



(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)



1-0026680

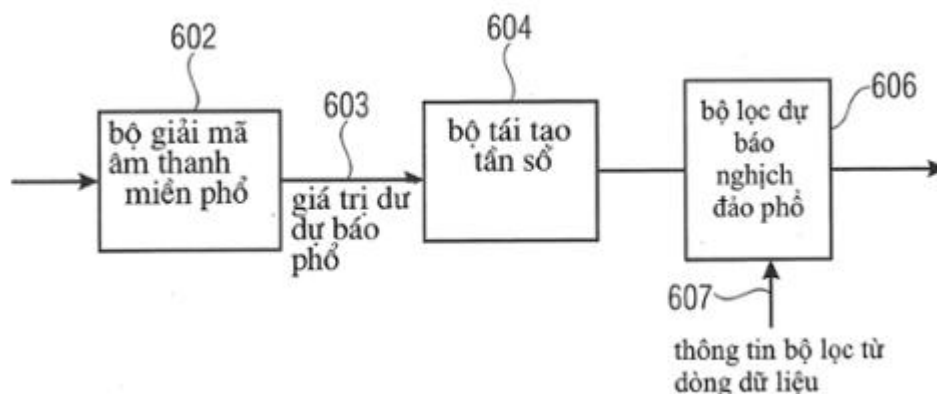
(51)⁷ G10L 19/03; G10L 21/0388; G10L
19/02; G10L 19/028

(13) B

- (21) 1-2015-01191 (22) 15/07/2014
(86) PCT/EP2014/065123 15/07/2014 (87) WO 2015/010954 29/01/2015
(30) 13177346.7 22/07/2013 EP; 13177348.3 22/07/2013 EP; 13177350.9 22/07/2013 EP;
13177353.3 22/07/2013 EP; 13189358.8 18/10/2013 EP
(45) 25/12/2020 393 (43) 25/01/2016 334A
(73) Fraunhofer-Gesellschaft zur Foerderung der angewandten Forschung e.V. (DE)
Hansastrasse 27c, 80686 Muenchen, Germany
(72) DISCH, Sascha (DE); NAGEL, Frederik (DE); GEIGER, Ralf (DE);
THOSKAHNA, Balaji Nagendran (DE); SCHMIDT, Konstantin (DE); BAYER,
Stefan (DE); NEUKAM, Christian (DE); EDLER, Bernd (DE); HELMRICH,
Christian (DE).
(74) CÔNG TY LUẬT TRÁCH NHIỆM HỮU HẠN AMBYS HÀ NỘI (AMBYS
HANOI)

(54) THIẾT BỊ VÀ PHƯƠNG PHÁP MÃ HÓA VÀ GIẢI MÃ TÍN HIỆU ÂM THANH
ĐƯỢC MÃ HÓA SỬ DỤNG VIỆC TẠO HÌNH NHIỀU ÂM/VÁ LỖI THEO THỜI
GIAN

(57) Thiết bị để giải mã tín hiệu âm thanh được mã hóa, bao gồm: bộ giải mã âm thanh miền phổ (602) để tạo ra phép biểu diễn được giải mã thứ nhất của tập hợp thứ nhất của các phần phổ thứ nhất là các giá trị dự dự báo phổ; bộ tái tạo tần số (604) để tạo ra phần phổ thứ hai được khôi phục sử dụng phần phổ thứ nhất của tập hợp thứ nhất của các phần phổ thứ nhất, trong đó phần phổ thứ hai được khôi phục bao gồm các giá trị dự dự báo phổ; và bộ lọc dự báo nghịch đảo (606) để thực hiện việc dự báo nghịch đảo vượt quá tần số sử dụng các giá trị dự phổ cho tập hợp thứ nhất của các phần phổ thứ nhất và phần phổ thứ hai được khôi phục sử dụng thông tin bộ lọc dự báo (607) được chứa trong tín hiệu âm thanh được mã hóa.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến mã hóa/giải mã âm thanh và, cụ thể, đề cập đến mã hóa âm thanh sử dụng việc điền đầy khoảng trống thông minh (Intelligent Gap Filling - IGF).

Tình trạng kỹ thuật của sáng chế

Mã hóa âm thanh là miền nén tín hiệu mà thực hiện với việc khai thác phần dư và không thích hợp trong tín hiệu âm thanh sử dụng kiến thức cảm quan. Hiện nay, các bộ mã hóa-giải mã âm thanh điển hình cần khoảng 60 kilobit trên giây (Kilobit per second – kbps) trên một kênh để mã hóa rõ ràng cảm giác của hầu hết bất kỳ dạng tín hiệu âm thanh nào. Các bộ mã hóa-giải mã mới hơn nhằm mục đích làm giảm tốc độ bit mã hóa bằng việc khai thác các đặc điểm giống nhau về phổ trong tín hiệu sử dụng các kỹ thuật như mở rộng băng thông (bandwidth extension - BWE). Sơ đồ BWE sử dụng tham số tốc độ bit thấp để biểu diễn các thành phần tần số cao (hight frequency - HF) của tín hiệu âm thanh. Phổ HF được điền đầy với lượng phổ từ vùng tần số thấp (low frequency - LF) và tạo hình phổ, độ nghiêng và tính liên tục theo thời gian được điều chỉnh để duy trì được màu và âm sắc của tín hiệu gốc. Các phương pháp BWE này làm cho các bộ mã hóa-giải mã có khả năng giữ được chất lượng tốt thậm chí tại tốc độ bit thấp khoảng 24kbps/kênh.

Việc lưu trữ và truyền tín hiệu âm thanh là vấn đề thường xuyên đề hạn chế tốc độ bit chặt chẽ. Trước đây, bộ mã hóa-giải mã bắt buộc phải giảm đáng kể băng thông âm thanh được truyền khi mà chỉ tốc độ bit rất thấp có thể có được.

Bộ mã hóa-giải mã âm thanh hiện đại ngày nay có khả năng mã hóa được tín hiệu dải rộng bằng việc sử dụng phương pháp mở rộng băng thông BWE [1]. Những thuật toán này dựa trên việc biểu diễn theo tham số của lượng tần số cao (HF) - mà được tạo ra từ phần tần số thấp được mã hóa dạng sóng (LF) của tín hiệu được giải mã bởi việc chuyển thành vùng phổ HF (việc vá lỗi) và ứng dụng sau xử lý dẫn theo tham số. Trong sơ đồ BWE, việc khôi phục vùng phổ HF ở trên đã được gọi là tần số giao nhau

thường được dựa trên việc vá lỗi phổ. Điển hình, vùng HF gồm nhiều bản vá liền kề và mỗi bản vá này có nguồn gốc từ các vùng dải giữa (band-pass - BP) của phổ LF phía dưới tần số giao nhau đã có. Các hệ thống thuộc tình trạng kỹ thuật thực hiện hiệu quả việc vá lỗi trong biểu diễn giàn lọc, ví dụ, Giàn lọc gương cầu phương (Quadrature Mirror Filterbank - QMF), bằng việc sao chép bộ hệ số dải phụ liền kề từ nguồn tới vùng đích.

