu nhiên được chèn được định tỉ lệ để thu được các giá trị nhiễu âm độc lập hoặc không tương quan hoặc được giải tương quan thứ hai.

Tiếp theo, Fig.11 được thảo luận để minh họa thông tin về việc định tỉ lệ các giá trị điền đầy nhiễu âm trong vùng phổ nâng cao, tức là, bằng cách nào, từ các giá trị ngẫu nhiên, các giá trị ngẫu nhiên thứ hai thu được.

Trong bước 1100, thông tin năng lượng về nhiễu âm trong phạm vi nguồn thu được. Sau đó, thông tin năng lượng được xác định từ các giá trị ngẫu nhiên, tức là, từ các giá trị được tạo ra bởi quy trình ngẫu nhiên hoặc giả ngẫu nhiên như được minh họa trong bước 1102. Hơn nữa, bước 1104 minh họa cách làm thế nào để tính toán thừa số định tỉ lệ, tức là, bằng cách sử dụng thông tin năng lượng về nhiễu âm trong phạm vi nguồn và bằng cách sử dụng thông tin năng lượng về các giá trị ngẫu nhiên. Sau đó, trong bước 1106, các giá trị ngẫu nhiên, tức là, từ đó năng lượng đã được tính toán trong bước 1102, được nhân với thừa số định tỉ lệ được tạo ra bằng bước 1104. Do đó, quy trình được minh họa trên Fig.11 tương ứng với sự tính toán thừa số định tỉ lệ g được minh họa trước trong phương án. Tuy nhiên, tất cả các tính toán này cũng được thực hiện trong miền logarit hoặc trong miền bất kì khác và bước nhân 1106 có thể được thay thế bằng phép cộng hoặc phép trừ trong phạm vi logarit.

Sự tham chiếu nữa được thực hiện đến Fig.12 để minh họa sự nhúng của sáng chế này trong sự điền đầy khoảng trống thông minh chung hoặc sơ đồ mở rộng băng thông. Trong bước 1200, thông tin đường bao phổ được lấy từ tín hiệu đầu vào. Thông tin đường bao phổ có thể, ví dụ, được tạo ra bằng bộ trích theo tham số 1306 của Fig.13A và có thể được cung cấp bởi bộ giải mã tham số 1324 của Fig.13b. Sau đó, các giá trị thứ hai và các giá trị khác trong phạm vi đến được định tỉ lệ sử dụng thông tin đường bao phổ này như được minh họa trong 1202. Tiếp theo, bất kỳ việc xử lý sau 1204 nào có thể được thực hiện để thu được tín hiệu nâng cao miền thời gian có băng thông tăng trong trường hợp mở rộng băng thông hoặc có số lượng giảm hoặc không có lỗi hồng nào trong ngữ cảnh của việc điền đầy khoảng trống thông minh.

Trong trường hợp này, đã được chỉ ra rằng, cụ thể đối với phương án của Fig.9, một số thay thế có thể được áp dụng. Đối với phương án, bước 902 được thực hiện với tất cả phổ của tín hiệu đầu vào hoặc ít nhất với phần của phổ của tín hiệu đầu vào mà ở trên tần số biên điền đầy nhiễu âm. Tần số này đảm bảo rằng dưới tần số nhất định, tức là, dưới tần số này, bất kỳ sự điền đầy nhiễu âm nào không được thực hiện.

Sau đó, bất kể bất kỳ thông tin ánh xạ phạm vi nguồn/phạm vi đích cụ thể nào, toàn bộ phổ tín hiệu đầu vào, tức là, phạm vi nguồn tiềm năng hoàn chỉnh nào được sao chép đến vùng đệm ô nguồn 902 và sau đó được xử lý với bước 904 và 906 và bước 908 sau đó lựa chọn vùng nguồn được yêu cầu cụ thể nhất định từ vùng đệm ô nguồn này.

Tuy nhiên, trong các phương án khác, chỉ các phạm vi nguồn được yêu cầu cụ thể mà có thể chỉ các phần của tín hiệu đầu vào được sao chép đến vùng đệm ô nguồn đơn hoặc đến các vùng đệm ô nguồn riêng lẻ dựa trên thông tin phạm vi nguồn/phạm vi đích được nằm trong tín hiệu đầu vào, tức là, được liên kết như thông tin phụ đến tín hiệu đầu vào âm thanh. Phụ thuộc vào tình huống, sự thay thế thứ hai, trong đó chỉ các phạm vi nguồn được yêu cầu cụ thể được xử lý bằng các bước 902, 904, 906, sự phức tạp hoặc ít nhất các yêu cầu bộ nhớ có thể được giảm so với tình huống trong đó luôn luôn, độc lập với tình huống ánh xạ cụ thể, toàn bộ phạm vi nguồn ít nhất trên tần số biên điền đầy nhiễu âm được xử lý bởi các bước 902, 904, 906.

Tiếp theo, viện dẫn đến các hình vẽ từ Fig.1a đến Fig.5c để minh họa sự thực hiện cụ thể của sáng chế nằm trong sự tái tạo tần số 116, mà được đặt trước bộ chuyển đổi phổ - thời gian 118.

Fig.1a minh họa thiết bị mã hóa tín hiệu âm thanh 99. Tín hiệu âm thanh 99 được nhập vào bộ chuyển đổi phổ theo thời gian 100 để chuyển đổi tín hiệu âm thanh có tốc độ lấy mẫu thành sự biểu diễn phổ 101 được phát ra bởi bộ biến đổi phổ theo thời gian. Phổ 101 được nhập vào trong bộ phân tích phổ 102 để phân tích sự biểu diễn phổ 101. Bộ phân tích phổ 101 được tạo cấu hình để xác định tập hợp thứ nhất gồm các phần phổ thứ nhất 103 để được mã hóa với độ phân giải phổ thứ nhất và tập hợp thứ hai khác gồm các phần phổ thứ hai 105 sẽ được mã hóa với độ phân giải phổ thứ hai. Độ phân giải phổ thứ hai nhỏ hơn độ phân giải phổ thứ nhất. Tập hợp thứ hai gồm các phần phổ thứ hai 105 được nhập vào bộ tính toán theo tham số hoặc bộ mã hóa theo tham số 104 để tính toán thông tin đường bao phổ có độ phân giải phổ thứ hai. Hơn nữa, bộ mã hóa âm thanh miền phổ 106 được cung cấp để tạo ra sự biểu diễn được mã hóa thứ nhất 107 của tập hợp thứ nhất gồm các phần phổ thứ nhất có độ phân giải phổ thứ nhất. Thêm nữa, bộ tính toán theo tham số/bộ mã hóa theo tham số 104 được tạo

cấu hình để tạo ra sự biểu diễn được mã hóa thứ hai 109 của tập hợp thứ hai gồm các phần phổ thứ hai. Sự biểu diễn được mã hóa thứ nhất 107 và sự biểu diễn được mã hóa thứ hai 109 được nhập vào bộ dồn kênh dòng bit hoặc bộ tạo dòng bit 108 và khối 108 cuối cùng phát ra tín hiệu âm thanh được mã hóa để truyền hoặc lưu trữ trên thiết bị lưu trữ.

Cụ thể, phần phổ thứ nhất như 306 của Fig.3a sẽ được bao quanh bởi hai phần phổ thứ hai là 307a, 307b. Đây không phải là trường hợp trong HE AAC, tại đó khoảng tần số bộ mã hóa lỗi bị giới hạn bằng.

Fig.1b minh họa bộ giải mã phù hợp với bộ mã hóa của Fig.1a. Sự biểu diễn được mã hóa thứ nhất 107 được nhập vào bộ giải mã âm thanh miền phổ 112 để tạo ra sự biểu diễn được giải mã thứ nhất của tập hợp thứ nhất gồm các phần phổ thứ nhất, sự biểu diễn được giải mã có độ phân giải phổ thứ nhất. Hơn nữa, sự biểu diễn được mã hóa thứ nhất 109 được nhập vào bộ giải mã âm thanh miền phổ 114 để tạo ra sự biểu diễn được giải mã thứ hai của tập hợp thứ hai gồm các phần phổ thứ hai có độ phân giải phổ thứ hai thấp hơn độ phân giải phổ thứ nhất.

Bộ giải mã khác bao gồm bộ tái tạo tần số 116 để tái tạo phần phổ thứ hai được khôi phục có độ phân giải phổ thứ nhất sử dụng phần phổ thứ nhất. Bộ tái tạo tần số 116 thực hiện thuật toán điền đầy ô, cụ thể là, sử dụng ô hoặc phần của bộ thứ nhất của các phần phổ thứ nhất và sao chép bộ thứ nhất này của các phần phổ thứ nhất thành khoảng khôi phục hoặc băng khôi phục có phần phổ thứ hai và cụ thể thực hiện việc tạo hình đường bao phổ hoặc thuật toán khác như được biểu thị bởi sự biểu diễn thứ hai được phát ra bởi bộ giải mã theo tham số 114, cụ thể, bằng cách sử dụng thông tin trên tập hợp thứ hai gồm các phần phổ thứ hai. Tập hợp thứ nhất được giải mã gồm các phần phổ thứ nhất và tập hợp thứ hai được khôi phục gồm các phần phổ như được biểu thị tại đầu ra của bộ tái tạo tần số 116 trên đường dẫn 117 được nhập vào bộ chuyển đổi phổ - thời gian 118 được tạo cấu hình để chuyển đổi sự biểu diễn được giải mã thứ nhất trong phần phổ thứ hai được khôi phục thành sự biểu diễn thời gian 119, sự biểu diễn thời gian 119 có tốc độ lấy mẫu cao nhất định.

Fig.2b minh họa sự thực hiện của bộ mã hóa trên Fig.1a. Tín hiệu đầu vào âm thanh 99 được nhập vào giàn lọc phân tích 220 tương ứng với bộ chuyển đổi phổ - thời

gian 100 của Fig.1a. Sau đó, thuật toán tạo hình nhiễu âm theo thời gian được thực hiện trong khối TNS 222. Do đó, việc nhập vào bộ phân tích phổ 102 của Fig.1a tương ứng với màn chắn âm khối 226 của Fig.2b có thể hoặc là giá trị phổ đầy đủ, khi thực hiện thuật toán tạo hình nhiễu âm theo thời gian/tạo hình ô theo thời gian không được áp dụng hoặc có thể là các giá trị dư phổ, khi thuật toán TNS như được minh họa trên Fig.2b, khối 222 được áp dụng. Đối với các tín hiệu hai kênh hoặc các tín hiệu đa kênh, sự mã hóa kênh nói 228 có thể được thực hiện bổ sung, để bộ mã hóa miền phổ 106 của Fig.1a có thể bao gồm khối mã hóa kênh nói 228. Hơn nữa, bộ mã hóa entropi 232 để thực hiện việc nén dữ liệu không tổn hao được cung cấp mà cũng là phần của bộ mã hóa miền phổ 106 của Fig.1a.

Bộ phân tích phổ/màn chắn âm 226 phân tách đầu ra của khối TNS 222 thành băng lỗi và các thành phần âm tương ứng với tập hợp thứ nhất gồm các phần phổ thứ nhất 103 và các phần dư tương ứng với tập hợp thứ hai gồm các phần phổ thứ hai 105 của Fig.1a. Khối 224 được biểu thị như việc mã hóa trích theo tham số tương ứng với bộ mã hóa theo tham số 104 của Fig.1a và bộ dồn kênh dòng bit 230 tương ứng với bộ dồn kênh dòng bit 108 của Fig.1a.

Tốt hơn là, giàn lọc phân tích 222 được thực hiện như giàn lọc biến đổi cosin rời rạc được thay đổi (modified discrete cosine transform filterbank - MDCT) và MDCT được sử dụng để biến đổi tín hiệu 99 thành miền thời gian - tần số với phép biến đổi cosin rời rạc được cải biên đóng vai trò như công cụ phân tích tần số.

Bộ phân tích phổ 226 tốt hơn là áp dụng màn chắn âm. Giai đoạn ước lượng màn chắn âm được sử dụng để phân tách các thành phần âm từ thành phần tương tự nhiễu âm trong tín hiệu. Điều này cho phép bộ mã hóa lỗi 228 mã hóa tất cả các thành phần âm với môđun tâm thính học. Giai đoạn ước lượng màn chắn âm có thể được thực hiện bằng nhiều cách khác nhau và tốt hơn là được thực hiện tương tự trong chức năng của nó đến giai đoạn ước lượng rãnh đường hình sin được sử dụng trong hình sin và mô hình hóa nhiễu âm cho việc mã hóa tiếng nói/âm thanh [8, 9] hoặc bộ mã hóa âm thanh trên cơ sở mẫu HILN được mô tả trong [10]. Tốt hơn là, việc thực hiện được sử dụng mà dễ dàng thực hiện không cần duy trì quỹ đạo sinh - tử, nhưng bất kỳ bộ phát hiện âm hoặc nhiễu âm khác có thể cũng được sử dụng tốt.

Môđun IGF tính toán sự giống nhau mà tồn tại giữa vùng nguồn và vùng đích. Vùng đích được biểu diễn bởi phổ từ vùng nguồn. Phép đo sự giống nhau giữa vùng nguồn và vùng đích được thực hiện sử dụng phương pháp tương quan chéo. Vùng đích được chia thành n_{Tar} các ô tần số không chồng lấp. Đối với mỗi ô trong vùng đích, n_{Src} các ô nguồn được tạo ra từ tần số bắt đầu cố định. Các ô nguồn này chồng lấp với thừa số nằm giữa 0 và 1, tại đó 0 nghĩa là chồng lấp 0% và 1 nghĩa là chồng lấp 100%. Từng ô nguồn này được tương quan với ô đích tại các độ trễ khác nhau để tìm ra ô nguồn phù hợp nhất với ô đích. Số ô phù hợp nhất được lưu trữ trong $tileNum[idx_tar]$, độ trễ mà tại đó nó tương quan nhất với đích được lưu trữ trong $xcorr_lag[idx_tar][idx_src]$ và ký hiệu của sự tương quan được lưu trữ trong $xcorr_sign[idx_tar][idx_src]$. Trong trường hợp sự tương quan là âm cao, ô nguồn cần được nhân lên với -1 trước quy trình điền đầy ô tại bộ giải mã. Môđun IGF cũng lưu ý không ghi đè các thành phần âm trong phổ vì các thành phần âm đã bảo toàn sử dụng màn chắn âm. Tham số năng lượng theo băng được sử dụng để lưu trữ năng lượng của vùng đích cho phép chúng ta khôi phục phổ một cách chính xác.

Phương pháp này có ưu điểm nhất định hơn cả SBR cổ điển [1] trong đó lưới sóng hài của tín hiệu đa âm được bảo toàn bởi bộ mã hóa lõi trong khi chỉ các khoảng trống giữa đường hình sin được điền đầy với "nhiều âm được tạo hình" phù hợp nhất từ vùng nguồn. Ưu điểm khác của hệ thống này khi so sánh với sự thay thế phổ chính xác ASR (Accurate Spectral Replacement - ASR) [2-4] là việc thiếu giai đoạn tổng hợp tín hiệu mà tạo các phần quan trọng của tín hiệu tại bộ giải mã. Thay vào đó, nhiệm vụ này được tiếp quản bởi bộ mã hóa lõi, cho phép việc bảo toàn các thành phần quan trọng của phổ. Ưu điểm khác của hệ thống được đề xuất là khả năng mở rộng liên tục mà các tính năng được đề ra.. Chỉ sử dụng $tileNum[idx_tar]$ và $xcorr_lag = 0$, cho mỗi ô được gọi là sự phù hợp độ kết hạt và có thể được sử dụng cho các tốc độ bit thấp trong khi sử dụng nhiều $xcorr_lag$ cho mỗi ô cho phép chúng ta làm phù hợp các phổ đích và nguồn hơn.

Ngoài ra, kỹ thuật làm ổn định chọn ô được đề xuất mà loại bỏ thành phần lạ miền tần số như nhiều âm âm rung và nhiều âm âm nhạc.

Trong trường hợp của cặp kênh âm lập thể việc xử lý âm lập thể nổi bổ sung được áp dụng. Điều này là cần thiết, vì đối với phạm vi đến nhất định tín hiệu có thể là nguồn âm thanh được quét tương quan cao. Trong trường hợp các vùng nguồn được chọn cho vùng cụ thể này không tương quan tốt, mặc dù năng lượng phù hợp cho các vùng đến, hình ảnh không gian có thể tổn thất do các vùng nguồn không tương quan. Bộ mã hóa phân tích từng băng năng lượng vùng đến, cụ thể thực hiện sự tương quan chéo của các giá trị phổ và nếu vượt quá ngưỡng nhất định, thiết lập cờ hiệu nổi cho băng năng lượng này. Trong bộ giải mã các băng năng lượng kênh phải và trái được xử lý riêng biệt nếu cờ hiệu âm lập thể nổi này không được thiết lập. Trong trường hợp cờ hiệu âm lập thể nổi được thiết lập, cả năng lượng và sự ráp nối đều được thực hiện trong miền âm lập thể nổi. Thông tin âm lập thể ghép cho các vùng IGF được ký hiệu tương tự thông tin âm lập thể ghép cho việc mã hóa lỗi, bao gồm việc chỉ ra dấu hiệu trong trường hợp dự báo nếu hướng dự báo từ bộ trộn giảm tới phần dư hoặc ngược lại.