Kỹ thuật khác được tìm thấy trong mã hóa-giải mã âm thanh hiện nay nhằm làm tăng hiệu quả nén và bằng cách đó cho phép mở rộng băng thông âm thanh tại tốc độ bit thấp là việc thay thế tổng hợp dẫn theo tham số của các phần thích hợp của phổ âm thanh. Ví dụ, các phần tín hiệu nhiễu âm tương tự của tín hiệu âm thanh gốc có thể được thay thế mà chất lượng âm thanh không bị giảm đáng kể bởi nhiễu âm nhân tạo được tạo ra trong bộ giải mã và được định tỷ lệ bằng các tham số thông tin phụ. Một ví dụ là công cụ Thay thế Nhiễu âm Cảm giác (Perceptual Noise Substitution - PNS) có trong Mã hóa Âm thanh Tiên tiến MPEG-4 (Advanced Audio Coding - AAC) [5].

Một điều nữa là cũng cho phép băng thông âm thanh được mở rộng tại tốc độ bit thấp chính là kỹ thuật điền đầy nhiễu âm có trong Mã hóa Âm thanh và Tiếng nói Hợp nhất (Unified Speech and Audio Coding - USAC) MPEG-D [7]. Khoảng trống phổ (trống không) được phỏng đoán bởi vùng chết của bộ lượng tử hóa do việc lượng tử hóa quá thô, được điền đầy liên tiếp với nhiễu âm nhân tạo trong bộ giải mã và được định tỷ lệ bởi sau xử lý dẫn theo tham số.

Hệ thống thuộc tình trạng kỹ thuật khác được đặt tên là Thay thế Phổ Chính xác (Accurate Spectral Replacement - ASR) [2-4]. Ngoài ra để mã hóa-giải mã dạng sóng, ASR sử dụng mức tổng hợp tín hiệu riêng cho phép phục hồi lại các phần hình sin quan trọng theo cảm giác của tín hiệu tại bộ giải mã. Ngoài ra, hệ thống được mô tả trong tài liệu [5] dựa vào mô hình hình sin trong vùng HF của bộ mã hóa-giải mã dạng sóng để cho phép băng thông âm thanh được mở rộng có chất lượng cảm giác khá tốt được ở tốc độ bit thấp. Tất cả các phương pháp này liên quan đến việc biến đổi dữ liệu vào trong miền thứ hai ngoại trừ Biến đổi Cosin Rời rạc được biến đổi (Modified Discrete Cosine Transform - MDCT) và ngoài ra tăng phân tích/tổng hợp khá phức tạp cho việc biểu diễn thành phần hình sin HF.

Fig.13a minh họa biểu đồ dưới dạng sơ đồ của bộ mã hóa âm thanh cho kỹ thuật mở rộng băng thông, 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 ở đường 1300 là đầu vào vào hệ thống bộ lọc bao gồm thông thấp 1320 và thông cao 1340. Tín hiệu đầu ra được lọc bởi bộ lọc thông cao 1340 là đầu vào vào trong bộ phân tách/bộ mã hóa theo tham số 1306. Bộ phân tách/bộ mã hóa theo tham số 1306 được tạo cấu hình để tính toán và mã hóa theo tham số ví dụ như tham số đường bao phổ, tham số bổ sung nhiễu âm, tham số sóng hài khuyết, hoặc tham số lọc nghịch đảo. Tham số đã phân tách là đầu vào vào bộ trộn kênh dòng bit 1308. Tín hiệu đầu ra thông thấp là đầu vào vào bộ xử lý điện hình chứa chức năng của bộ lấy mẫu giảm 1310 và bộ mã hóa lõi 1312. Thông thấp 1302 hạn chế băng thông để được mã hóa thành băng thông nhỏ hơn đáng kể so với băng thông xuất hiện 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 là toàn bộ chức năng xuất hiện trong bộ mã hóa lõi chỉ hoạt động trên tín hiệu với băng thông bị giảm. Khi đó, ví dụ, băng thông của tín hiệu âm thanh trên đường 1300 là 20kHz và khi bộ lọc thông thấp 1320 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 8kHz, đó là sự giảm bớt đáng kể để tốc độ lấy mẫu cần thiết cho tín hiệu âm thanh 1300 tối thiểu là 40kHz.

Fig.13b minh họa biểu đồ dưới dạng sơ đồ bộ giải mã mở rộng băng thông tương ứng. Bộ giải mã bao gồm bộ trộn kênh dòng bit 1320. Bộ tách kênh dòng bit 1320 phân tách tín hiệu đầu ra cho bộ giải mã lõi 1322 và tín hiệu đầu vào cho bộ giải mã theo tham số 1324. Trong ví dụ trên, tín hiệu đầu ra bộ giải mã lõi có tốc độ lấy mẫu 8kHz, do đó, trong khi băng thông là 4kHz, để hoàn thành việc khôi phục băng thông, tín hiệu đầu ra của bộ khôi phục tần số cao 1330 phải là 20kHz bắt buộc tốc độ tải mẫu tối thiểu là 40kHz. Để thực hiện tính khả thi này, bộ xử lý giải mã cần có chức năng của bộ lấy mẫu tăng 1325 và giàn lọc 1326. Bộ khôi phục tần số cao 1330 sau đó nhận đầu ra tín hiệu tần số thấp đã phân tích tần số cung cấp bởi giàn lọc 1326 và khôi phục khoảng tần số được xác định bằng bộ lọc thông cao 1304 của Fig.13a sử dụng biểu diễn tham số của dải tần cao. Bộ khôi phục tần số cao 1330 có chức năng ghép như tái tạo khoảng tần số phía trên sử dụng phạm vi nguồn trong khoảng tần số thấp, việc điều

chỉnh đường bao phổ, chức năng bổ sung nhiều âm và chức năng đưa sóng hài khuyết trong khoảng tần số phía trên, nếu được đặt vào và tính toán trong bộ mã hóa của Fig.13a, hoạt động lọc nghịch đảo để tính toán với thực tế là khoảng tần số cao hơn là đặc trưng không như âm điệu khoảng tần số thấp hơn. Trong HE-AAC, sóng hài khuyết được tái tổng hợp ở phía bộ giải mã và được đặt chính xác ở giữa dải khôi phục. Do đó, tất cả các đường sóng hài khuyết được xác định trong dải khôi phục nhất định không được đặt tại giá trị tần số tại đó mà chúng được đặt trong tín hiệu gốc. Thay vào đó, những đường sóng hài khuyết này được đặt tại các tần số trong trung tâm của dải nhất định. Do vậy, khi đường sóng hài khuyết trong tín hiệu gốc được đặt rất gần viền dải khôi phục trong tín hiệu gốc, lỗi trong tần số được đưa vào bằng việc đặt đường sóng hài khuyết trong tín hiệu đã được khôi phục tại trung tâm của dải gần tới 50% dải khôi phục riêng, để tham số được tạo ra và được truyền.

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 mà sau đó, một lần nữa, được biến đổi thành miền phổ bởi chức năng 1326 bằng bộ lọc. Việc đưa vào các độ trễ xử lý bổ sung, có thể đưa các thành phần nhân tạo do việc xử lý điện từ của biến đổi thứ nhất từ miền phổ thành miền tần số và một lần nữa biến đổi điển hình thành miền tần số khác, và tất nhiên, điều này cũng cần lượng đáng kể việc tính toán phức tạp và bởi nguồn điện, mà là lợi ích cụ thể khi kỹ thuật 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.

Bộ mã hóa-giải mã âm thanh hiện có thực hiện việc mã hóa âm thanh tốc độ bit thấp sử dụng BWE như phần đầy đủ của sơ đồ mã hóa. Tuy nhiên, các kỹ thuật BWE bị hạn chế việc chỉ thay thế lượng tần số cao (HF). Hơn nữa, các kỹ thuật này không cho phép lượng quan trọng cảm giác ở trên có được tần số giao nhau để được mã hóa dạng sóng. Do đó, bộ mã hóa-giải mã âm thanh hiện thời hoặc là giảm chi tiết HF hoặc âm sắc khi BWE được thực hiện, vì không có được sự giống nhau chính xác của sóng hài âm trong tín hiệu trong việc xem xét trong hầu hết các hệ thống.

Thiếu sót khác trong tình trạng hiện có của hệ thống BWE ứng dụng là cần phải biến đổi tín hiệu âm thanh thành miền mới để thực hiện của BWE (ví dụ chuyển từ

miền MDCT sang miền QMF). Điều này dẫn tới sự phức tạp của việc đồng bộ hóa, sự phức tạp tính toán bổ sung và cần bộ nhớ phải tăng.

Đặc biệt, nếu hệ thống mở rộng băng thông được triển khai trong giàn lọc hoặc miền biến đổi thời gian-tần số, chỉ có khả năng hạn chế để kiểm soát hình dạng theo thời gian của tín hiệu mở rộng băng thông. Thông thường, độ chi tiết theo thời gian bị giới hạn bởi kích thước bước nhảy được sử dụng giữa các cửa sổ biến đổi liên kề. Điều này có thể dẫn tới tiếng vang trước hoặc sau tiếng vang không mong muốn trong dải phổ mở rộng băng thông. Để tăng độ chi tiết theo thời gian, có thể sử dụng các kích thước bước nhảy ngắn hơn hoặc các khung mở rộng băng thông ngắn hơn, nhưng điều này dẫn đến chi phí tốc độ bit bởi thực tế là, trong khoảng thời gian nhất định, số lượng các tham số cao hơn, điển hình là tập hợp các tham số nhất định cho từng khung thời gian phải được truyền. Mặt khác, nếu từng khung thời gian riêng được tạo quá rộng, thì tiếng vang trước và sau cụ thể cho các phần chuyển tiếp của tín hiệu âm thanh được tạo ra.

Tài liệu tham khảo

- [1] Dietz, L. Liljeryd, K. Kjörling and O. Kunz, “Spectral Band Replication, a novel approach in audio coding”, tại Hội nghị AES lần thứ 112, Munich, tháng 05/2002.
- [2] Ferreira, D. Sinha, “Accurate Spectral Replacement”, Hội nghị Hiệp hội Kỹ thuật âm thanh, Barcelona, Tây Ban Nha 2005.
- [3] D. Sinha, A. Ferreira¹ and E. Harinarayanan, “A Novel Integrated Audio Bandwidth Extension Toolkit (ABET)”, Hội nghị Hiệp hội Kỹ thuật âm thanh, Paris, Pháp 2006.
- [4] R. Annadana, E. Harinarayanan, A. Ferreira and D. Sinha, “New Results in Low Dữ liệu Rate Speech Coding and Bandwidth Extension”, Hội nghị Hiệp hội Kỹ thuật âm thanh, San Francisco, Mỹ 2006.
- [5] T. Żernicki, M. Bartkowiak, “Audio bandwidth extension by frequency scaling of sinusoidal partials”, Hội nghị Hiệp hội Kỹ thuật âm thanh, San Francisco, Mỹ 2008.

- [6] J. Herre, D. Schulz, Extending the MPEG-4 AAC Codec by Perceptual Noise Substitution, Hội nghị AES lần thứ 104, Amsterdam, 1998, Preprint 4720.
- [7] M. Neuendorf, M. Multus, N. Rettelbach, et al., MPEG Unified Speech and Audio Coding-The ISO/MPEG Standard for High-Efficiency Audio Coding of all Content Types, Hội nghị AES lần thứ 132, Budapest, Hungary, tháng 04/2012.
- [8] McAulay, Robert J., Quatieri, Thomas F. "Speech Analysis/Synthesis Based on a Sinusoidal Representation". IEEE Transactions on Acoustics, Speech, And Signal Processing, Vol 34(4), tháng 08/1986.
- [9] Smith, J.O., Serra, X. "PARSHL: An analysis/synthesis program for non-harmonic sounds based on a sinusoidal representation", Chuyên đề của Hội thảo âm nhạc máy tính quốc tế, 1987.
- [10] Purnhagen, H.; Meine, Nikolaus, "HILN-the MPEG-4 parametric audio coding tools," *Circuits and Systems, 2000. Proceedings. ISCAS 2000 Geneva. The 2000 IEEE International Symposium on* , vol.3, no., pp.201,204 vol.3, 2000
- [11] International Standard ISO/IEC 13818-3, Generic Coding of Moving Pictures and Associated Audio: Audio", Geneva, 1998.
- [12] M. Bosi, K. Brandenburg, S. Quackenbush, L. Fielder, K. Akagiri, H. Fuchs, M. Dietz, J. Herre, G. Davidson, Oikawa: "MPEG-2 Advanced Audio Coding", Hội nghị AES lần thứ 101, Los Angeles 1996
- [13] J. Herre, "Temporal Noise Shaping, Quantization and Coding methods in Perceptual Audio Coding: A Tutorial introduction", Hội nghị quốc tế AES lần thứ 17 về Mã hóa âm thanh chất lượng cao, tháng 08/1999
- [14] J. Herre, "Temporal Noise Shaping, Quantization and Coding methods in Perceptual Audio Coding: A Tutorial introduction", Hội nghị quốc tế AES lần thứ 17 về Mã hóa âm thanh chất lượng cao, tháng 08/1999
- [15] International Standard ISO/IEC 23001-3:2010, Unified speech and audio coding Audio, Geneva, 2010.

[16] International Standard ISO/IEC 14496-3:2005, Information technology - Coding of audio-visual objects - Part 3: Audio, Geneva, 2005.

[17] P. Ekstrand, "Bandwidth Extension of Audio Signals by Spectral Band Replication", trong Chuyên đề của Hội thảo IEEE Benelux lần thứ nhất về MPCA, Leuven, tháng 11/2002

[18] F. Nagel, S. Disch, S. Wilde, A continuous modulated single sideband bandwidth extension, Hội nghị quốc tế ICASSP về âm học, xử lý tín hiệu và âm thanh, Dallas, Texas (Mỹ), tháng 04/2010

Bản chất kỹ thuật của sáng chế

Mục đích của sáng chế là đề xuất khái niệm mã hóa/giải mã được cải thiện.

Mục đích này đạt được bằng các thiết bị giải mã tín hiệu âm thanh được mã hóa theo yêu cầu bảo hộ điểm 1, thiết bị để mã hóa tín hiệu âm thanh theo điểm 10, phương pháp giải mã của điểm 16, phương pháp mã hóa theo điểm 18 hoặc chương trình máy tính theo điểm 19.