Năng lượng có thể được tính toán từ năng lượng được truyền trong miền L/R .

$$midNrg[k] = leftNrg[k] + rightNrg[k];$$

$$sideNrg[k] = leftNrg[k] - rightNrg[k];$$

với k chỉ số tần số trong miền biến đổi.

Giải pháp khác để tính toán và truyền năng lượng trực tiếp trong miền âm thanh lập thể ghép để các băng nơi mà âm thanh lập thể ghép hoạt động, vì thế nên không cần việc biến đổi năng lượng bổ sung tại phía bộ giải mã.

Ô nguồn luôn được tạo ra theo Ma trận giữa/bên:

$$midTile[k] = 0.5 \cdot (leftTile[k] + rightTile[k])$$

$$sideTile[k] = 0.5 \cdot (leftTile[k] - rightTile[k])$$

Điều chỉnh năng lượng:

$$midTile[k] = midTile[k] * midNrg[k];$$

$$sideTile[k] = sideTile[k] * sideNrg[k];$$

Âm lập thể ghép $\rightarrow$ Biến đổi LR:

Nếu không tham số dự báo bổ sung được mã hóa:

$$leftTile[k] = midTile[k] + sideTile[k]$$

$$rightTile[k] = midTile[k] - sideTile[k]$$

Nếu tham số dự báo bổ sung được mã hóa và nếu hướng được tạo tín hiệu là từ giữa tới bên:

$$sideTile[k] = sideTile[k] - predictionCoeff \cdot midTile[k]$$

$$leftTile[k] = midTile[k] + sideTile[k]$$

$$rightTile[k] = midTile[k] - sideTile[k]$$

Nếu hướng được tạo tín hiệu từ phía bên tới phía giữa:

$$midTile1[k] = midTile[k] - predictionCoeff \cdot sideTile[k]$$

$$leftTile[k] = midTile1[k] - sideTile[k]$$

$$rightTile[k] = midTile1[k] + sideTile[k]$$

Quy trình này đảm bảo rằng từ các ô được sử dụng để tái tạo các vùng đến tương quan cao và vùng đến được quét, các kênh trái và phải thu được vẫn biểu diễn nguồn âm thanh tương quan và được quét thậm chí nếu các vùng nguồn không tương quan, bảo toàn hình ảnh âm lập thể cho các vùng này.

Nói cách khác, trong dòng bit, các cờ hiệu âm lập thể nổi được truyền mà biểu thị L/R hoặc M/S như là ví dụ cho việc mã hóa âm lập thể nổi thông thường sẽ được sử dụng. Trong bộ giải mã, thứ nhất, tín hiệu lỗi được giải mã như được biểu thị bởi các cờ hiệu âm lập thể nổi cho các băng lỗi. Thứ hai, tín hiệu lỗi được lưu trữ trong cả sự biểu diễn L/R và M/S. Với việc điền đầy ô IGF, việc biểu diễn ô nguồn được chọn để phù hợp với việc biểu diễn ô mục tiêu như được biểu thị bởi thông tin âm lập thể ghép cho các băng IGF.

Tạo hình nhiễu âm theo thời gian (Temporal Noise Shaping-TNS) là kỹ thuật tiêu chuẩn và là một phần của AAC [11 – 13]. TNS có thể được xem như việc mở rộng của giản đồ cơ bản của bộ mã hóa cảm giác, chèn bước xử lý tùy chọn giữa giàn lọc và giai đoạn lượng tử hóa. Nhiệm vụ chính của môđun TNS là làm ẩn nhiễu âm lượng tử hóa được tạo ra trong vùng chắn theo thời gian của sự chuyển tiếp như các tín hiệu và do

đó dẫn đến giản đồ mã hóa hiệu quả hơn. Thứ nhất, TNS tính toán tập hợp các hệ số dự báo sử dụng "dự báo phía trước" trong miền biến đổi, ví dụ, MDCT. Những hệ số này sau đó được sử dụng để làm phẳng đường bao theo thời gian của tín hiệu. Vì lượng tử hoá tác động tới phổ đã lọc TNS, nhiều âm lượng tử hóa cũng phẳng theo thời gian. Bằng việc áp dụng việc lọc TNS nghịch đảo trên phía bộ giải mã, nhiều âm lượng tử hóa được tạo hình theo đường bao theo thời gian của bộ lọc TNS và do đó nhiều âm lượng tử hóa có được chặn bởi phần chuyển tiếp.

IGF được dựa trên sự biểu diễn MDCT. Để mã hóa hiệu quả, tốt hơn là sử dụng các khối dài xấp xỉ 20 miligiây (milisecond - ms). Nếu tín hiệu trong khối dài này chứa các phần chuyển tiếp, các tiếng vang trước và sau có thể nghe được xuất hiện trong các băng phổ IGF do việc điền đầy ô. Fig.7c cho thấy hiệu quả tiếng vang trước cụ thể trước khi bắt đầu phần chuyển tiếp do IGF. Trên phía trái, ảnh phổ của tín hiệu gốc được thể hiện và trên phía phải ảnh phổ của tín hiệu được mở rộng bằng thông không lọc TNS được thể hiện.

Hiệu quả tiếng vang trước này bị làm giảm bằng cách sử dụng TNS trong ngữ cảnh IGF. Ở đây, TNS được sử dụng như công cụ tạo hình ô theo thời gian (temporal tile shaping - TTS) như sự tái tạo phổ trong bộ giải mã được thực hiện trên tín hiệu dư TNS. Các hệ số dự báo TTS yêu cầu được tính toán và áp dụng sử dụng phổ đầy đủ trên phía bộ mã hóa như thường lệ. Các tần số bắt đầu và tần số dừng TNS/TTS không bị ảnh hưởng bởi tần số bắt đầu IGF $f_{IGFstart}$ của công cụ IGF. Trong so sánh với TNS kế thừa, tần số dừng TTS tăng đến tần số dừng của công cụ IGF, mà cao hơn $f_{IGFstart}$. Trên phía bộ giải mã, các hệ số TNS/TTS được áp dụng trên phổ đầy đủ lần nữa, cụ thể là, phổ lõi cộng với phổ được tái tạo cộng với các thành phần âm từ bản đồ âm (xem Fig.7e). Việc áp dụng TTS là cần thiết để tạo thành đường bao theo thời gian của phổ được tái tạo để phù hợp với đường bao của tín hiệu ban đầu lần nữa. Vì tiếng vang trước đã thể hiện được làm giảm. Ngoài ra, nó vẫn tạo hình nhiều âm lượng tử hóa trong tín hiệu dưới $f_{IGFstart}$ như thường lệ với TNS.

Trong các bộ giải mã kế thừa, việc ráp nối phổ trên tín hiệu âm thanh làm sai độ tương quan phổ tại biên nối và do đó làm suy giảm đường bao thời gian của tín hiệu âm thanh bằng việc đưa vào sự phân tán. Vì thế, lợi ích khác của việc thực hiện điền

đầy ô IGF trên tín hiệu dư là, sau việc áp dụng bộ lọc tạo hình, các biên ô được tương quan liền mạch, thu được sự tái tạo theo thời gian chính xác của tín hiệu.

Trong bộ mã hóa theo sáng chế, phổ có trải qua việc lọc TNS/TTS, xử lý màn chắn âm và sự ước lượng tham số IGF không có bất kỳ tín hiệu ở trên tần số bắt đầu IGF ngoại trừ các thành phần âm. Phổ rải rác này bây giờ được mã hóa bằng bộ mã hóa lõi sử dụng các nguyên lý của việc mã hóa số học và mã hóa dự báo. Các thành phần được mã hóa này cùng với các bit tạo tín hiệu từ dòng bit âm thanh.

Fig.2a minh họa việc thực hiện bộ giải mã tương ứng. Dòng bit trong Fig.2a tương ứng với tín hiệu âm thanh được nhập vào trong bộ tách kênh/bộ giải mã mà sẽ được nối, đối với Fig.1b, đến các khối 112 và 114. Bộ tách kênh dòng bit tách tín hiệu âm thanh đầu vào vào sự biểu diễn được mã hóa thứ nhất 107 của Fig.1b và sự biểu diễn được mã hóa thứ hai 109 của Fig.1b. Sự biểu diễn được mã hóa thứ nhất có tập hợp thứ nhất của các phần phổ thứ nhất được đưa vào khối giải mã kênh nối 204 tương ứng với bộ giải mã miền phổ 112 của Fig.1b. Sự biểu diễn được mã hóa thứ hai được nhập vào bộ giải mã theo tham số 114 không được minh họa trên Fig.2a sau đó được nhập vào khối IGF 202 tương ứng với bộ tái tạo tần số 116 của Fig.1b. Tập hợp thứ nhất gồm các vị trí phổ thứ nhất được yêu cầu cho sự tái tạo tần số được nhập vào khối IGF 202 qua đường dẫn 203. Hơn nữa, tiếp theo sự giải mã kênh nối 204, sự giải mã lõi cụ thể được áp dụng trong khối màn chắn âm 206 để đầu ra màn chắn âm 206 tương ứng với đầu ra của bộ giải mã miền phổ 112. Sau đó, sự tổ hợp bởi bộ tổ hợp 208 được thực hiện, tức là, sự xây dựng khung trong đó đầu ra của bộ tổ hợp 208 bây giờ có phổ phạm vi đầy đủ, nhưng vẫn trong miền được lọc TNS/TTS. Sau đó, trong khối 210, thuật toán TNS/TTS nghịch đảo được thực hiện sử dụng thông tin bộ lọc TNS/TTS được cung cấp qua đường dẫn 109, tức là, thông tin phụ TTS tốt hơn là được chứa trong sự biểu diễn được mã hóa thứ nhất được tạo ra bởi bộ mã hóa miền phổ 106 mà có thể, ví dụ, là bộ mã hóa lõi AAC hoặc USAC truyền thẳng, hoặc có thể cũng được chứa trong sự biểu diễn được mã hóa thứ hai. Tại đầu ra của khối 210, phổ hoàn chỉnh cho đến tần số tối đa được cung cấp mà là tần số phạm vi đầy đủ được định rõ bởi tốc độ lấy mẫu của tín hiệu đầu vào ban đầu. Sau đó, việc chuyển đổi phổ/thời gian được thực hiện trong giàn lọc tổng hợp 212 để cuối cùng thu được tín hiệu đầu ra âm thanh.

Fig.3a minh họa sự biểu diễn dưới dạng giản đồ của phổ. Phổ được chia nhỏ ra trong các băng thừa số định tỉ lệ SCB tại đó có bảy băng thừa số định tỉ lệ SCB1 tới SCB7 trong ví dụ minh họa của Fig.3a. Các băng thừa số định tỉ lệ có thể là băng thừa số định tỉ lệ AAC mà được định rõ trong tiêu chuẩn AAC và có băng thông tăng đến tần số cao như được minh họa trên Fig.3a dưới dạng giản đồ. Ưu tiên thực hiện sự điền đầy khoảng trống thông minh không từ phần đầu phổ, tức là, tại tần số thấp, nhưng để bắt đầu việc hoạt động IGF tại tần số bắt đầu IGF như được minh họa tại 309. Do đó, băng tần số lõi mở rộng từ tần số thấp nhất đến tần số bắt đầu IGF. Phía trên tần số bắt đầu IGF, việc phân tích phổ được áp dụng để phân tách các thành phần phổ có độ phân giải cao 304, 305, 306, 307 (tập hợp thứ nhất gồm các phần phổ thứ nhất) từ các thành phần có độ phân giải thấp được biểu diễn bởi bộ thứ hai của phần phổ thứ hai. Fig.3a minh họa phổ mà là đầu vào minh họa vào bộ mã hóa miền phổ 106 hoặc bộ mã hóa kênh nổi 228, tức là, bộ mã hóa lõi hoạt động trong phạm vi đầy đủ, nhưng mã hóa một lượng đáng kể các giá trị phổ 0, tức là, các giá trị phổ bằng 0 này được lượng tử hóa đến 0 hoặc là được thiết lập đến 0 trước khi lượng tử hóa hoặc sau khi lượng tử hóa. Dù thế nào đi nữa, bộ mã hóa lõi hoạt động trong phạm vi đầy đủ, tức là, nếu phổ như là được minh họa, tức là, bộ giải mã lõi không nhất thiết phải nhận thấy bất kì việc điền đầy khoảng trống thông minh hoặc mã hóa tập hợp thứ hai gồm các phần phổ thứ hai với độ phân giải phổ thấp hơn.

Tốt hơn là, độ phân giải được định rõ bởi sự mã hóa theo vạch của các vạch phổ như các vạch MDCT, trong khi độ phân giải thứ hai hoặc độ phân giải thấp được xác định bằng cách, ví dụ, tính toán chỉ giá trị phổ đơn mỗi băng thừa số định tỉ lệ, trong đó băng thừa số định tỉ lệ che phủ một số vạch tần số. Do đó, độ phân giải thấp thứ hai là, đối với độ phân giải phổ của nó, thấp hơn nhiều so với độ phân giải thứ nhất hoặc độ phân giải cao được định rõ bởi việc mã hóa theo vạch được ứng dụng bởi bộ mã hóa lõi như AAC hoặc USAC.

Đối với thừa số định tỉ lệ hoặc việc tính toán năng lượng, tình huống này được minh họa trong Fig.3b. Do thực tế rằng bộ mã hóa là bộ mã hóa lõi và do thực tế rằng có thể, nhưng không nhất thiết phải là, các thành phần của tập hợp thứ nhất gồm các phần phổ trong mỗi băng, bộ mã hóa lõi tính toán thừa số định tỉ lệ cho mỗi băng không chỉ trong phạm vi lõi dưới tần số bắt đầu IGF 309, mà cũng trên tần số bắt đầu

IGF cho đến khi tần số tối đa $f_{IGFstop}$ mà nhỏ hơn hoặc bằng một nửa tần số lấy mẫu, tức là, $f_s/2$. Do đó, các phần âm được mã hóa 302, 304, 305, 306, 307 của Fig.3a và, trong phương án này cùng với các thừa số định tỉ lệ SCB1 đến SCB7 tương ứng với dữ liệu phổ độ phân giải cao. Dữ liệu phổ độ phân giải thấp được tính toán bắt đầu từ tần số bắt đầu IGF và tương ứng với các giá trị thông tin năng lượng E_1, E_2, E_3, E_4 , mà được truyền cùng với các thừa số định tỉ lệ SF4 đến SF7.

Cụ thể, khi bộ mã hóa lỗi dưới điều kiện tốc độ bit thấp, hoạt động điền đầy nhiều âm bổ sung trong băng lỗi, tức là, thấp hơn trong tần số so với tần số bắt đầu IGF, tức là, trong băng thừa số định tỉ lệ SB1 tới SB3 ngoài ra có thể được áp dụng. Trong việc điền đầy nhiều âm, một số vạch phổ liền kề tồn tại mà đã được lượng tử hóa tới 0. Trên phía bộ giải mã, các giá trị phổ được lượng tử hóa tới 0 này được tái tổng hợp và các giá trị phổ được tái tổng hợp được điều chỉnh trong độ lớn của chúng sử dụng năng lượng điền đầy nhiều âm như NF_2 được minh họa tại 308 trong Fig.3b. Năng lượng điền đầy nhiều âm, mà có thể được cho theo các số hạng tuyệt đối hoặc các số hạng tương đối cụ thể đối với hệ số tỉ lệ như trong USAC tương ứng với năng lượng của bộ giá trị phổ được lượng tử hóa tới 0. Các vạch phổ điền đầy nhiều âm cũng có thể được xem như tập hợp thứ ba gồm các phần phổ thứ ba mà được tái tạo bằng sự tổng hợp điền đầy nhiều âm dễ dàng mà không có bất kỳ thuật toán IGF nào phụ thuộc vào sự tái tạo tần số sử dụng các ô tần số từ các tần số để khôi phục các ô tần số sử dụng các giá trị phổ từ phạm vi nguồn và thông tin năng lượng E_1, E_2, E_3, E_4 .

Tốt hơn là, các băng, mà thông tin năng lượng được tính toán cho các băng này trùng với các băng thừa số định tỉ lệ. Trong các phương án khác, tạo nhóm giá trị thông tin năng lượng được áp dụng sao cho, ví dụ, đối với các băng thừa số định tỉ lệ 4 và 5, chỉ giá trị thông tin năng lượng đơn được truyền, nhưng thậm chí trong phương án này, các biên của các băng khôi phục được nhóm trùng với các biên của băng thừa số định tỉ lệ. Nếu việc phân tách băng khác nhau được áp dụng, sau đó việc tính toán lại hoặc tính toán việc lượng tử hóa nhất định có thể được áp dụng, và điều này có thể có ý nghĩa phụ thuộc vào việc thực hiện nhất định.