Sáng chế dựa trên việc tìm chất lượng được cải thiện và làm giảm tốc độ bit đặc biệt cho các tín hiệu bao gồm các phần chuyển tiếp mà chúng xuất hiện rất thường xuyên trong tín hiệu âm thanh thu được bởi sự kết hợp kỹ thuật Tạo hình nhiễu âm theo thời gian (Temporal Noise Shaping - TNS) hoặc Tạo hình ô theo thời gian (Temporal Tile Shaping - TTS) với việc khôi phục tần số cao. Xử lý TNS/TTS ở phía bộ mã hóa được thực hiện bằng dự báo qua tần số khôi phục đường bao theo thời gian của tín hiệu âm thanh. Phụ thuộc vào việc thực hiện, cụ thể là, khi bộ lọc tạo hình nhiễu âm theo thời gian được xác định trong khoảng tần số biến đổi không chỉ khoảng tần số nguồn mà còn khoảng tần số mục tiêu được khôi phục trong bộ giải mã khôi phục tần số, đường bao theo thời gian không chỉ được áp dụng cho tín hiệu âm thanh lõi tăng đến khi tần số bắt đầu điền đầy khoảng trống, mà đường bao theo thời gian cũng được áp dụng cho khoảng phổ của phần phổ thứ hai đã khôi phục. Do vậy, tiếng vang trước hoặc tiếng vang sau có thể sẽ xuất hiện mà không có việc tạo dạng ô theo thời gian bị giảm hoặc bị loại trừ. Việc này được hoàn thành bằng việc ứng dụng dự báo nghịch đảo trên tần số không chỉ trong khoảng tần số lõi tăng đến tần số bắt đầu

điền đầy khoảng trống nhất định mà còn trong các khoảng tần số trên khoảng tần số lỗi. Để đạt được điều này, việc tái tạo tần số hoặc tạo ra ô tần số được thực hiện ở phía bộ giải mã trước khi áp dụng dự báo vượt quá tần số. Tuy nhiên, tần số vượt quá dự báo có thể được áp dụng trước hoặc sau để tạo hình đường bao phổ phụ thuộc vào có hay không việc tính toán thông tin năng lượng được thực hiện trên giá trị dư phổ để lọc hoặc giá trị phổ (đầy đủ) trước khi tạo hình đường bao.

Việc xử lý TTS vượt quá một hoặc nhiều ô tần số cũng thiết lập tính liên tục của mối tương quan giữa phạm vi nguồn và phạm vi khôi phục hoặc trong hai phạm vi khôi phục liên kế hoặc hai ô tần số liên kế.

Trong việc thực hiện, ưu tiên sử dụng dạng lọc TNS/TTS phức hợp. Bằng cách đó, tránh được các giả âm sai số (theo thời gian) của việc biểu diễn thực tế đã lấy mẫu tới hạn, như MDCT. Bộ lọc TNS phức hợp có thể được tính toán ở phía bộ mã hóa bằng việc áp dụng không chỉ biến đổi cosin rời rạc được biến đổi mà còn áp dụng thêm biến đổi sin rời rạc được biến đổi để thu được biến đổi thay đổi phức hợp. Tuy nhiên, chỉ giá trị biến đổi cosin rời rạc được biến đổi, cụ thể là, phần thực của biến đổi phức hợp được truyền. Tuy nhiên, ở phía bộ giải mã, có thể ước lượng phần ảo của biến đổi bằng cách sử dụng phổ MDCT của các khung trước hoặc sau để ở phía bộ giải mã, bộ lọc phức hợp có thể lại được áp dụng trong dự báo ngược vượt quá tần số và, cụ thể, dự báo biên giữa phạm vi nguồn và phạm vi khôi phục và cũng vượt quá biên giữa ô tần số liên kế tần số trong phạm vi khôi phục.

Khía cạnh nữa dựa trên việc phát hiện ra vấn đề liên quan tới sự phân tách của một mặt là mở rộng băng thông và mặt khác là việc mã hóa lỗi có thể được giải quyết và khắc phục bằng cách thực hiện mở rộng băng thông trong cùng một miền phổ trong đó bộ giải mã lỗi hoạt động. Do đó, bộ giải mã lỗi tốc độ đầy đủ được đề xuất để mã hóa và giải mã toàn bộ khoảng tín hiệu âm thanh. Điều này không đòi hỏi phải có bộ lấy mẫu giảm ở phía bộ mã hóa và bộ lấy mẫu tăng ở phía bộ giải mã. Thay vào đó, toàn bộ quá trình xử lý được thực hiện trong toàn bộ tốc độ lấy mẫu hoặc toàn bộ miền băng thông. Để thu được độ khuếch đại mã hóa cao, tín hiệu âm thanh được phân tích để tìm tập hợp thứ nhất của phần phổ thứ nhất mà được mã hóa với độ phân giải cao, tại đó tập hợp thứ nhất này của phần phổ thứ nhất có thể bao gồm, trong phương án,

phần âm điệu của tín hiệu âm thanh. Mặt khác, các thành phần không âm hoặc thành phần nhiễu âm trong tín hiệu âm thanh tạo thành tập hợp thứ hai của phần phổ thứ hai được mã hóa theo tham số với độ phân giải phổ thấp. Tín hiệu âm thanh được mã hóa sau đó chỉ yêu cầu tập hợp thứ nhất của phần phổ thứ nhất được mã hóa theo cách bảo toàn dạng sóng với độ phân giải phổ cao và thêm vào đó, tập hợp thứ hai của phần phổ thứ hai mã hóa theo tham số với độ phân giải thấp sử dụng "các ô" tần số có nguồn gốc từ tập hợp thứ nhất. Ở phía bộ giải mã, bộ giải mã lỗi, là bộ giải mã đầy đủ, khôi phục tập hợp thứ nhất của phần phổ thứ nhất trong cách thức bảo toàn dạng sóng, tức là, không có bất kỳ kiến thức nào về việc tái tạo tần số bổ sung. Tuy nhiên, phổ đã tạo ra có nhiều khoảng trống phổ. Các khoảng trống này sau đó được lấp đầy bằng kỹ thuật Điền đầy khoảng trống thông minh sáng tạo (IGF) một mặt bằng việc sử dụng tái tạo tần số áp dụng dữ liệu theo tham số, mặt khác sử dụng khoảng phổ nguồn, tức là, phần phổ thứ nhất được khôi phục bởi bộ giải mã âm thanh tốc độ đầy đủ.

Trong phương án khác, các phần phổ, mà được khôi phục bởi việc điền đầy nhiễu âm chứ không phải bởi việc tái tạo băng thông hoặc điền đầy ô tần số, tạo thành bộ thứ ba của phần phổ thứ ba. Do thực tế là khái niệm mã hóa một mặt hoạt động trong miền đơn cho việc mã hóa/giải mã lỗi và mặt khác là tái tạo tần số, IGF không chỉ bị giới hạn để điền tới khoảng tần số cao hơn mà còn có thể điền tới các khoảng tần số thấp hơn, hoặc bởi việc điền đầy nhiễu âm không có tái tạo tần số hoặc bằng việc tái tạo tần số sử dụng ô tần số tại khoảng tần số khác nhau.