Tốt hơn là, bộ mã hóa miền phổ 106 của Fig.1a là bộ mã hóa dẫn theo tâm thính học như minh họa trong Fig.4a. Thông thường, như ví dụ được minh họa trong tiêu

chuẩn MPEG2/4 AAC hoặc tiêu chuẩn MPEG1/2, Layer 3, tín hiệu âm thanh sẽ được mã hóa sau khi đã được biến đổi thành phạm vi phổ (401 trên Fig.4a) được chuyển tiếp đến bộ tính toán thừa số định tỉ lệ 400. Bộ tính toán thừa số định tỉ lệ được điều khiển bằng mô hình tâm thính học, ngoài ra nhận tín hiệu âm thanh sẽ được lượng tử hóa hoặc nhận, như trong tiêu chuẩn MPEG1/2 Layer 3 hoặc tiêu chuẩn MPEG AAC, sự biểu diễn phổ phức của tín hiệu âm thanh. Mô hình tâm thính học tính toán, cho từng băng thừa số định tỉ lệ, thừa số định tỉ lệ biểu diễn ngưỡng tâm thính học. Ngoài ra, các thừa số định tỉ lệ sau đó, bằng việc cùng hoạt động của các vòng lặp bên trong và bên ngoài đã biết hoặc bằng bất kỳ quy trình mã hóa thích hợp nào khác được điều chỉnh để đáp ứng các điều kiện tốc độ bit nhất định. Sau đó, một mặt các giá trị phổ sẽ được lượng tử hóa và mặt khác các thừa số định tỉ lệ được tính toán được nhập vào bộ xử lý lượng tử hóa 404. Trong thuật toán bộ mã hóa âm thanh dễ dàng, các giá trị phổ sẽ được định tỉ lệ được gán trọng số bằng các thừa số định tỉ lệ và, các giá trị phổ được gán trọng số sau đó được đưa vào bộ lượng tử hóa cố định thường có chức năng nén đến các phạm vi biên độ phía trên. Sau đó, tại đầu ra của bộ xử lý lượng tử hóa tại đó thực hiện lượng tử hóa chỉ số mà sau đó được chuyển vào trong bộ mã hóa entropi đặc trưng có sự mã hóa cụ thể và rất hiệu quả cho bộ lượng tử hóa 0 chỉ thị cho các giá trị tần số liên kề hoặc, cũng được gọi trong kỹ thuật, là "sự vận hành" của các giá trị 0.

Tuy nhiên, trong bộ mã hóa âm thanh của Fig.1a, bộ xử lý lượng tử hóa đặc trưng nhận thông tin trên các phần phổ thứ hai từ bộ phân tích phổ. Do đó, bộ xử lý lượng tử hóa 404 chắc chắn là, trong đầu ra của bộ xử lý lượng tử hóa 404, các phần phổ thứ hai được nhận biết bởi bộ phân tích phổ 102 là 0 hoặc có sự biểu diễn được xác nhận bởi bộ mã hóa hoặc bộ giải mã như sự biểu diễn 0 mà có thể được mã hóa rất hiệu quả, đặc biệt khi có "sự vận hành" của giá trị 0 trong phổ.

Fig.4b minh họa sự thực hiện của bộ xử lý lượng tử hóa. Các giá trị phổ MDCT có thể được đưa vào thành tập hợp đến khối 0 410. Khi đó, các phần phổ thứ hai đã được thiết lập về 0 trước khi sự gán trọng số bằng các thừa số định tỉ lệ trong khối 412 được thực hiện. Trong sự thực hiện bổ sung, khối 410 không được đề xuất, nhưng sự hợp tác được thiết lập về 0 được thực hiện trong khối 418 tiếp theo khối gán trọng số 412. Trong sự thực hiện điều đặn khác, thuật toán được thiết lập về 0 có thể được thực hiện trong khối được thiết lập về 0 422 tiếp theo sự lượng tử hóa trong khối lượng tử

hóa 420. Trong việc thực hiện này, các khối 410 và 418 sẽ không được trình bày. Nhìn chung, ít nhất một trong số các khối 410, 418, 422 được cung cấp phụ thuộc vào sự thực hiện cụ thể.

Sau đó, tại đầu ra của khối 422, phổ được lượng tử hóa thu được tương ứng với điều đã minh họa trên Fig.3a. Phổ được lượng tử hóa này sau đó được nhập vào bộ mã hóa entropi như 232 trong Fig.2b mà có thể là bộ mã hóa Huffman hoặc bộ mã hóa số học như, ví dụ, được định rõ trong tiêu chuẩn USAC.

Các khối được thiết lập về 0 410, 418, 422, mà được đề xuất thay thế các khối xen kẽ hoặc song được điều khiển bởi bộ phân tích phổ 424. Bộ phân tích phổ tốt hơn là bao gồm bất kỳ việc thực hiện của bộ phát hiện âm đã biết hoặc bao gồm bất kỳ dạng khác của bộ phát hiện hoạt động để phân tách phổ thành các thành phần sẽ được mã hóa với độ phân giải cao và các thành phần sẽ được mã hóa với độ phân giải thấp. Các thuật toán khác được thực hiện trong bộ phân tích phổ có thể là bộ phát hiện hoạt động giọng nói, bộ phát hiện nhiều âm, bộ phát hiện tiếng nói hay bất kỳ bộ phát hiện khác quyết định, phụ thuộc vào thông tin phổ hoặc siêu dữ liệu được kết hợp trên các yêu cầu độ phân giải cho các phần phổ khác nhau.

Fig.5a minh họa việc thực hiện được ưu tiên của bộ chuyển đổi phổ thời gian 100 của Fig.1a như ví dụ, đã bổ sung trong AAC hoặc USAC. Bộ chuyển đổi phổ thời gian 100 bao gồm bộ tạo cửa sổ 502 được điều khiển bởi bộ phát hiện sự chuyển tiếp 504. Khi bộ phát hiện sự chuyển tiếp 504 phát hiện sự chuyển tiếp, khi đó sự luân phiên từ các cửa sổ dài đến các cửa sổ ngắn được tạo tín hiệu đến bộ tạo cửa sổ. Bộ tạo cửa sổ 502 sau đó tính toán, cho các khối chồng lấp, các khung được tạo cửa sổ, trong đó mỗi khung được tạo cửa sổ thường có hai giá trị N như 2048 giá trị. Sau đó, việc biến đổi trong bộ chuyển khối 506 được thực hiện, và bộ biến đổi khối điển hình này cung cấp bổ sung lấy thập phân sao cho lấy thập phân/biến đổi kết hợp được thực hiện để thu được khung phổ với các giá trị N như các giá trị phổ MDCT. Do đó, hoạt động cửa sổ dài, khung tại đầu ra của khối 506 bao gồm hai giá trị N như các giá trị 2048 và khung phổ sau đó có giá trị 1024. Sau đó, tuy nhiên, bộ chuyển được thực hiện cho các khối ngắn, khi tám khối ngắn được thực hiện tại từng khối ngắn có 1/8 giá trị miền thời gian cửa sổ được so sánh với cửa sổ dài và từng khối phổ có 1/8 giá trị phổ khi so sánh với

khối dài. Do đó, sự lấy thập phân này được tổ hợp với thuật toán chồng lấp 50% của bộ tạo cửa sổ, phổ là phiên bản bản được lấy mẫu tới hạn của tín hiệu âm thanh miền thời gian 99.

Tiếp theo, viện dẫn đến Fig.5b minh họa thực hiện cụ thể bộ tái tạo tần số 116 và bộ chuyển đổi phổ-thời gian 118 của Fig.1b, hoặc của thuật toán được tổ hợp của các khối 208, 212 của Fig.2a. Trên Fig.5b, băng khôi phục cụ thể được xem xét như băng thừa số định tỉ lệ 6 của Fig.3a. Phần phổ thứ nhất trong băng khôi phục này, tức là, phần phổ thứ nhất 306 của Fig.3a được nhập vào khối dựng/điều chỉnh khung 510. Hơn nữa, phần phổ thứ hai được khôi phục cho băng thừa số định tỉ lệ 6 cũng được đưa vào bộ dựng/bộ điều chỉnh khung 510. Hơn nữa, thông tin năng lượng như E_3 của Fig.3b cho băng thừa số định tỉ lệ 6 cũng được đưa vào khối 510. Phần phổ thứ hai được khôi phục trong băng khôi phục đã được tạo ra bằng sự điền đầy ô tần số bằng cách sử dụng phạm vi nguồn và băng khôi phục sau đó tương ứng với phạm vi đích. Sau đây, việc điều chỉnh năng lượng của khung được thực hiện để cuối cùng thu được khung được khôi phục hoàn chỉnh có các giá trị N, chẳng hạn, được thu tại đầu ra của bộ tổ hợp 208 của Fig.2a. Sau đó, trong khối 512, phép biến đổi/phép nội suy khối nghịch đảo được thực hiện để thu được 248 giá trị miền thời gian cho, ví dụ, 124 giá trị phổ tại đầu vào của khối 512. Sau đó, thuật toán tạo cửa sổ tổng hợp được thực hiện trong khối 514 mà lần nữa được điều khiển bởi sự chỉ thị cửa sổ dài/cửa sổ ngắn được truyền như thông tin phụ trong tín hiệu âm thanh được mã hóa. Sau đó, trong khối 516, thuật toán chồng lấp/cộng với khung thời gian trước được thực hiện. Tốt hơn là, MDCT áp dụng sự chồng lấp 50% sao cho, đối với từng khung thời gian mới của $2N$ giá trị, N giá trị miền thời gian là đầu ra cuối cùng. Sự chồng lấp 50% được ưu tiên hơn do thực tế là nó cung cấp việc lấy mẫu tới hạn và sự giao chéo liên tục từ một khung tới khung kế tiếp do thuật toán chồng lấp/cộng trong khối 516.

Như được miêu tả tại 301 trên Fig.3a, thuật toán điền đầy nhiễu âm có thể được áp dụng bổ sung không chỉ dưới tần số bắt đầu IGF, mà còn trên tần số bắt đầu IGF như vậy cho băng khôi phục được dự tính trùng với băng thừa số định tỉ lệ 6 của Fig.3a. Sau đó, các giá trị phổ điền đầy nhiễu âm có thể cũng được đưa vào bộ dựng/bộ điều chỉnh khung 510 và việc điều chỉnh các giá trị phổ điền đầy nhiễu âm có thể cũng được áp dụng trong khối này hoặc giá trị phổ điền đầy nhiễu âm có thể được

điều chỉnh sử dụng năng lượng điền đầy nhiều âm trước khi được nhập vào bộ dụng/bộ điều chỉnh khung 510.

Tốt hơn là, thuật toán IGF, tức là, thuật toán điền đầy ô tần số sử dụng các giá trị phổ từ các phần khác có thể được áp dụng trong phổ hoàn chỉnh. Do đó, hoạt động điền đầy ô phổ có thể không chỉ được áp dụng trong băng cao trên tần số bắt đầu IGF mà còn được áp dụng trong cả băng thấp. Hơn nữa, việc điền đầy nhiều âm ô tần số không điền đầy ô tần số có thể cũng được áp dụng không chỉ dưới tần số bắt đầu IGF mà còn trên tần số bắt đầu IGF. Tuy nhiên, các tác giả sáng chế đã tìm ra rằng việc mã hóa âm thanh có chất lượng cao và hiệu quả cao có thể thu được khi hoạt động điền đầy nhiều âm được giới hạn tới phạm vi tần số dưới tần số bắt đầu IGF và khi hoạt động điền đầy ô tần số bị hạn chế tới phạm vi tần số trên tần số bắt đầu IGF như mô tả trên Fig.3a.

Tốt hơn là, các ô đích (target tile - TT) (có các tần số lớn hơn tần số bắt đầu IGF) được giới hạn tới các biên băng thừa số định tỉ lệ của bộ mã hóa tốc độ đầy đủ. Các ô nguồn (source tile - ST), từ thông tin thu được, cụ thể là, đối với các tần số thấp hơn tần số bắt đầu IGF không được giới hạn bởi các biên băng thừa số định tỉ lệ. Kích thước của ST nên tương ứng với kích thước của TT được kết hợp. Điều này được mô tả sử dụng ví dụ dưới đây. TT[0] có chiều dài là 10 ngăn MDCT. Điều này tương ứng chính xác với chiều dài của hai SCB tiếp theo (như $4 + 6$). Sau đó, tất cả ST khả thi mà tương quan với TT[0], cũng có chiều dài 10 ngăn. Ô đích thứ hai TT[1] liên kế với TT[0] có chiều dài là 15 ngăn 1 (SCB có chiều dài $7+8$). Sau đó, ST để có chiều dài là 15 ngăn hơn 10 ngăn đối với TT[0].

Nếu trường hợp xảy ra rằng người ta không thể tìm TT cho ST với chiều dài của ô đích (khi, ví dụ, chiều dài của TT lớn hơn phạm vi nguồn hiện có), khi đó sự tương quan không được tính toán và phạm vi nguồn được sao chép số lượng thời gian vào trong TT này (việc sao chép được thực hiện lần này đến lần khác khác để đường tần số cho tần số thấp nhất của việc sao chép thứ hai theo ngay sau - theo tần số - đường tần số cho tần số cao nhất của việc sao chép thứ nhất), đến khi ô mục tiêu TT được điền đầy hoàn toàn.

Tiếp theo, viện dẫn đến Fig.5c minh họa phương án được ưu tiên khác của bộ tái

tạo tần số 116 của Fig.1b hoặc khối IGF 202 của Fig.2a. Khối 522 là bộ tái tạo ô tần số nhận, không chỉ ID băng mục tiêu mà còn nhận thêm ID băng nguồn. Ví dụ, đã được xác định trên phía bộ mã hóa là băng thừa số định tỉ lệ 3 của Fig.3a rất phù hợp cho việc tái tạo băng thừa số định tỉ lệ 7. Do đó, ID băng nguồn sẽ là 2 và ID băng đích sẽ là 7. Dựa trên thông tin này, bộ tái tạo ô tần số 522 áp dụng sự sao chép hoặc thuật toán điền đầy ô sóng hài hoặc bất kỳ thuật toán điền đầy ô nào khác để tạo ra phần thứ hai thô của các thành phần phổ 523. Phần thứ hai thô của các thành phần phổ có độ phân giải tần số giống với độ phân giải tần số nằm trong tập hợp thứ nhất gồm các phần phổ thứ nhất.

Khi đó, phần phổ thứ nhất của băng khôi phục như 307 của Fig.3a được nhập vào bộ dựng khung 524 và phần thứ hai thô 523 cũng được nhập vào bộ dựng khung 524. Khi đó, khung được khôi phục được điều chỉnh bởi bộ điều chỉnh 526 sử dụng thừa số khuếch đại cho băng khôi phục được tính toán bằng bộ tính toán thừa số khuếch đại 528. Tuy nhiên, quan trọng hơn, phần phổ thứ nhất trong khung không bị ảnh hưởng bởi bộ điều chỉnh 526, mà chỉ phần thứ hai thô cho khung khôi phục được ảnh hưởng bởi bộ điều chỉnh 526. Để đạt được mục đích này, bộ tính toán thừa số khuếch đại 528 phân tích băng nguồn hoặc phần thứ hai thô 523 và phân tích thêm phần phổ thứ nhất trong băng khôi phục để cuối cùng tìm ra rằng thừa số khuếch đại đúng 527 để năng lượng của khung được điều chỉnh được phát ra bằng bộ điều chỉnh 526 có năng lượng E4 khi băng thừa số định tỉ lệ 7 được dự tính.

Trong phạm vi này, rất quan trọng để ước tính độ chính xác khôi phục tần số cao của sáng chế so với HE-AAC. Điều này được giải thích với băng thừa số định tỉ lệ 7 trên Fig.3a. Giả định rằng bộ mã hóa đã có trước trong tình trạng kỹ thuật như minh họa trong Fig.13a sẽ phát hiện phần phổ 307 sẽ được mã hóa với độ phân giải cao như "sóng hài khuyết". Sau đó, năng lượng của thành phần phổ này sẽ được truyền cùng với thông tin đường bao phổ cho băng khôi phục như băng thừa số định tỉ lệ 7 đến bộ giải mã. Sau đó, bộ giải mã sẽ tái tạo sóng hài bị mất. Tuy nhiên, giá trị phổ, tại đó sóng hài 307 bị mất sẽ được khôi phục bởi bộ giải mã đã có trong tình trạng kỹ thuật của Fig.13b sẽ ở giữa băng 7 tại tần số được chỉ thị bởi tần số khôi phục 390. Do đó, sáng chế này tránh được lỗi tần số 391 mà được giới thiệu trong bộ giải mã đã có trong tình trạng kỹ thuật của Fig.13d.

Trong thực hiện, bộ phân tích phổ cũng được thực hiện để tính toán sự tương đồng giữa các phần phổ thứ nhất và các phần phổ thứ hai và để xác định dựa trên sự tương đồng được tính toán, đối với phần phổ thứ hai trong phạm vi khôi phục, phần phổ thứ nhất trùng khớp với phần phổ thứ hai càng nhiều càng tốt. Sau đó, trong thực hiện khoảng nguồn/khoảng đích có thể thay đổi, bộ mã hóa theo tham số sẽ được đưa bổ sung vào trong sự biểu diễn được mã hóa thứ hai thông tin phù hợp biểu thị cho từng khoảng đích khoảng nguồn phù hợp. Trên phía bộ giải mã, thông tin này sau đó sẽ được sử dụng bằng bộ tạo ô tần số 522 của Fig.5c minh họa sự tạo ra phần thứ hai thô 523 dựa trên ID băng nguồn và ID băng đích.

Hơn nữa, như được minh họa trên Fig.3a, bộ phân tích phổ được tạo cấu hình để phân tích sự biểu diễn phổ tăng đến tần số phân tích tối đa chỉ là lượng nhỏ dưới một nửa tần số lấy mẫu và tốt hơn ít nhất là một phần tư của tần số lấy mẫu hoặc cao hơn.

Như được minh họa, bộ mã hóa hoạt động không lấy mẫu giảm và bộ giải mã hoạt động không lấy mẫu tăng. Nói cách khác, bộ mã hóa âm thanh miền phổ được tạo cấu hình để tạo ra sự biểu diễn phổ có tần số Nyquist được xác định bởi tốc độ lấy mẫu của tín hiệu âm thanh đầu vào gốc.

Thêm vào đó, như minh họa trên Fig.3a, bộ phân tích phổ được tạo cấu hình để phân tích sự biểu diễn phổ bắt đầu với tần số bắt đầu điền đầy khoảng trống và kết thúc với tần số tối đa được biểu diễn bởi tần số tối đa nằm trong sự biểu diễn phổ, trong đó phần phổ mở rộng từ tần số tối thiểu lên đến tần số bắt đầu điền đầy khoảng trống thuộc tập hợp thứ nhất gồm các phần phổ và trong đó phần phổ khác như 304, 305, 306, 307 có giá trị tần số trên tần số điền đầy khoảng trống nằm bổ sung trong tập hợp thứ nhất gồm các phần phổ thứ nhất.

Như được tóm tắt, bộ giải mã âm thanh miền phổ 112 được tạo cấu hình để tần số tối đa được biểu diễn bằng giá trị phổ trong sự biểu diễn được giải mã thứ nhất là bằng với tần số tối đa nằm trong sự biểu diễn theo thời gian có tốc độ lấy mẫu trong đó giá trị phổ để tần số tối đa trong tập hợp thứ nhất gồm các phần phổ thứ nhất bằng 0 hoặc khác 0. Dù thế nào, để tần số tối đa trong tập hợp thứ nhất gồm các thành phần phổ thừa số định tỉ lệ cho bằng thừa số định tỉ lệ tồn tại, mà được tạo ra và truyền bắt chấp việc tất cả giá trị phổ trong bằng thừa số định tỉ lệ này được thiết lập về 0 hay không

như đã thảo luận trong phạm vi của Fig.3a và Figs.3b.

Do đó, sáng chế này có ưu điểm là đối với các kỹ thuật theo tham số khác để làm tăng hiệu quả nén, ví dụ, việc thay thế nhiều âm và điền đầy nhiều âm (các kỹ thuật này dành riêng cho việc biểu diễn hiệu quả của nội dung tín hiệu cục bộ như nhiều âm) sáng chế cho phép tái tạo tần số chính xác của các thành phần âm. Cho đến nay, không kỹ thuật thuộc tình trạng kỹ thuật nào hướng đến sự biểu diễn theo tham số hiệu quả của nội dung tín hiệu tùy ý bằng cách điền đầy khoảng trống phổ không hạn chế của sự phân chia ưu tiên được cố định trong băng thấp (LF) và băng cao (HF).

Các phương án của hệ thống theo sáng chế cải thiện cách tiếp cận của tình trạng kỹ thuật và do đó cung cấp hiệu quả nén cao, không hoặc chỉ một chút sự không hài lòng về cảm giác của băng thông âm thanh đầy đủ thậm chí với tốc độ bit thấp.

Hệ thống thông thường gồm có:

- mã hóa lõi băng đầy đủ
- điền đầy khoảng trống thông minh (điền đầy ô hoặc điền đầy nhiều âm)
- các phần âm điệu rải rác trong lõi được chọn bởi màn chắn âm.
- mã hóa cặp âm lập thể ghép cho băng đầy đủ, gồm có việc điền đầy ô
- TNS trên ô
- làm trắng phổ trong khoảng IGF

Bước thứ nhất hướng về hệ thống hiệu quả hơn là để loại bỏ sự cần thiết để biến đổi dữ liệu phổ thành miền biến đổi thứ hai khác với miền của bộ mã hóa lõi. Vì đa số bộ mã hóa giải mã âm thanh, ví dụ, như AAC, sử dụng MDCT như phép biến đổi cơ bản, cũng hữu ích để thực hiện BWE trong miền MDCT. Yêu cầu thứ hai cho hệ thống BWE sẽ cần thiết để bảo toàn lưới âm bằng cách thậm chí các thành phần âm HF được bảo toàn và chất lượng của âm thanh được mã hóa là ưu việt hơn các hệ thống hiện có. Việc đảm bảo cả các yêu cầu đã đề cập ở trên, hệ thống được đề xuất được gọi là điền đầy khoảng trống thông minh (Intelligent Gap Filling - IGF). Fig.2b thể hiện sơ đồ khối của hệ thống được đề xuất trên phía bộ mã hóa và Fig.2 thể hiện hệ thống trên phía bộ giải mã.

Tiếp theo, khung xử lý sau được mô tả đối với Fig.13A và Fig.13B để minh họa rằng sáng chế cũng có thể được thực hiện trong bộ khôi phục tần số cao 1330 trong phương án xử lý sau này.

Fig.13 minh họa sơ đồ giản lược của bộ mã hóa âm thanh cho công nghệ mở rộng băng thông như, ví dụ, được sử dụng trong mã hóa âm thanh tiên tiến hiệu quả cao (High Efficiency Advanced Audio Coding - HE-AAC). Tín hiệu âm thanh tại đường 1300 được đưa vào trong hệ thống bộ lọc bao gồm thông thấp 1302 và thông cao 1304. Đầu ra tín hiệu bởi bộ lọc thông cao 1304 được đưa vào trong bộ trích/bộ mã hóa tham số 1306. Bộ trích/bộ mã hóa tham số 1306 được tạo cấu hình để tính toán và mã hóa các tham số như, ví dụ, tham số đường bao phổ, tham số bổ sung nhiễu âm, tham số sóng hài bị thiếu, hoặc tham số lọc nghịch đảo. Các tham số được chiết xuất này được nhập vào trong bộ dồn kênh dòng bit 1308. Tín hiệu đầu ra thông thấp được nhập vào bộ xử lý biến hình chứa chức năng của bộ lấy mẫu giảm 1310 và bộ mã hóa lõi 1312. Thông thấp 1302 thu hẹp băng thông để được mã hóa đến băng thông nhỏ hơn đáng kể hơn là diễn ra trong tín hiệu âm thanh đầu vào ban đầu trên đường 1300. Điều này cung cấp độ khuếch đại mã hóa đáng kể do thực tế rằng toàn bộ các chức năng xảy ra trong bộ mã hóa lõi chỉ phải hoạt động trên tín hiệu với băng thông được giảm. Khi đó, ví dụ, băng thông của tín hiệu âm thanh trên đường 1300 là 20kHz và khi bộ lọc thông thấp 1320 ví dụ có băng thông là 4kHz, để hoàn thành lý thuyết lấy mẫu, đây là tính lý thuyết đủ để tín hiệu tiếp theo tới bộ lấy mẫu giảm có tần số lấy mẫu là 8kHz, đó là sự giảm bớt đáng kể để tốc độ lấy mẫu được yêu cầu cho tín hiệu âm thanh 1300 tối thiểu là 40kHz.

Fig.13b minh họa biểu đồ giản lược bộ giải mã mở rộng băng thông tương ứng. Bộ giải mã bao gồm bộ dồn kênh dòng bit 1320. Bộ tách kênh dòng bit 1320 chiết tín hiệu đầu vào cho bộ giải mã lõi 1322 và tín hiệu đầu vào cho bộ giải mã theo tham số 1324. Bộ giải mã lõi xuất ra tín hiệu có, trong ví dụ nêu trên, tốc độ lấy mẫu là 8kHz và, do đó, băng thông là 4kHz trong khi, đối với sự khôi phục băng thông hoàn toàn, tín hiệu đầu ra của bộ khôi phục tần số cao 1330 phải là tại 20kHz yêu cầu tốc độ lấy mẫu ít nhất là 40kHz. Để thực hiện tính khả thi này, bộ xử lý giải mã có chức năng của bộ lấy mẫu tăng 1325 và giàn lọc 1326 được yêu cầu. Bộ khôi phục tần số cao 1330 sau đó nhận tín hiệu tần số thấp được phân tích tần số được phát ra bởi giàn lọc 1326

và khôi phục phạm vi tần số được xác định bằng bộ lọc thông cao 1304 của Fig.13a sử dụng sự biểu diễn theo tham số của băng tần số cao. Bộ khôi phục tần số cao 1330 có các chức năng khác nhau như tái tạo phạm vi tần số phía trên sử dụng phạm vi nguồn trong phạm vi tần số thấp, điều chỉnh đường bao phổ, chức năng bổ sung nhiễu âm và chức năng đưa sóng hài bị thiếu vào trong phạm vi tần số phía trên và, nếu được áp dụng và tính toán trong bộ mã hóa của Fig.13a, hoạt động lọc nghịch đảo để tính toán với thực tế là phạm vi tần số cao hơn thường không có âm như khoảng tần số thấp hơn. Trong HE-AAC, sóng hài khuyết được tái tổng hợp trên phía bộ giải mã và được đặt chính xác ở giữa băng khôi phục. Vì thế, tất cả các đường sóng hài bị thiếu đã được xác định trong băng khôi phục nhất định không được đặt tại các giá trị tần số tại đó mà tại đó chúng đã được đặt trong tín hiệu ban đầu. Thay vào đó, những đường sóng hài bị thiếu này được đặt tại các tần số trong trung tâm của băng nhất định. Do vậy, khi đường sóng hài bị thiếu trong tín hiệu ban đầu được đặt rất gần biên băng khôi phục trong tín hiệu ban đầu, lỗi trong tần số được đưa vào bằng việc đặt đường sóng hài bị thiếu trong tín hiệu đã được khôi phục tại trung tâm của băng gần tới 50% băng khôi phục riêng, mà tham số được tạo ra và được truyền cho nó.

Hơn nữa, mặc dù bộ mã hóa lõi âm thanh điển hình hoạt động trong miền phổ, tuy nhiên bộ giải mã lõi tạo ra tín hiệu miền thời gian cái mà sau đó, một lần nữa, được chuyển đổi thành miền phổ bởi chức năng băng bộ lọc 1326. Điều này dẫn đến các độ trễ xử lý bổ sung, có thể đưa các thành phần lạ do việc xử lý trước sau là thứ nhất biến đổi từ miền phổ thành miền tần số và một lần nữa biến đổi cụ thể thành miền tần số khác, và tất nhiên, điều này cũng yêu cầu lượng đáng kể việc tính toán phức tạp và bởi nguồn điện, mà cụ thể là vấn đề khi công nghệ mở rộng băng thông được ứng dụng trong các thiết bị di động như điện thoại di động, máy tính bảng hoặc máy tính xách tay, v.v..

Mặc dù một số khía cạnh đã được mô tả trong phạm vi của thiết bị mã hóa hoặc giải mã, rõ ràng rằng các khía cạnh này cũng biểu diễn sự mô tả của phương pháp tương ứng, trong đó khối hoặc thiết bị tương ứng với bước phương pháp hoặc dấu hiệu của bước phương pháp. Tương tự, các khía cạnh được mô tả trong phạm vi của bước phương pháp cũng biểu diễn sự mô tả của khối hoặc mục hoặc dấu hiệu tương ứng của thiết bị tương ứng. Một số hoặc tất cả các bước phương pháp có thể được thực hiện bởi

(hoặc sử dụng) thiết bị phần cứng, như ví dụ, bộ vi xử lý, máy tính có thể lập trình hoặc mạch điện tử. Theo một số phương án, một số một hoặc nhiều bước phương pháp quan trọng nhất có thể được thực hiện bởi thiết bị như vậy.

Phụ thuộc vào các yêu cầu thực hiện nhất định, các phương án của sáng chế có thể được thực hiện trong phần cứng hoặc trong phần mềm. Phương án có thể được thực hiện sử dụng vật ghi lưu trữ không tạm thời như vật ghi lưu trữ kỹ thuật số, ví dụ, đĩa mềm, ổ đĩa cứng (Hard Disk Drive - HDD), DVD, Blu-Ray CD, ROM, PROM, và EPROM, EEPROM hoặc bộ nhớ FLASH, có các tín hiệu điều khiển có thể đọc được bằng điện được lưu trữ trên đó, mà kết hợp (hoặc có thể kết hợp) với hệ thống máy tính có thể lập trình được sao cho phương pháp tương ứng được thực hiện. Do đó, vật ghi lưu trữ dạng số có thể đọc được bằng máy tính.

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

Nói chung, các phương án của sáng chế có thể được thực thi như sản phẩm chương trình máy tính với mã chương trình, mã chương trình có hiệu lực để thực thi một trong các phương pháp khi sản phẩm chương trình máy tính chạy trên máy tính. Ví dụ, mã chương trình có thể được lưu trữ trên vật mang có thể đọc được bằng máy.

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

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

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

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

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

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

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

Trong một số phương án, thiết bị logic có thể lập trình được (ví dụ, mảng cổng có thể lập trình được dạng trường) có thể được sử dụng để thực hiện một số hoặc tất cả các chức năng của các phương pháp được mô tả ở đây. Theo một số phương án, mảng cổng lập trình được dạng trường có thể kết hợp với bộ vi xử lý để thực hiện một trong số các phương pháp được mô tả ở đây. Thông thường, các phương pháp được ưu tiên thực hiện bởi thiết bị phần cứng bất kỳ.

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

YÊU CẦU BẢO HỘ

1. Thiết bị tạo ra tín hiệu âm thanh nâng cao từ tín hiệu đầu vào (600), trong đó tín hiệu âm thanh nâng cao có các giá trị phổ cho vùng phổ nâng cao, các giá trị phổ cho vùng phổ nâng cao không nằm trong tín hiệu đầu vào (600), thiết bị bao gồm:

bộ ánh xạ (602) để ánh xạ vùng phổ nguồn của tín hiệu đầu vào thành vùng đích trong vùng phổ nâng cao; và

bộ điền đầy nhiễu âm (604) được tạo cấu hình để tạo ra các giá trị nhiễu âm thứ nhất cho vùng điền đầy nhiễu âm (302) trong vùng phổ nguồn của tín hiệu đầu vào và để tạo ra các giá trị nhiễu âm thứ hai cho vùng nhiễu âm trong vùng đích, trong đó các giá trị nhiễu âm thứ hai được giải tương quan từ các giá trị nhiễu âm thứ nhất hoặc để tạo ra các giá trị nhiễu âm thứ hai cho vùng nhiễu âm trong vùng đích, trong đó các giá trị nhiễu âm thứ hai được giải tương quan từ các giá trị nhiễu âm thứ nhất trong vùng nguồn, trong đó các giá trị nhiễu âm thứ nhất trong vùng nguồn không là kết quả từ thuật toán điền đầy nhiễu âm,

trong đó bộ điền đầy nhiễu âm (604) được tạo cấu hình để nhận biết các vị trí nhiễu âm sử dụng vectơ nhận biết (706) có các mục nhập cho các vị trí phổ chỉ trong vùng phổ nguồn, hoặc các mục nhập chỉ các vị trí phổ trong vùng phổ nguồn và trong vùng phổ đích,

trong đó bộ điền đầy nhiễu âm (604) được tạo cấu hình để tính toán thông tin năng lượng về các giá trị nhiễu âm được biểu thị bởi vectơ nhận biết (706),

trong đó bộ điền đầy nhiễu âm (604) được tạo cấu hình để tính toán thông tin năng lượng về các giá trị ngẫu nhiên được chèn được dự định cho vùng đích,

trong đó bộ điền đầy nhiễu âm được tạo cấu hình để tính toán thừa số khuếch đại để định tỉ lệ các giá trị ngẫu nhiên được chèn được dự định cho vùng đích sử dụng thông tin năng lượng về các giá trị nhiễu âm được biểu thị bởi vectơ nhận biết (706) và sử dụng thông tin năng lượng về các giá trị ngẫu nhiên được chèn được dự định cho vùng đích, và

trong đó bộ điền đầy nhiễu âm được tạo cấu hình để áp dụng thừa số khuếch đại cho các giá trị ngẫu nhiên được chèn được dự định cho vùng đích.