Hơn nữa, cần được nhấn mạnh rằng thông tin về các năng lượng phổ, thông tin về các năng lượng riêng rẽ hoặc thông tin năng lượng riêng rẽ, thông tin về năng lượng tồn tại hoặc thông tin năng lượng tồn tại, thông tin ô năng lượng hoặc thông tin năng lượng ô, hoặc thông tin về năng lượng thiếu hoặc thông tin năng lượng thiếu có thể bao gồm không chỉ giá trị năng lượng mà còn bao gồm giá trị biên độ (tuyệt đối), giá trị mức hoặc giá trị khác bất kỳ, từ đó có thể nhận được giá trị năng lượng cuối cùng. Do đó, thông tin về năng lượng có thể, ví dụ, bao gồm giá trị năng lượng của chính nó, và/hoặc giá trị của mức và/hoặc biên độ và/hoặc biên độ tuyệt đối.

Khía cạnh nữa được dựa trên việc tìm ra trạng thái tương quan không chỉ quan trọng cho phạm vi nguồn mà còn quan trọng cho cả khoảng đích. Hơn nữa, sáng chế

này xác nhận trạng thái mà trạng thái tương quan khác nhau có thể xuất hiện trong phạm vi nguồn hoặc khoảng đích. Ví dụ, khi tín hiệu tiếng nói với nhiều âm tần số cao được xem xét, trạng thái có thể là dải tần thấp bao gồm tín hiệu tiếng nói với lượng nhỏ âm bội là tương quan cao trong kênh trái và kênh phải, khi loa được đặt ở giữa. Tuy nhiên, phần tần số cao có thể không được tương quan mạnh do thực tế là có thể có nhiều âm tần số cao khác nhau trên phía bên trái được so sánh với nhiều âm tần số cao khác hoặc không có nhiều âm tần số cao trên phía bên phải. Do vậy, khi hoạt động điền đầy khoảng trống đơn giản sẽ được thực hiện mà bỏ qua trạng thái này, sau đó phần tần số cao sẽ cũng được tương quan, và điều này có thể tạo ra giả âm phân chia không gian nghiêm ngặt trong tín hiệu đã phục hồi. Để giải quyết vấn đề này, dữ liệu theo tham số cho dải phục hồi hoặc, thông thường, cho tập hợp thứ hai của phần phổ thứ hai mà được khôi phục sử dụng tập hợp thứ nhất của phần phổ thứ nhất được tính toán để nhận dạng hoặc biểu diễn kênh đôi khác nhau thứ nhất hoặc thứ hai cho phần phổ thứ hai hoặc trạng thái khác nhau cho dải khôi phục. Do đó, ở phía bộ mã hóa, việc nhận dạng kênh đôi được tính toán cho các phần phổ thứ hai, tức là, cho các phần bổ sung thông tin năng lượng cho các dải khôi phục được tính toán. Bộ tái tạo tần số ở phía bộ giải mã sau đó tái tạo phần phổ thứ hai phụ thuộc vào phần thứ nhất của tập hợp thứ nhất của các phần phổ thứ nhất, tức là, phạm vi nguồn và dữ liệu theo tham số cho phần thứ hai như thông tin năng lượng đường bao phổ hoặc dữ liệu đường bao phổ bất kỳ khác và, ngoài ra, phụ thuộc vào việc nhận dạng kênh đôi cho phần thứ hai, tức là, cho dải khôi phục dưới việc xem xét lại..

Việc nhận dạng kênh đôi tốt hơn là được truyền như dấu hiệu cho từng dải khôi phục và dữ liệu này được truyền từ bộ mã hóa đến bộ giải mã và bộ giải mã sau đó giải mã tín hiệu lỗi như đã biểu thị bởi dấu hiệu tính toán tốt nhất cho các dải lỗi. Sau đó, trong việc thực hiện, tín hiệu lỗi được lưu trữ trong cả biểu diễn âm lập thể (ví dụ trái/phải và ở giữa/bên cạnh) và, để việc điền đầy ô tần số IGF, việc biểu diễn ô nguồn được chọn để thích hợp với việc biểu diễn ô đích như được biểu thị bởi các dấu hiệu nhận dạng kênh đôi cho điền đầy khoảng trống thông minh hoặc các dải khôi phục, tức là, cho khoảng đích.

Cần được nhấn mạnh rằng quy trình này không chỉ thực hiện cho các tín hiệu âm lập thể, tức là, cho kênh trái và kênh phải mà còn hoạt động cho các tín hiệu đa kênh.

Trong trường hợp các tín hiệu đa kênh, nhiều cặp kênh khác nhau có thể được xử lý theo cách thức này như kênh bên trái và phải như cặp thứ nhất, kênh bao quanh bên trái và bao quanh bên phải như cặp thứ hai, kênh trung tâm và kênh LFE như cặp thứ ba. Việc tạo cặp khác có thể được xác định cho định dạng kênh đầu ra cao hơn như 7.1, 11.1, v.v.

Một khía cạnh nữa dựa trên việc tìm ra sự suy yếu nhất định trong chất lượng âm thanh có thể được khắc phục bằng việc áp dụng sơ đồ điền đầy ô tần số thích ứng. Để đạt được điều này, sự phân tích ở phía bộ mã hóa được thực hiện để tìm ra ứng viên vùng nguồn phù hợp nhất cho vùng đích nhất định. Việc nhận dạng thông tin phù hợp cho vùng đích vùng nguồn nhất định cùng với việc tùy chọn một số thông tin bổ sung được tạo ra và truyền như thông tin phụ tới bộ giải mã. Bộ giải mã sau đó áp dụng hoạt động điền đầy ô tần số sử dụng thông tin phù hợp. Để đạt được điều này, bộ giải mã đọc thông tin phù hợp từ dòng dữ liệu đã được truyền hoặc các tập dữ liệu và truy cập vùng nguồn đã được nhận dạng cho dải khôi phục nhất định và nếu được biểu thị trong thông tin phù hợp, thực hiện bổ sung một số xử lý dữ liệu vùng nguồn này để tạo ra dữ liệu phổ dạng thô cho dải khôi phục. Sau đó, kết quả này của hoạt động điền đầy ô tần số, cụ thể, dữ liệu phổ thô cho dải khôi phục, được tạo hình sử dụng thông tin đường bao phổ để cuối cùng thu được dải khôi phục mà chứa các phần phổ thứ nhất cũng như các phần âm. Tuy nhiên, những phần âm này không được tạo ra bởi sơ đồ điền đầy ô thích ứng, nhưng những phần phổ thứ nhất này là đầu ra bởi bộ giải mã âm thanh hoặc bộ giải mã lỗi trực tiếp.

Sơ đồ lựa chọn ô phổ thích ứng có thể hoạt động với độ chi tiết thấp. Trong việc thực hiện này, vùng nguồn được chia nhỏ thành vùng nguồn chồng lấp điển hình và vùng đích hoặc các dải khôi phục có được bởi những vùng đích tần số không chồng lấp. Sau đó, những điểm tương đồng giữa từng vùng nguồn và từng vùng đích được xác định ở phía bộ mã hóa và cặp phù hợp nhất của vùng nguồn và vùng đích được nhận dạng bằng thông tin phù hợp và, ở phía bộ giải mã, vùng nguồn được nhận dạng trong thông tin phù hợp được sử dụng cho việc tạo ra dữ liệu phổ thô cho dải khôi phục.

Với mục đích thu được độ chi tiết cao hơn, từng vùng nguồn được cho phép dịch chuyển để thu được độ trễ nhất định trong đó những điểm tương đồng là tối đa. Độ trễ này có thể khả quan như ngăn tần số và cho phép sự kết hợp thậm chí tốt hơn giữa vùng nguồn và vùng đích.

Ngoài ra, ngoài việc chỉ nhận dạng một cặp phù hợp nhất, độ trễ tương quan này cũng có thể được truyền trong thông tin phù hợp và, ngoài ra, ngay cả dấu hiệu cũng có thể được truyền. Khi dấu hiệu được xác định là âm ở phía bộ mã hóa, thì dấu hiệu biểu hiện tương quan cũng được truyền trong thông tin phù hợp và, ở phía bộ giải mã, các giá trị phổ vùng nguồn được nhân với “-1” hoặc, trong biểu diễn phức hợp, được “xoay” 180 độ.

Việc thực hiện khác của sáng chế này áp dụng hoạt động làm trắng ô. Việc làm trắng phổ loại bỏ thông tin đường bao phổ thô và nhấn mạnh cấu trúc phổ tinh là mối quan tâm hàng đầu để đánh giá độ tương tự của ô. Do đó, một mặt ô tần số và/hoặc mặt khác tín hiệu nguồn được làm trắng trước khi tính toán qua phép đo tương quan. Khi chỉ ô bị làm trắng bằng cách sử dụng quy trình được xác định trước, dấu hiệu làm trắng được truyền cho bộ giải mã rằng chu trình làm trắng được xác định trước đó sẽ được áp dụng cho ô tần số trong IGF.

Đối với việc lựa chọn ô, ưu tiên sử dụng độ trễ tương quan để dịch chuyển theo phổ phổ được tái tạo bằng số nguyên của các ngăn biến đổi. Phụ thuộc vào việc biến đổi cơ bản, việc dịch chuyển phổ có thể yêu cầu các hiệu chỉnh bổ sung. Trong trường hợp độ trễ lẻ, ô được biến đổi bổ sung thông qua sự khếch đại bởi chuỗi theo thời gian luân phiên của -1/1 để bù cho việc biểu diễn tần số được nghịch đảo của mọi dải khác trong MDCT. Ngoài ra, dấu hiệu của kết quả tương quan được áp dụng khi tạo ra ô tần số.

Ngoài ra, ưu tiên sử dụng việc lược bớt ô và làm ổn định để chắc chắn rằng giả âm được tạo ra bởi các vùng nguồn thay đổi nhanh cho cùng một vùng khôi phục hoặc cùng một vùng đích được tránh. Để đạt được điều này, phân tích điểm giống nhau giữa các vùng nguồn đã nhận dạng khác nhau được thực hiện và khi ô nguồn giống với ô nguồn khác với ngưỡng tương tự như trên, thì ô nguồn này có thể được loại bỏ khỏi tập hợp các ô nguồn tiềm năng vì nó tương quan cao với các ô nguồn khác. Hơn nữa,

dạng ổn định lựa chọn ô tốt hơn là để giữ thứ tự ô từ khung đứng trước nếu không có ô nguồn nào trong tương quan khung hiện tại (tốt hơn ngưỡng đã có) với các ô đích trong khung hiện tại.

Hệ thống mã hóa âm thanh theo sáng chế mã hóa hiệu quả các tín hiệu âm thanh tùy ý ở nhiều tốc độ bit. Trong khi đó, với tốc độ bit cao, hệ thống theo sáng chế hướng tới tính thông suốt, với tốc độ bit thấp sự không hài lòng trong cảm giác được giảm đến mức tối thiểu. Do đó, phần chính của tốc độ bit có sẵn được sử dụng để mã hóa dạng sóng chỉ là cấu trúc có liên quan nhất của tín hiệu trong bộ mã hóa, và các khoảng trống phổ thu được được điền đầy trong bộ giải mã bởi lượng tín hiệu xấp xỉ phổ ban đầu. Khối bit rất hạn chế được sử dụng để điều khiển việc được dẫn theo tham số mà được gọi là điền đầy khoảng trống thông minh (IGF) phổ bằng thông tin phụ chuyên biệt được truyền từ bộ mã hóa tới bộ giải mã.

Mô tả vắn tắt các hình vẽ

Các phương án ưu tiên của sáng chế sau đây được mô tả với việc tham chiếu đến các hình vẽ đi kèm trong đó:

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

Fig.1b minh họa bộ giải mã để giải mã tín hiệu âm thanh được 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 phép biểu diễn dưới dạng sơ đồ của phổ như được tạo ra bởi bộ giải mã miền phổ của Fig.1b;

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

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

Fig.4b minh họa việc 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 kỹ thuật MDCT.

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

Fig.6a minh họa bộ mã hóa âm thanh với chức năng tạo hình nhiễu âm theo thời gian/tạo hình ô theo thời gian;

Fig.6b minh họa bộ giải mã với kỹ thuật tạo hình nhiễu âm theo thời gian/tạo hình ô theo thời gian.

Fig.6c minh họa chức năng khác của chức năng tạo hình nhiễu âm theo thời gian/tạo hình ô theo thời gian với thứ tự khác của bộ lọc dự báo phổ và bộ tạo hình phổ;

Fig.7a minh họa việc thực hiện chức năng tạo hình ô theo thời gian (TTS);

Fig.7b minh họa việc thực hiện bộ giải mã phù hợp với việc thực hiện bộ mã hóa của Fig.7a;

Fig.7c minh họa ảnh phổ của tín hiệu gốc và tín hiệu được mở rộng không có TTS.

Fig.7d minh họa việc biểu diễn tần số mô tả sự tương đương giữa các tần số điền đầy khoảng trống thông minh và các năng lượng tạo hình ô theo thời gian;

Fig.7e minh họa ảnh phổ của tín hiệu gốc và tín hiệu được mở rộng với TTS;

Fig.8a minh họa bộ giải mã kênh đôi với việc tái tạo tần số;

Fig.8b minh họa bảng mô tả các tổ hợp khác nhau của việc biểu diễn và các phạm vi nguồn/đích;

Fig.8c minh họa đồ thị dòng mô tả chức năng của bộ giải mã kênh đôi với việc tái tạo tần số của Fig.8a;

Fig.8d minh họa việc thực hiện chi tiết hơn của bộ giải mã của Fig.8a;

Fig.8e minh họa việc thực hiện của bộ mã hóa cho xử lý kênh đôi để được giải mã bằng bộ giải mã của Fig.8a:

Fig.9a minh họa bộ giải mã với kỹ thuật tái tạo tần số sử dụng N giá trị lượng cho khoảng tần số tái tạo;

Fig.9b minh họa việc thực hiện chi tiết hơn của bộ tái tạo tần số của Fig.9a;

Fig.9c minh họa sơ đồ mô tả chức năng của Fig.9b;

Fig.9d minh họa việc thực hiện Ngoài ra cho bộ giải mã của Fig.9a;