2. Thiết bị theo điểm 1,

trong đó tín hiệu đầu vào là tín hiệu được mã hóa bao gồm các tham số điền đầy nhiễu âm cho vùng phổ nguồn của tín hiệu đầu vào,

trong đó bộ điền đầy nhiễu âm được tạo cấu hình để tạo ra các giá trị nhiễu âm thứ nhất sử dụng các tham số điền đầy nhiễu âm và để tạo ra các giá trị nhiễu âm thứ hai sử dụng thông tin năng lượng về các giá trị nhiễu âm thứ nhất.

3. Thiết bị theo điểm bất kì trong số các điểm nêu trên,

trong đó bộ điền đầy nhiễu âm (604) được tạo cấu hình để tạo ra giá trị nhiễu âm thứ hai tiếp theo thuật toán của bộ ánh xạ (602) hoặc để tạo ra (604) các giá trị nhiễu âm thứ nhất và thứ hai tiếp theo thuật toán của bộ ánh xạ (602).

4. Thiết bị theo điểm bất kì trong số các điểm nêu trên,

trong đó bộ ánh xạ (602) được tạo cấu hình để ánh xạ vùng phổ nguồn thành vùng đích, và

trong đó bộ điền đầy nhiễu âm (604) được tạo cấu hình để thực hiện sự điền đầy nhiễu âm trong các vùng phổ bằng cách tạo ra các giá trị nhiễu âm thứ nhất sử dụng sự điền đầy nhiễu âm và các tham số điền đầy nhiễu âm được truyền trong tín hiệu đầu vào như thông tin phụ, và để thực hiện sự điền đầy nhiễu âm trong vùng đích để tạo ra các giá trị phổ thứ hai sử dụng thông tin năng lượng trên các giá trị nhiễu âm thứ nhất.

5. Thiết bị theo điểm bất kì trong số các điểm nêu trên, thiết bị này còn bao gồm:

bộ điều chỉnh đường bao phổ để điều chỉnh (1202) các giá trị nhiễu âm thứ hai trong vùng phổ nâng cao sử dụng thông tin đường bao phổ nằm trong tín hiệu đầu vào như thông tin phụ.

6. Thiết bị theo điểm bất kì trong số các điểm nêu trên,

trong đó bộ điền đầy nhiễu âm (604) được tạo cấu hình để chỉ sử dụng thông tin phụ của tín hiệu đầu vào để nhận biết các vị trí phổ để điền đầy nhiễu âm, hoặc

trong đó bộ điền đầy nhiễu âm (604) được tạo cấu hình để phân tích đặc tính thời gian hoặc phổ của tín hiệu đầu vào có hoặc không có các giá trị phổ trong vùng điền đầy nhiễu âm để nhận biết các vị trí phổ để điền đầy nhiễu âm.

7. Thiết bị theo điểm bất kì trong số các điểm nêu trên,

trong đó bộ ánh xạ (602) được tạo cấu hình để thực hiện thuật toán điền đầy khoảng trống để tạo ra vùng phổ đích, thiết bị này bao gồm:

bộ giải mã âm thanh miền phổ (112) để tạo ra sự biểu diễn được giải mã thứ nhất của tập hợp thứ nhất gồm các phần phổ thứ nhất, sự biểu diễn được giải mã có độ phân giải phổ thứ nhất;

bộ giải mã theo tham số (114) để tạo ra sự biểu diễn được giải mã thứ hai của tập hợp thứ hai gồm các phần phổ thứ hai có độ phân giải phổ thứ hai thấp hơn độ phân giải phổ thứ nhất;

bộ tái tạo tần số (116) để tái tạo phần phổ thứ hai được khôi phục có độ phân giải phổ thứ nhất sử dụng phần phổ thứ nhất và thông tin đường bao phổ cho phần phổ thứ hai; và

bộ chuyển đổi phổ thời gian (118) để chuyển đổi sự biểu diễn được giải mã thứ nhất trong phần phổ thứ hai được khôi phục thành sự biểu diễn thời gian,

trong đó bộ ánh xạ (602) và bộ điền đầy nhiễu âm (604) ít nhất nằm một phần trong bộ tái tạo tần số (116).

8. Thiết bị theo điểm 7,

trong đó bộ giải mã âm thanh miền phổ được tạo cấu hình để xuất ra chuỗi khung được giải mã gồm các giá trị phổ, khung được giải mã là sự biểu diễn được giải mã thứ nhất, trong đó khung bao gồm các giá trị phổ cho tập hợp thứ nhất gồm các phần phổ và các chỉ số bằng 0 cho tập hợp thứ hai gồm các phần phổ thứ hai,

trong đó thiết bị giải mã còn bao gồm bộ tổ hợp (208) để tổ hợp các giá trị phổ được tạo ra bởi bộ tái tạo tần số cho tập hợp thứ hai gồm các phần phổ thứ hai và các giá trị phổ của tập hợp thứ nhất gồm các phần phổ thứ nhất trong băng khôi phục để thu được khung phổ được khôi phục bao gồm các giá trị phổ cho tập hợp thứ nhất gồm các phần phổ thứ nhất và tập hợp thứ hai gồm các phần phổ thứ hai; và

trong đó bộ chuyển đổi phổ-thời gian (118) được tạo cấu hình để chuyển đổi khung phổ được khôi phục thành sự biểu diễn thời gian.

9. Thiết bị theo điểm bất kì trong số các điểm nêu trên, thiết bị còn bao gồm:

đối với mỗi phạm vi tần số đích, sự nhận biết vùng nguồn, và

trong đó bộ ánh xạ (602) được tạo cấu hình để lựa chọn vùng nguồn sử dụng sự nhận biết vùng nguồn và để ánh xạ vùng nguồn được lựa chọn đến vùng đích.

10. Phương pháp tạo ra tín hiệu âm thanh nâng cao từ tín hiệu đầu vào (600), trong đó tín hiệu âm thanh nâng cao có các giá trị phổ cho vùng phổ nâng cao, các giá trị phổ cho vùng phổ nâng cao không nằm trong tín hiệu đầu vào (600), phương pháp bao gồm:

ánh xạ (602) vùng phổ nguồn của tín hiệu đầu vào thành vùng đích trong vùng phổ nâng cao; và

tạo ra (604) các giá trị nhiều âm thứ nhất cho vùng điền đầy nhiều âm (302) trong vùng phổ nguồn của tín hiệu đầu vào và tạo ra các giá trị nhiều âm thứ hai cho vùng nhiều âm trong vùng đích, trong đó các giá trị nhiều âm thứ hai được giải tương quan từ các giá trị nhiều âm thứ nhất hoặc để tạo ra các giá trị nhiều âm thứ hai cho vùng nhiều âm trong vùng đích, trong đó các giá trị nhiều âm thứ hai được giải tương quan từ các giá trị nhiều âm thứ nhất trong vùng nguồn, trong đó các giá trị nhiều âm thứ nhất trong vùng nguồn không là kết quả từ thuật toán điền đầy nhiều âm,

trong đó việc tạo ra (604) bao gồm:

nhận biết các vị trí nhiều âm sử dụng vectơ nhận biết (706) có các mục nhập cho các vị trí phổ chỉ trong vùng phổ nguồn, hoặc các mục nhập chỉ các vị trí phổ trong vùng phổ nguồn và trong vùng phổ đích,

tính toán thông tin năng lượng trên các giá trị nhiều âm được biểu thị bởi vectơ nhận biết (706),

tính toán thông tin năng lượng về các giá trị ngẫu nhiên được chèn được dự định cho vùng đích,

tính toán thừa số khuếch đại để định tỉ lệ các giá trị ngẫu nhiên được chèn được dự định cho vùng đích sử dụng thông tin năng lượng về các giá trị nhiều âm được biểu

thị bởi vectơ nhận biết (706) và sử dụng thông tin năng lượng về các giá trị ngẫu nhiên được chèn được dự định cho vùng đích, và

áp dụng thừa số khuếch đại cho các giá trị ngẫu nhiên được chèn được dự định cho vùng đích.

11. Hệ thống xử lý tín hiệu âm thanh, hệ thống bao gồm:

bộ mã hóa để tạo ra tín hiệu được mã hóa; và

thiết bị tạo ra tín hiệu âm thanh nâng cao phù hợp với điểm bất kỳ trong số các điểm từ 1 đến 9, trong đó tín hiệu được mã hóa chịu sự xử lý (700) để tạo ra tín hiệu đầu vào (600) vào trong thiết bị tạo ra tín hiệu âm thanh nâng cao.

12. Phương pháp xử lý tín hiệu âm thanh, phương pháp bao gồm:

tạo ra tín hiệu được mã hóa từ tín hiệu âm thanh; và

phương pháp tạo ra tín hiệu âm thanh nâng cao phù hợp với điểm 10, trong đó tín hiệu được mã hóa chịu sự xử lý được định trước (700) để tạo ra tín hiệu đầu vào (600) vào phương pháp tạo ra tín hiệu âm thanh nâng cao.

13. Vật ghi có thể đọc được bằng máy tính bao gồm chương trình máy tính để thực hiện, khi chạy trên máy tính, phương pháp theo điểm 10.

14. Vật ghi có thể đọc được bằng máy tính bao gồm chương trình máy tính để thực hiện, khi chạy trên máy tính, phương pháp theo điểm 12.

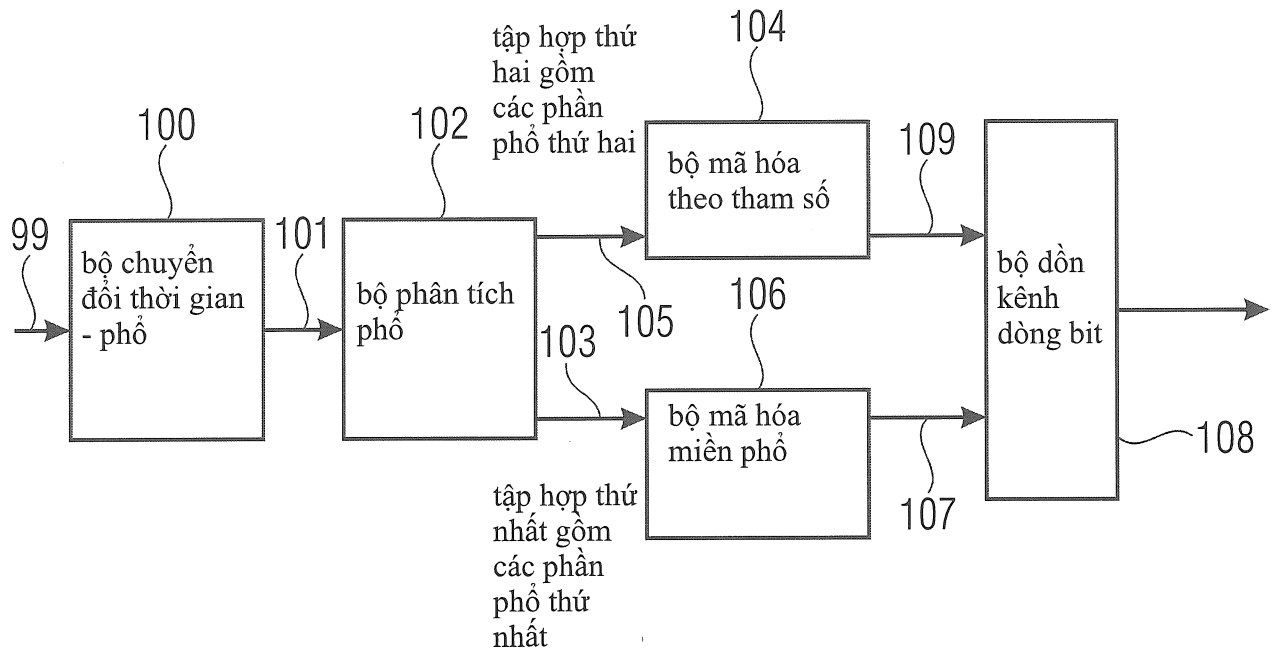


FIG 1A

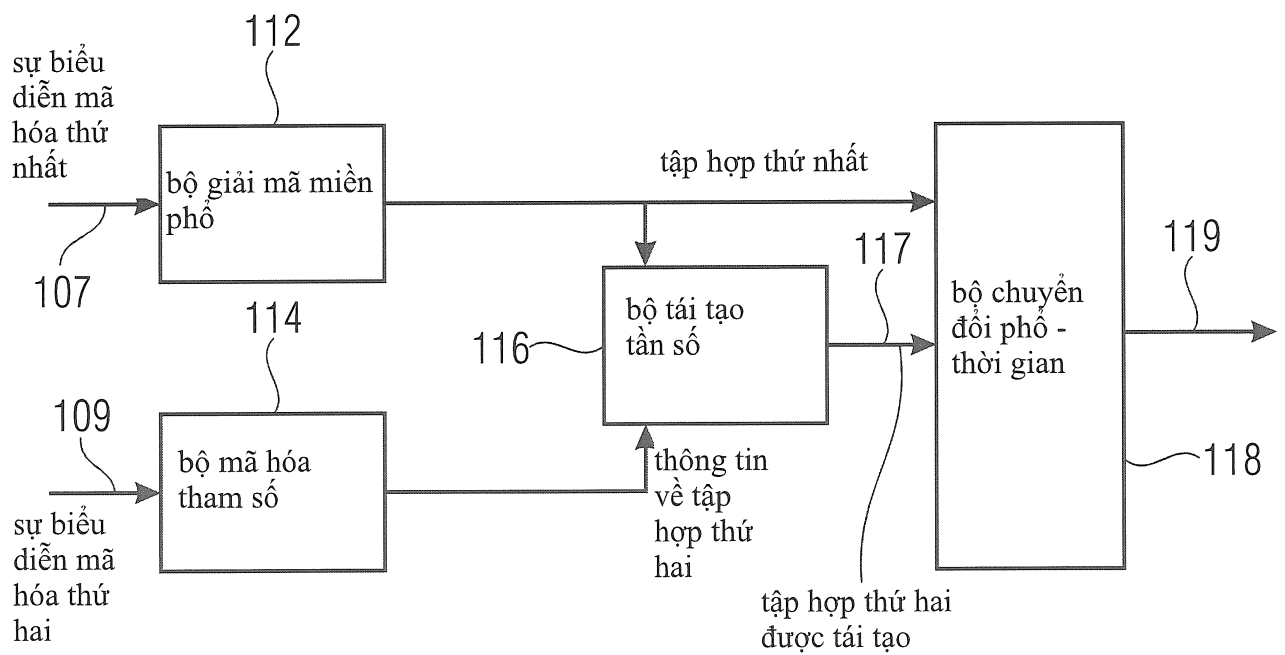


FIG 1B

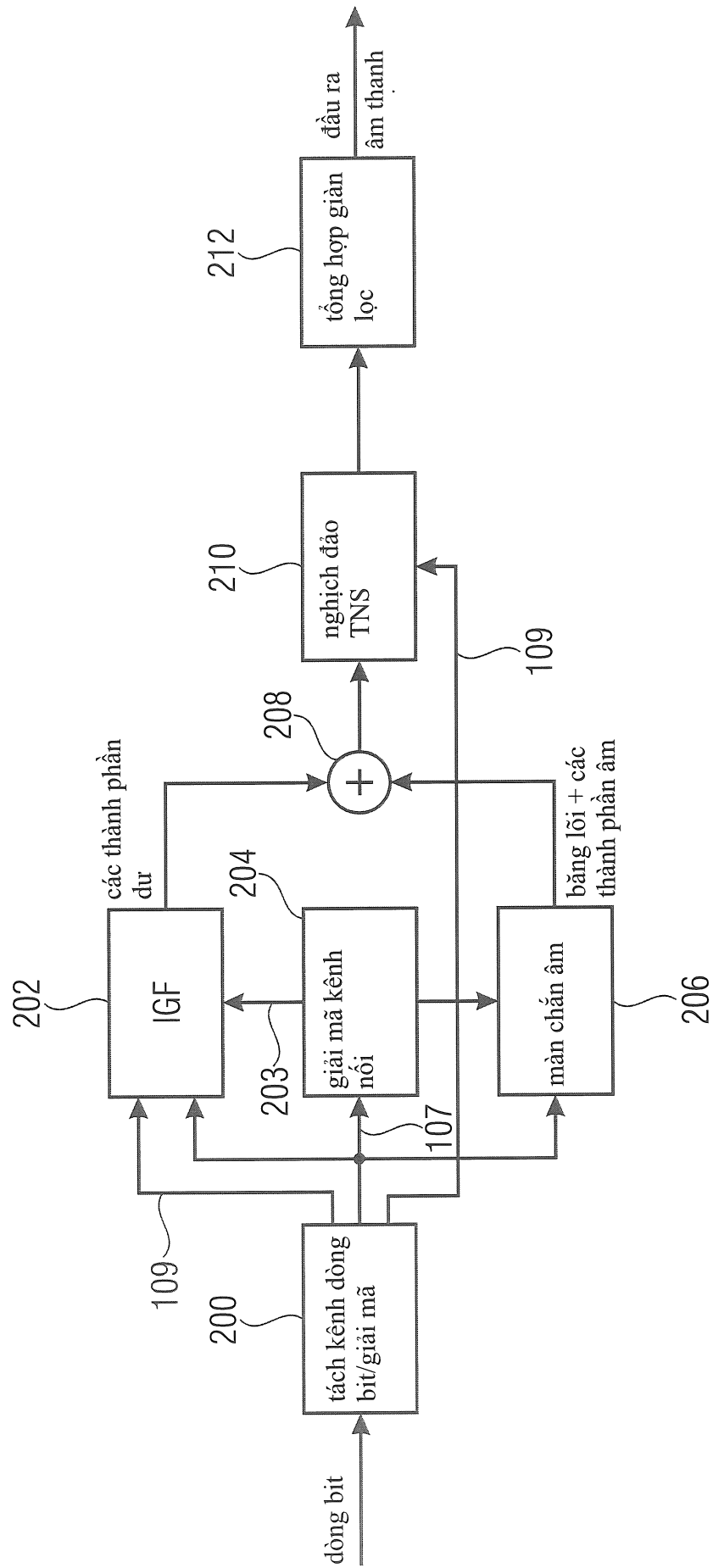


FIG 2A

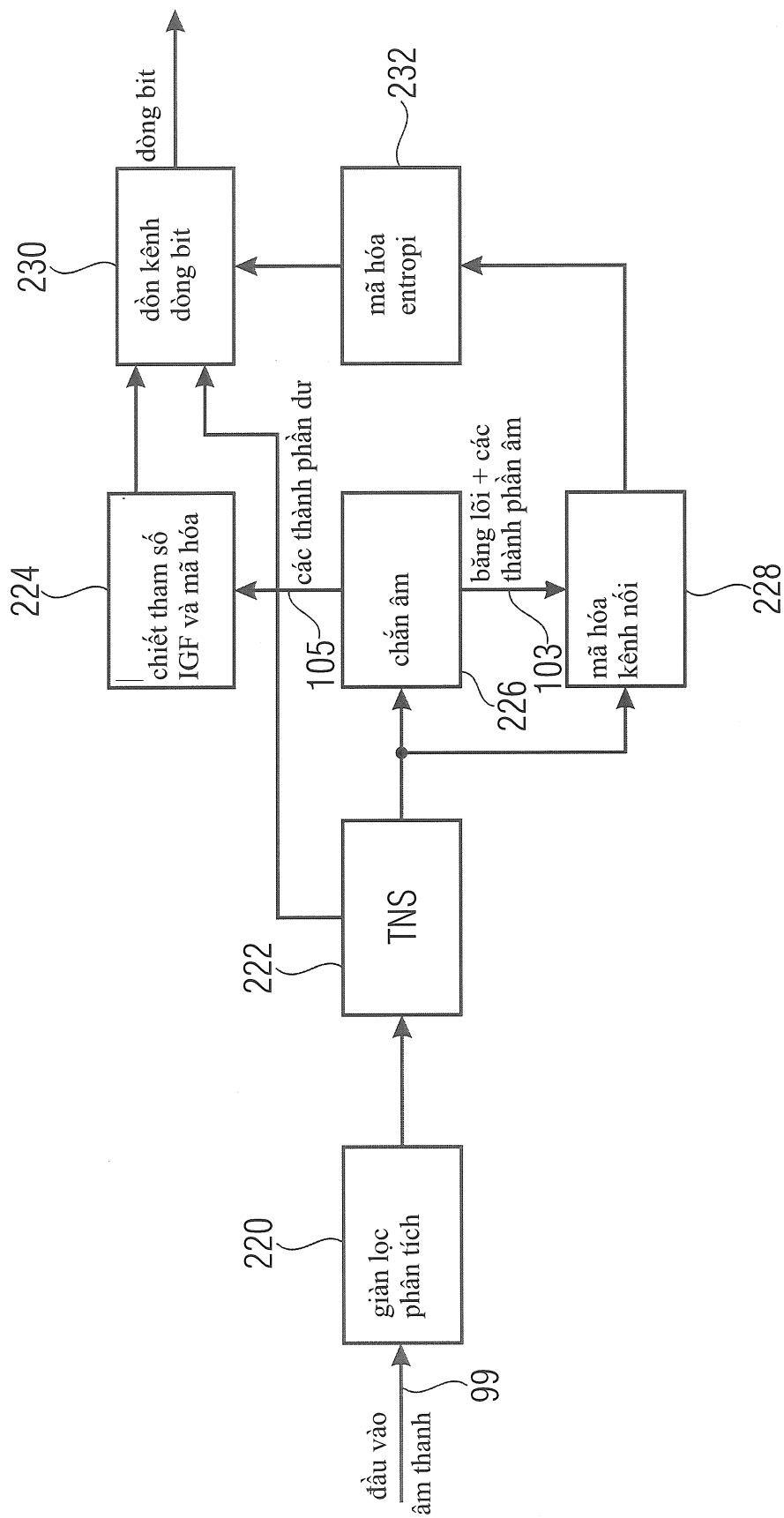


FIG 2B

- độ phân giải thứ nhất (độ phân giải cao) cho "đường bao" của tập hợp thứ nhất (mã hóa theo vạch);
- độ phân giải thứ hai (độ phân giải thấp) cho "đường bao" của tập hợp thứ hai (thừa số định tỉ lệ mỗi SCB);

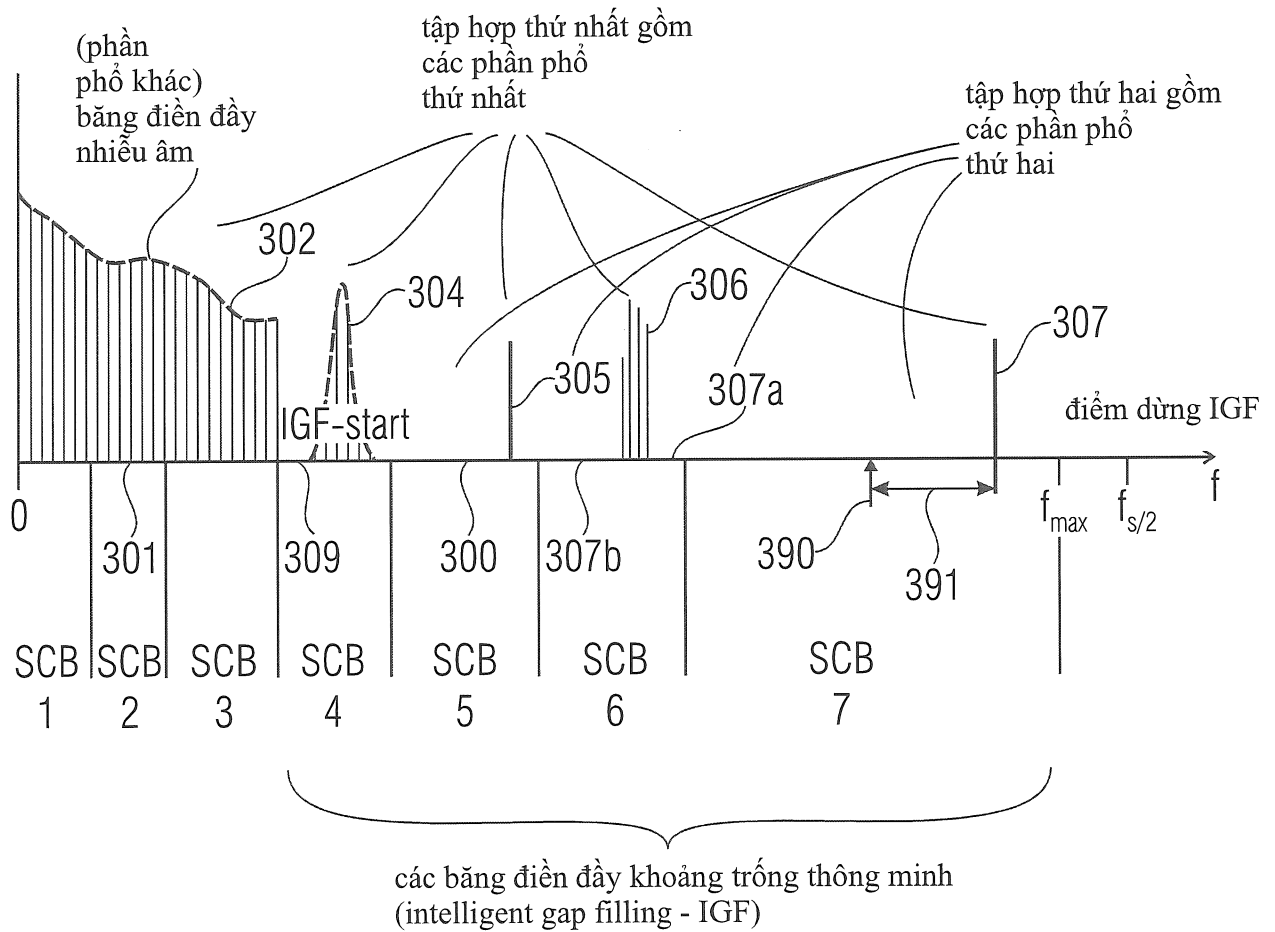


FIG 3A

SCB1	SCB2	SCB3	SCB4	SCB5	SCB6	SCB7
SF1	SF2	SF3	SF4	SF5	SF6	SF7
			E ₁	E ₂	E ₃	E ₄
	NF ₂					

308 310 312

FIG 3B

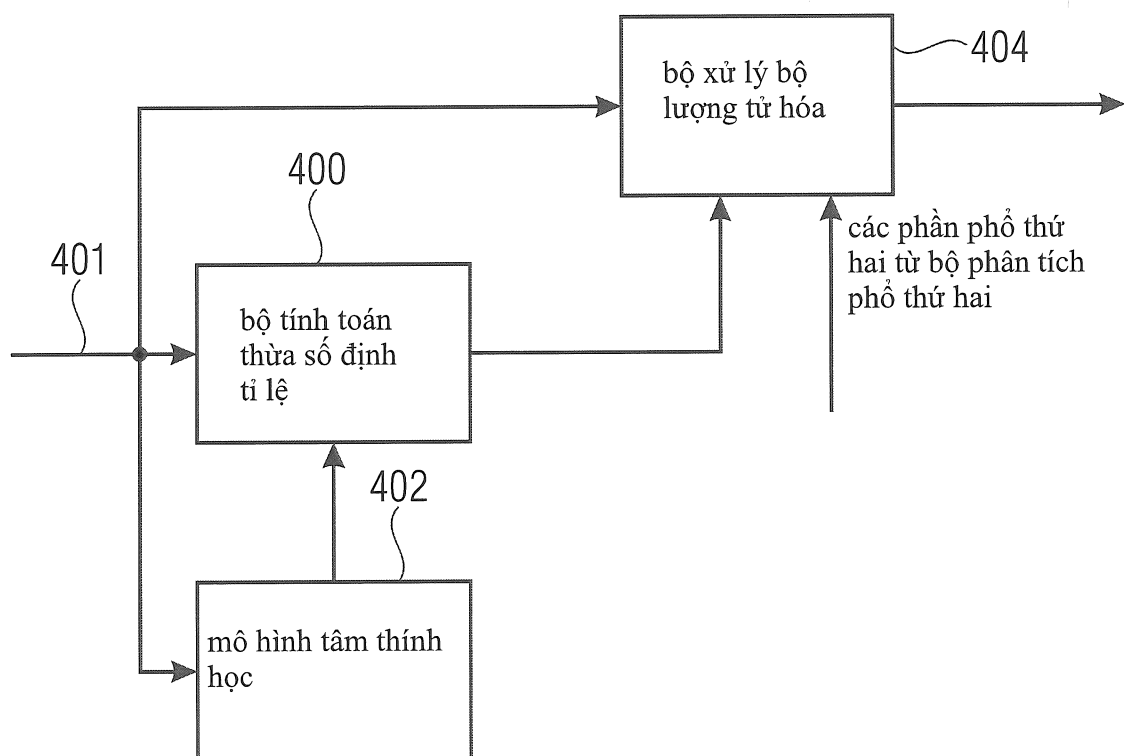


FIG 4A

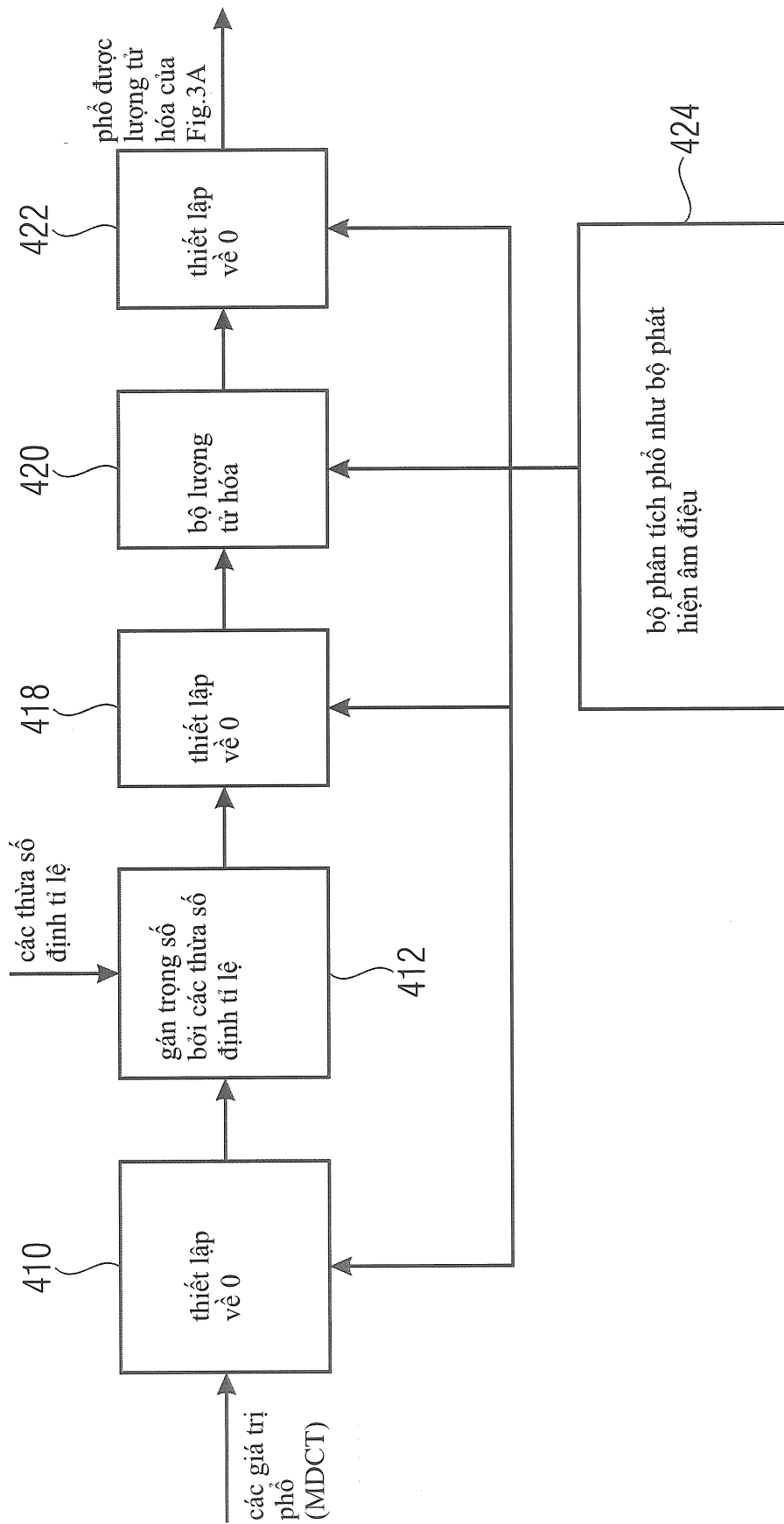


FIG 4B
(BỘ XỬ LÝ LƯỢNG TỬ HÓA)