Fig.10a minh họa sơ đồ khối của bộ mã hóa phù hợp với bộ giải mã của Fig.9a;

Fig.10b minh họa sơ đồ khối mô tả chức năng nữa của bộ tính toán theo tham số của Fig.10a;

Fig.10c minh họa sơ đồ khối mô tả chức năng nữa của bộ tính toán theo tham số của Fig.10a;

Fig.10d minh họa sơ đồ khối mô tả chức năng khác của bộ tính toán theo tham số của Fig.10a;

Fig.11a minh họa bộ giải mã nữa có sự nhận dạng phạm vi nguồn cụ thể cho hoạt động điền đầy ô phổ trong bộ giải mã;

Fig.11b minh họa chức năng khác của bộ tái tạo tần số của Fig.11a;

Fig.11c minh họa bộ mã hóa được sử dụng để phối hợp với bộ giải mã của Fig.11a;

Fig.11d minh họa sơ đồ khối của việc thực hiện bộ tính toán theo tham số của Fig.11c;

Fig.12a và 12b minh họa các phác họa tần số để mô tả phạm vi nguồn và khoảng đích;

Fig.12c minh họa đồ thị của tương quan ví dụ của hai tín hiệu;

Fig.13a minh họa bộ mã hóa đã có trong tình trạng kỹ thuật với việc mở rộng băng thông; và

Fig.13b minh họa bộ giải mã đã có trong tình trạng kỹ thuật với việc mở rộng băng thông.

Mô tả chi tiết sáng chế

Fig.1a minh họa thiết bị để mã hóa tín hiệu âm thanh 99. Tín hiệu âm thanh 99 là đầu vào vào bộ chuyển đổi phổ theo thời gian 100 để chuyển đổi tín hiệu âm thanh có tốc độ lấy mẫu thành biểu diễn phổ 101 xuất ra bởi bộ biến đổi phổ theo thời gian. Phổ 101 là đầu vào vào bộ phân tích phổ 102 để phân tích phép biểu diễn phổ 101. Bộ phân tích phổ 101 được tạo cấu hình để xác định tập hợp thứ nhất của các phần phổ thứ nhất 103 được mã hóa với độ phân giải phổ thứ nhất và tập hợp thứ hai khác của các phần phổ thứ hai 105 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 của các phần phổ thứ hai 105 là đầu vào 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 đề xuất để tạo ra phép biểu diễn được mã hóa thứ nhất 107 của tập hợp thứ nhất của các phần phổ thứ nhất có độ phân giải phổ thứ nhất. Ngoài ra, bộ tính toán theo tham số/bộ mã hóa theo tham số 104 được tạo cấu hình để tạo ra phép biểu diễn được mã hóa thứ hai 109 của tập hợp thứ hai của các phần phổ thứ hai. Phép biểu diễn được mã hóa thứ nhất 107 và phép biểu diễn được mã hóa thứ hai 109 là đầu vào vào bộ trộn kênh dòng bit hoặc bộ tạo dòng bit 108 và khối 108 cuối cùng xuất ra tín hiệu âm thanh được mã hóa để truyền hoặc lưu trữ trên thiết bị lưu trữ.

Cụ thể, phần phổ thứ nhất chẳng hạn 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 lõi bị giới hạn dải.

Fig.1b minh họa bộ giải mã phù hợp với bộ mã hóa của Fig.1a. Việc biểu diễn được mã hóa thứ nhất 107 là đầu vào vào bộ giải mã âm thanh miền phổ 112 để tạo ra biểu diễn được giải mã thứ nhất của tập hợp thứ nhất của các phần phổ thứ nhất, phép biểu diễn được giải mã có độ phân giải phổ thứ nhất. Ngoài ra, phép biểu diễn được mã hóa thứ hai 109 là đầu vào vào bộ giải mã theo tham số 114 để tạo ra phép biểu diễn được giải mã thứ hai của tập hợp thứ hai của các phần phổ thứ hai có độ phân giải phổ thứ hai thấp hơn độ phân giải phổ thứ nhất.

Bộ giải mã còn 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 hoạt động điền đầy ô, tức là, sử dụng ô hoặc phần của tập hợp thứ nhất của các phần phổ thứ nhất và sao chép tập hợp thứ nhất này của các phần phổ thứ nhất thành phạm vi khôi phục hoặc dải khôi phục có phần phổ thứ hai và thực hiện một cách đặc trưng việc tạo hình đường bao phổ hoặc hoạt động khác như được biểu thị bởi đầu ra biểu diễn thứ hai được giải mã bởi bộ giải mã theo tham số 114, tức là, bằng việc sử dụng thông tin về tập hợp thứ hai của các phần phổ thứ hai. Tập hợp thứ nhất được giải mã của các phần phổ thứ nhất và tập hợp thứ hai được khôi phục của các phần phổ như được biểu thị tại đầu ra của bộ tái tạo tần số 116 trên đường dẫn 117 là đầu vào vào bộ chuyển đổi phổ theo thời gian 118 được tạo cấu hình để chuyển đổi biểu diễn được giải mã thứ nhất và phần phổ thứ hai được khôi phục thành phép biểu diễn theo thời gian 119, phép biểu diễn theo thời gian 119 có tốc độ lấy mẫu cao nhất định.

Fig.2b minh họa việc thực hiện của bộ mã hóa theo Fig.1a. Tín hiệu đầu vào âm thanh 99 là đầu vào vào giàn lọc phân tích 220 tương ứng với bộ chuyển đổi phổ theo thời gian 100 của Fig.1a. Sau đó, hoạt động tạo hình nhiễu âm theo thời gian được thực hiện trong khối TNS 222. Do đó, đầu vào 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 hoạt động tạo hình nhiễu âm theo thời gian/hoạt động tạo hình ô theo thời gian không được áp dụng hoặc có thể là các giá trị dư phổ, khi hoạt động TNS như đã mô tả trên Fig.2b, khối 222 được áp dụng. Với các tín hiệu kênh đôi hoặc tín hiệu đa kênh, việc mã hóa kênh ghép 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 ghép 228. Ngoài ra, bộ mã hóa entropy 232 để thực hiện việc nén dữ liệu không tổn hao được đề xuất mà cũng là một 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 dải lõi và các thành phần âm tương ứng với tập hợp thứ nhất của các phần phổ thứ nhất 103 và các thành phần dư tương ứng với tập hợp thứ hai của các phần phổ thứ hai 105 của Fig.1a. Khối 224 được biểu thị như việc mã hóa chiết xuất theo tham số tương ứng với bộ mã hóa theo tham số 104 của Fig.1a và bộ trộn kênh dòng bit 230 tương ứng với bộ trộn kênh dòng bit 108 của Fig.1a.

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

Bộ phân tích phổ 226 ưu tiên 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 lõi 228 mã hóa tất cả các thành phần âm với mô đun cảm quan. Giai đoạn ước lượng màn chắn âm có thể được thực hiện theo nhiều cách khác nhau và tốt hơn là được thực hiện tương tự về chức năng của nó so với giai đoạn ước lượng vết hình sin được sử dụng trong mô hình sin và nhiễu âm mô hình hóa cho việc giải mã âm thanh/tiếng nói trong các tài liệu [8, 9] hoặc bộ mã hóa âm thanh dựa trên mô hình HILN được mô tả trong tài liệu [10]. Tốt hơn là, việc thực hiện được sử dụng dễ dàng thực hiện mà không cần duy trì quỹ đạo sinh tử, nhưng bộ phát hiện âm hoặc nhiễu âm khác bất kỳ có thể cũng được sử dụng.