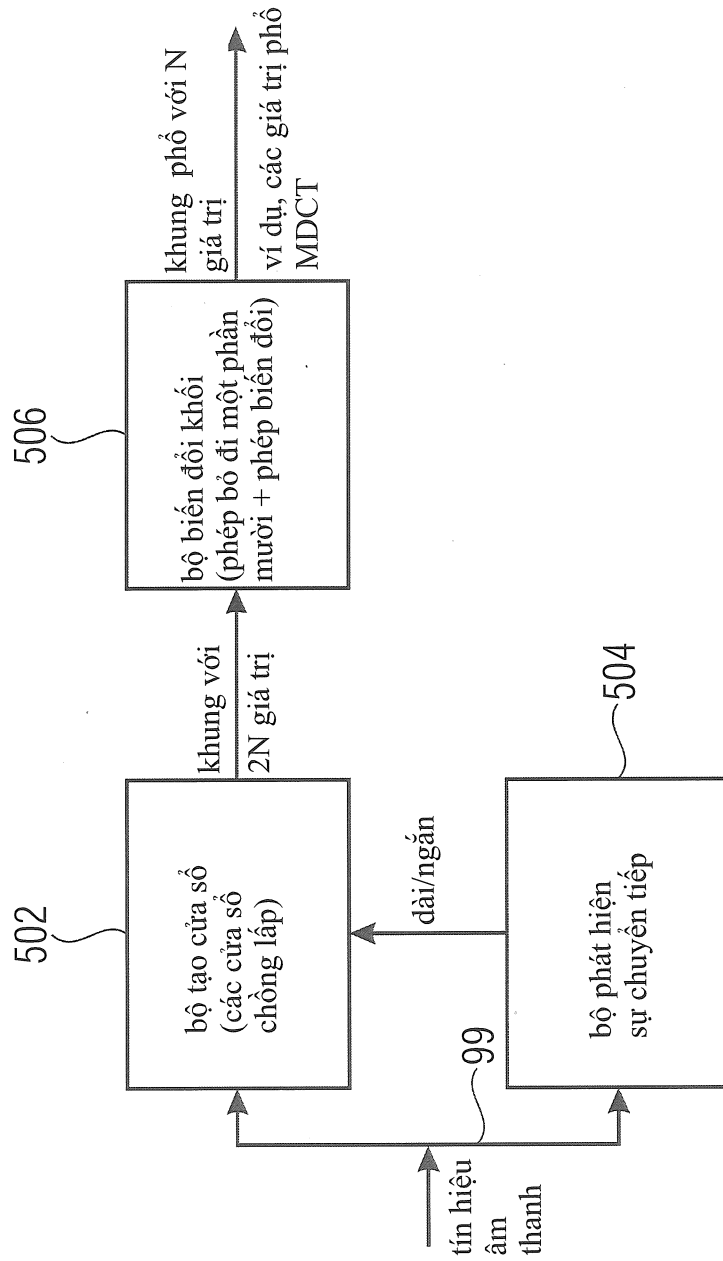


FIG 5A
(CÁC PHẦN PHỎ KHÁC)

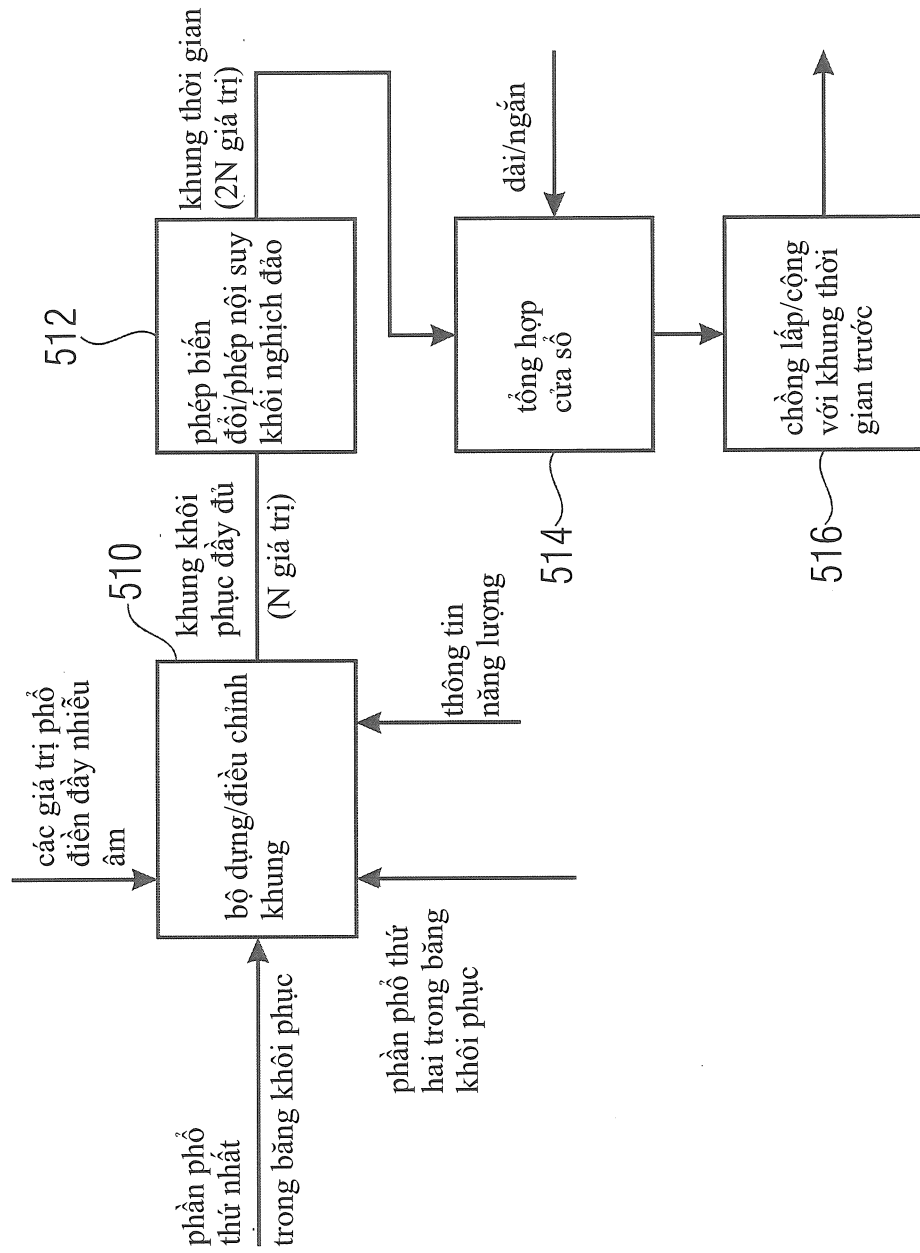


FIG 5B

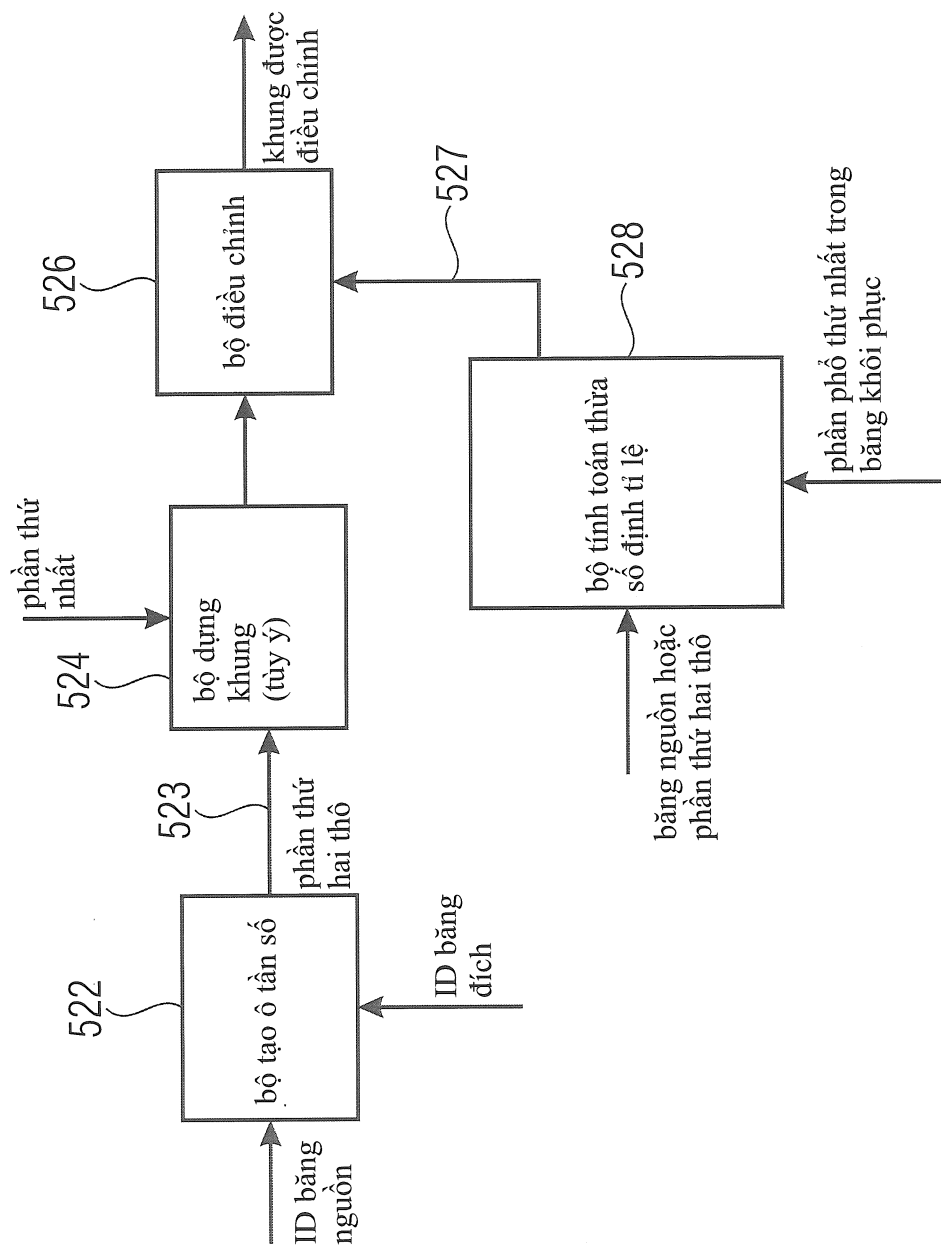


FIG 5C

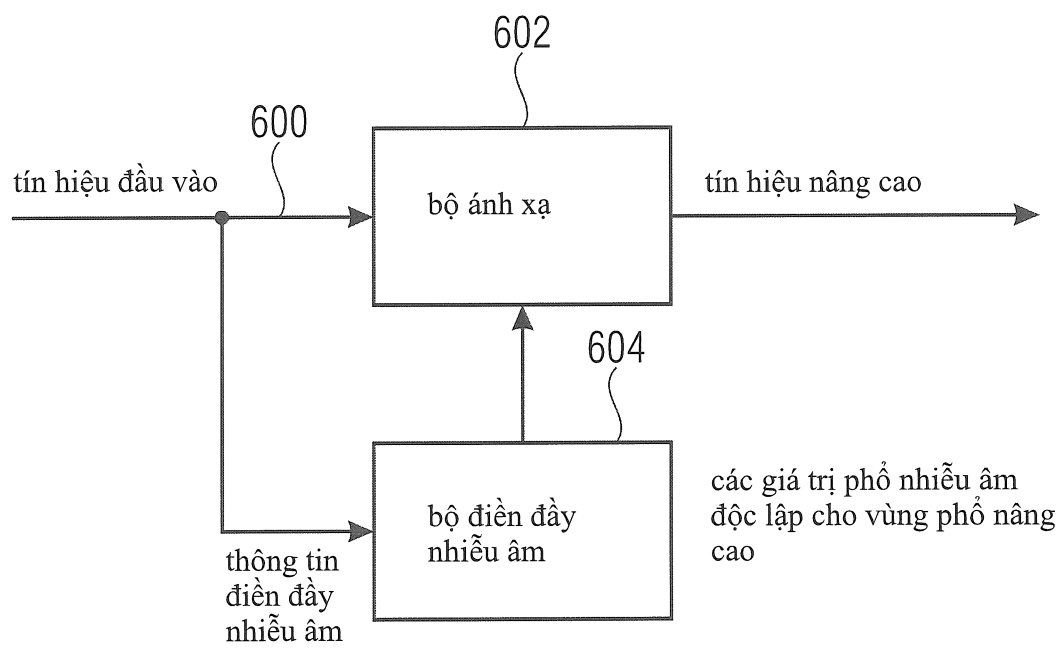


FIG 6

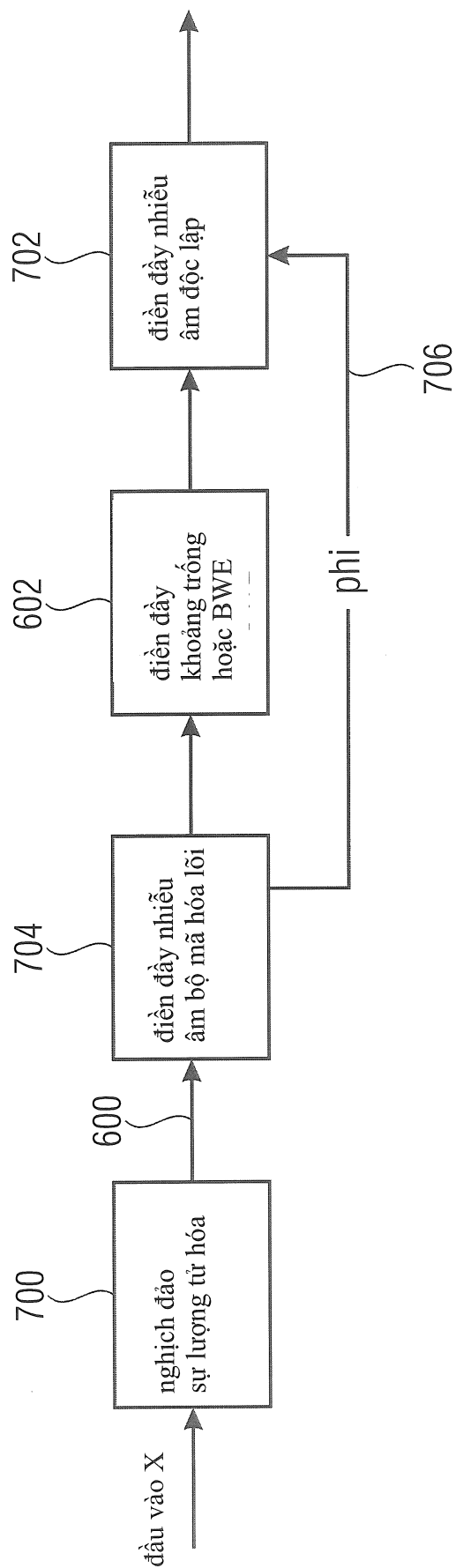


FIG 7

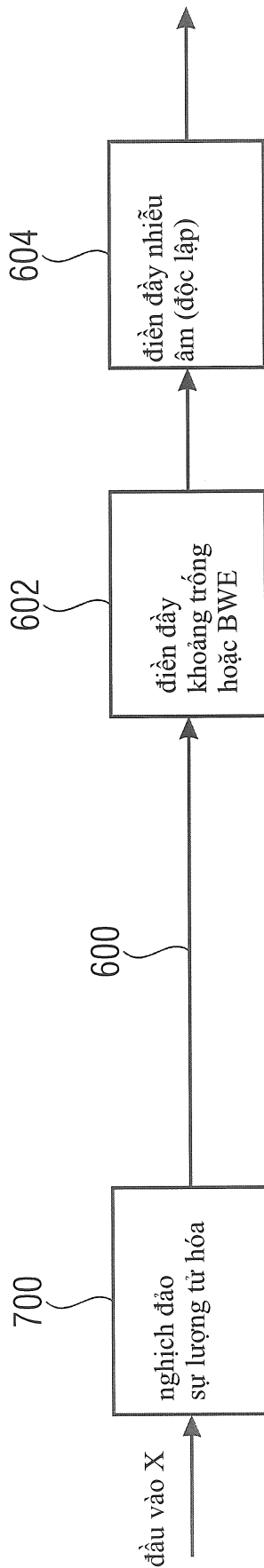


FIG 8

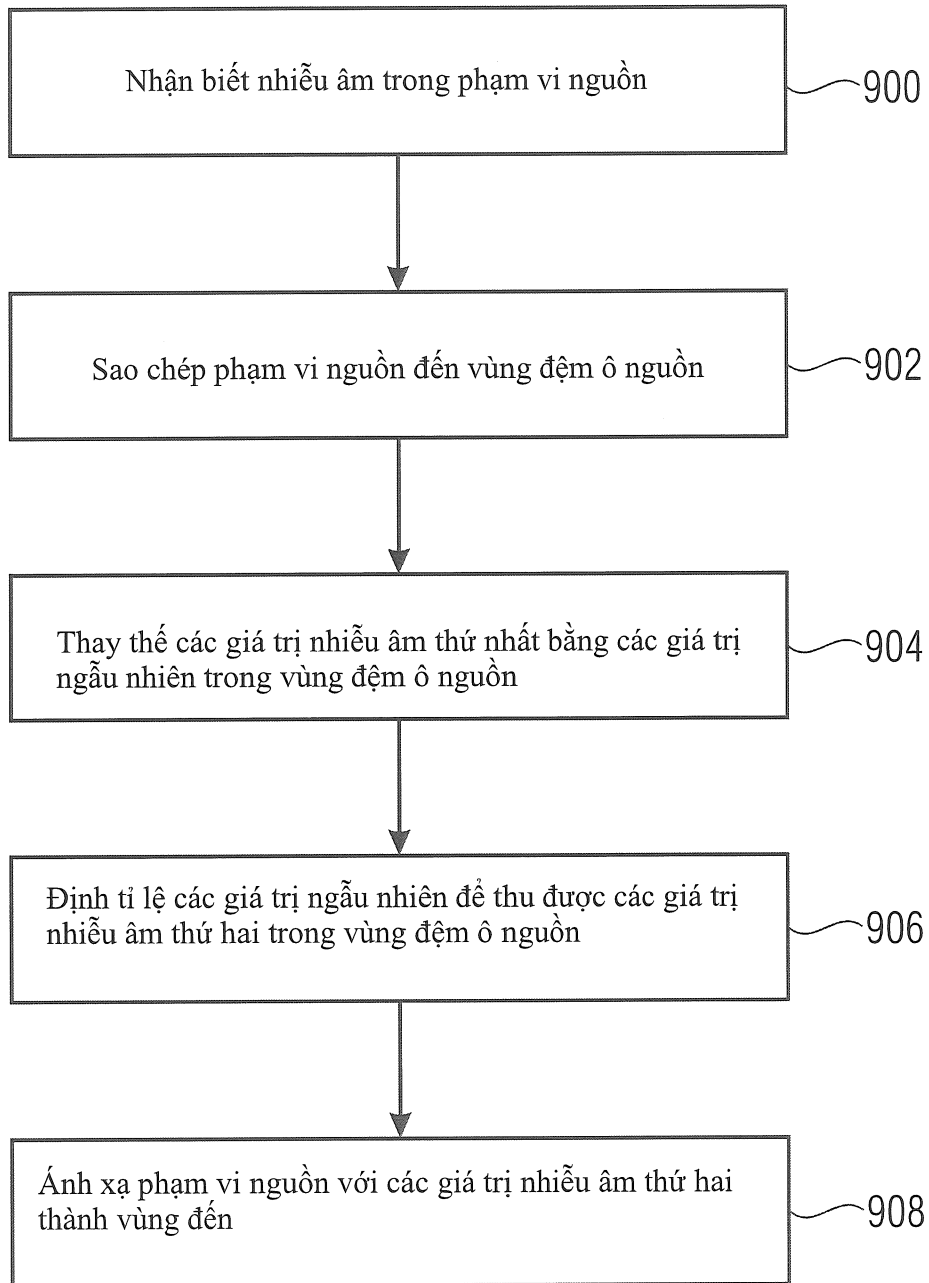


FIG 9

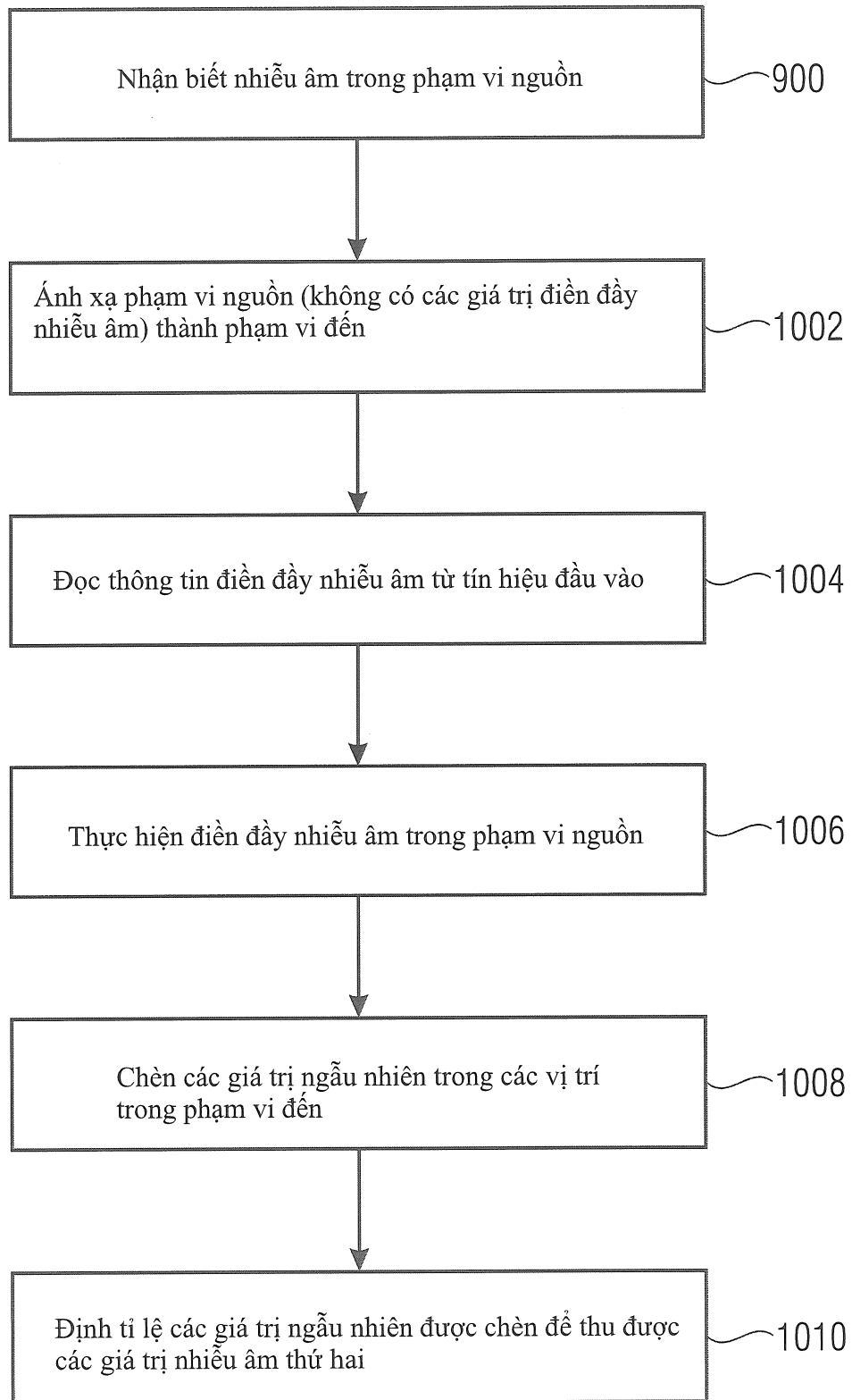


FIG 10

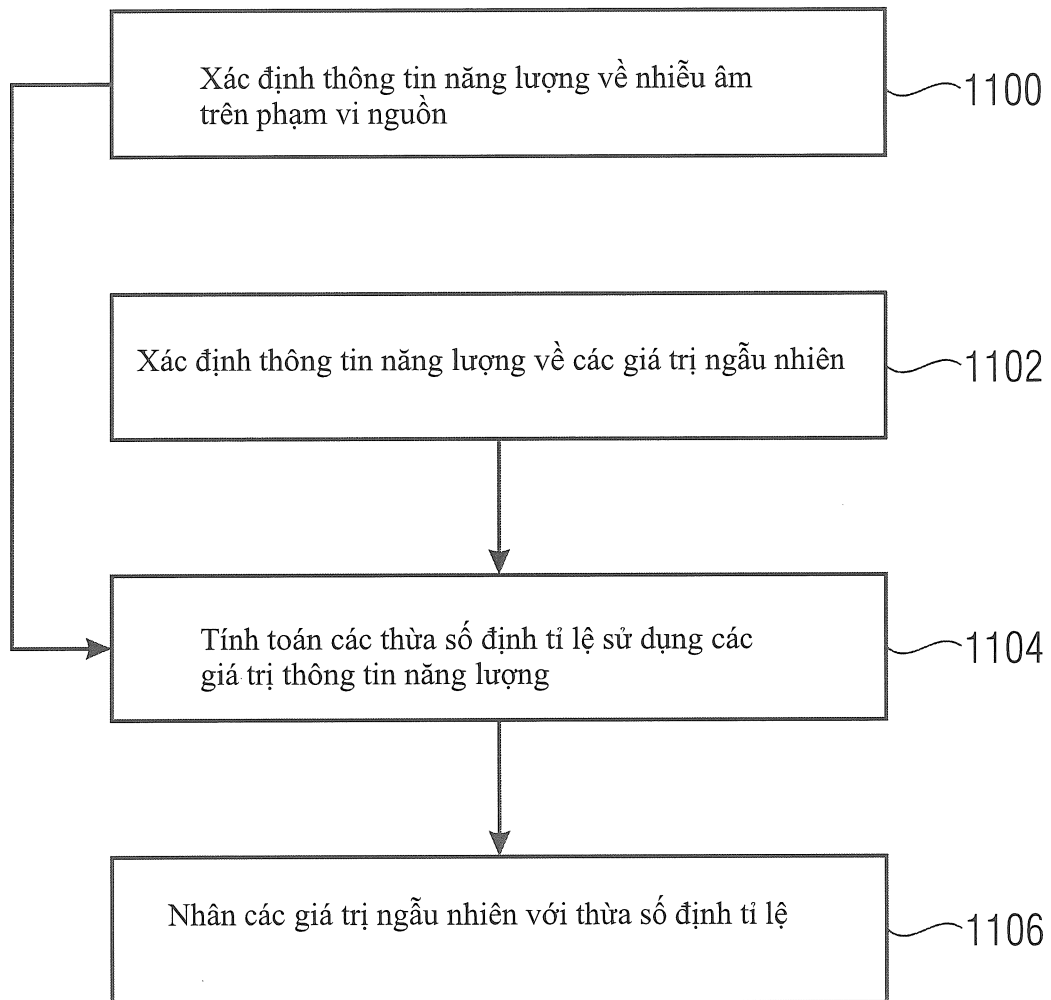


FIG 11

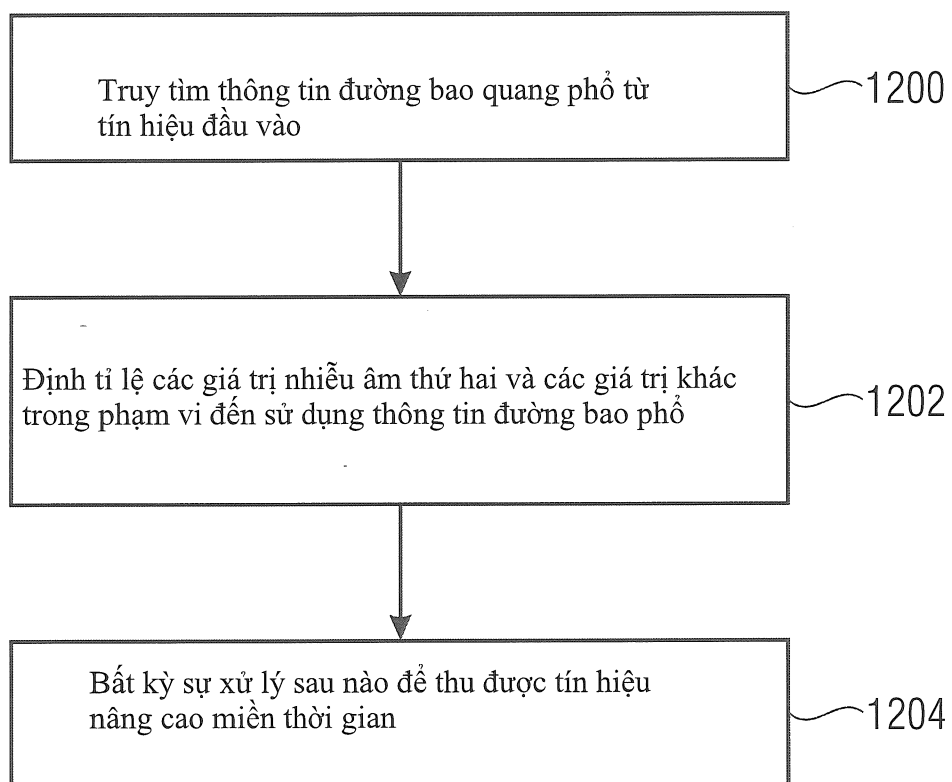


FIG 12

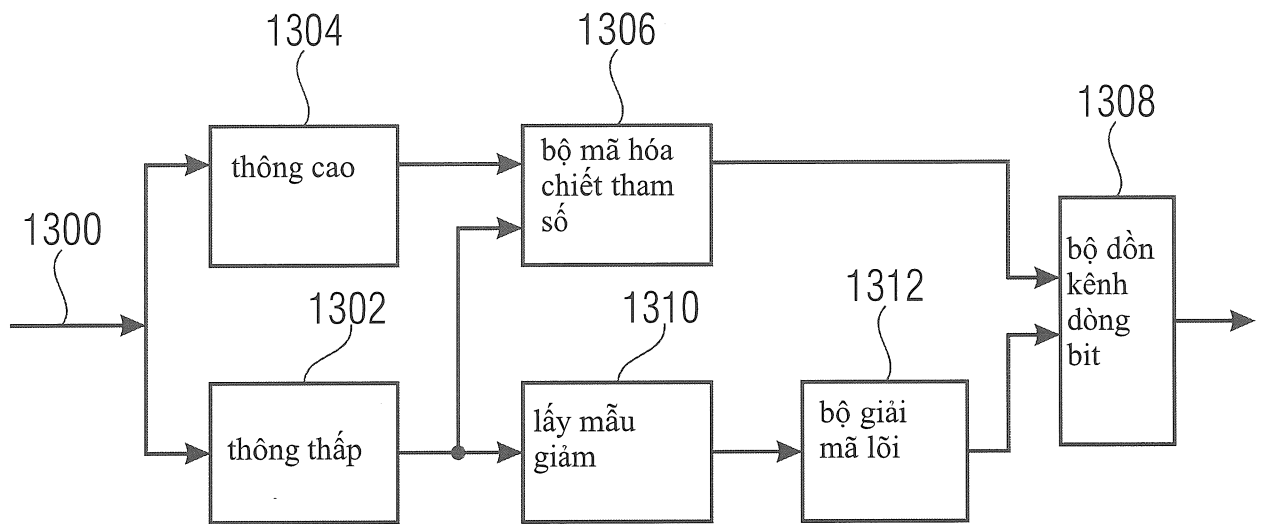


FIG 13A

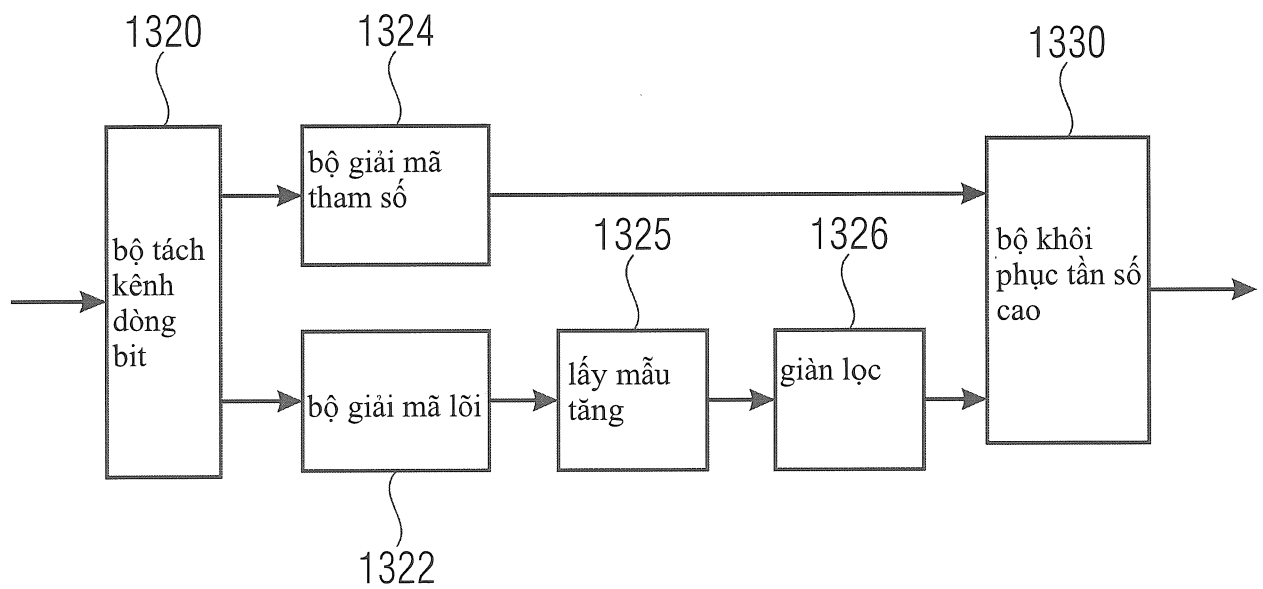


FIG 13B