Mô đun IGF tính toán sự giống nhau mà tồn tại giữa phạm vi nguồn và khoảng đích. Khoảng đích được biểu diễn bởi phổ từ phạm vi nguồn. Phép đo sự giống nhau giữa phạm vi nguồn và khoảng đích được thực hiện bằng cách sử dụng phương pháp tương quan chéo. Khoảng đích được tách ra thành n_{Tar} ô tần số không chồng lấp. Với mỗi ô trong vùng đích, n_{Src} ô nguồn được tạo ra từ tần số bắt đầu cố định. Các ô nguồn này chồng lấp bởi hệ số nằm trong khoảng 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 tương quan nhất với đích được lưu trữ trong $xcorr_lag[idx_tar][idx_src]$ và dấu hiệu tương quan được lưu trữ trong $xcorr_sign[idx_tar][idx_src]$. Trong trường hợp tính tương quan là rất âm, ô nguồn cần được làm nhân với -1 trước quy trình điền đầy ô tại bộ giải mã. Mô đun 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 dải được sử dụng để lưu trữ năng lượng của vùng đích cho phép khôi phục phổ chính xác.

Phương pháp này có một số ưu điểm nhất định so với SBR cổ điển [1] ở chỗ lưới sóng hài của tín hiệu đa âm được bảo toàn bởi bộ mã hóa lỗi 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. Một ưu điểm khác của hệ thống này so với Thay thế Phổ Chính xác (Accurate Spectral Replacement - ASR) [2-4] là không có giai đoạn tổng hợp tín hiệu tạo các phần quan trọng của tín hiệu tại bộ giải mã. Thay vào đó, nhiệm vụ này được đảm nhận bởi bộ mã hóa lỗi, cho phép bảo toàn các thành phần quan trọng của phổ. Một ưu điểm khác của hệ thống được đề xuất khả năng mở rộng liên tục mà các tính năng được đề ra. Chỉ cần sử dụng $tileNum[idx_tar]$ và $xcorr_lag = 0$, cho mọi ô được gọi là kết hợp độ chi tiết tổng thể và có thể được sử dụng cho tốc độ bit thấp trong khi sử dụng biến số $xcorr_lag$ cho mọi ô cho phép khớp với phổ nguồn và phổ đích tốt hơn.

Ngoài ra, kỹ thuật làm ổn định chọn ô được đề xuất nhằm loại bỏ các giả âm miền tần số như nhiễu âm âm rung và nhiễu âm âm nhạc.

Trong trường hợp các cặp kênh âm lập thể, việc xử lý âm lập thể chung bổ sung được áp dụng. Điều này là cần thiết, bởi vì đối với một phạm vi đích nhất định, tín hiệu có thể là nguồn âm thanh được điều chỉnh tương quan cao. Trong trường hợp các vùng nguồn được chọn cho vùng cụ thể này không tương quan tốt, mặc dù các năng lượng được khớp với các vùng đích, hình ảnh không gian có thể bị ảnh hưởng do các vùng nguồn không tương thích. Bộ mã hóa phân tích từng dải năng lượng của vùng đích, thường thực hiện mối tương quan chéo giữa các giá trị phổ và nếu vượt quá ngưỡng nhất định, sẽ thiết lập dấu hiệu chung cho dải năng lượng này. Trong bộ giải mã các dải năng lượng kênh phải và trái được xử lý riêng biệt nếu dấu hiệu âm lập thể chung này không được thiết lập. Trong trường hợp dấu hiệu âm lập thể chung được thiết lập, cả năng lượng và bản vá đều được thực hiện trong miền âm lập thể chung. Thông tin âm lập thể chung cho các vùng IGF được báo hiệu tương tự thông tin âm lập thể chung 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 là chỉ số tần số trong miền biến đổi.

Giải pháp khác để tính toán và truyền năng lượng trực tiếp trong miền âm lập thể chung cho các dải mà âm lập thể chung là hoạt động, Do đó không cần biến đổi năng lượng bổ sung tại phía bộ giải mã.

Các 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 chung $\rightarrow$ Biến đổi LR:

Nếu không có 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 tín hiệu 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 tín hiệu từ bên tới giữa:

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

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

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

Quy trình này đảm bảo rằng từ các ô được sử dụng để tái tạo các vùng đích tương quan cao và các vùng đích xoay chuyển, các kênh bên trái và bên phải thu được vẫn biểu diễn nguồn âm thanh tương quan và xoay chuyển ngay cả khi các vùng nguồn không tương quan, thì vẫn bảo toàn hình ảnh âm lập thể cho những vùng này.

Nói cách khác, trong dòng bit, các dấu hiệu âm lập thể chung được truyền mà biểu thị L/R hoặc M/S như là ví dụ cho việc mã hóa âm lập thể chung thông thường sẽ được sử dụng. Trong bộ giải mã, thứ nhất, tín hiệu lỗi được giải mã như được biểu thị bởi các dấu hiệu âm lập thể chung cho các dải lỗi. Thứ hai, tín hiệu lỗi được lưu trữ trong cả biểu diễn L/R và M/S. Với điền đầy ô IGF, việc biểu diễn ô nguồn được chọn để phù hợp với việc biểu diễn ô đích như được biểu thị bởi thông tin âm lập thể chung cho các dải 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 là việc mở rộng của sơ đồ cơ bản của bộ mã hóa theo cảm giác, chèn bước xử lý tùy chọn giữa giàn lọc và mức lượng tử hóa. Nhiệm vụ chính của môđun TNS là làm ảm nhiễu âm lượng tử hóa đã tạo ra trong vùng chắn theo thời gian của chuyển tiếp dạng các tín hiệu và do đó dẫn tới sơ đồ mã hóa hiệu quả hơn. Thứ nhất, TNS tính toán bộ 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 thời gian của tín hiệu. Vì lượng tử hoá tác động tới phổ đã lọc TNS, nhiễu âm lượng tử hóa cũng được làm phẳng theo thời gian. Bằng cách áp dụng việc lọc TNS nghịch đảo ở phía bộ giải mã, nhiễu âm lượng tử hóa được tạo hình theo đường bao theo thời gian của bộ lọc TNS và Do đó nhiễu âm lượng tử hóa được chắn bởi phần chuyển tiếp.

IGF được dựa trên biểu diễn MDCT. Để mã hóa hiệu quả, tốt hơn là sử dụng các khối dài xấp xỉ 20ms. Nếu tín hiệu trong khối dài này chứa các phần chuyển tiếp, các trước-tiếng vang và sau-tiếng vang có thể nghe được xuất hiện trong dải phổ IGF do việc điền đầy ô. Fig.7c cho thấy hiệu quả trước-tiếng vang điển hình trước sự 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 điền đầy TNS được thể hiện.

Hiệu quả trước-tiếng vang này bị giảm bởi việc sử dụng TNS trong phạm vi IGF. Ở đây, TNS được sử dụng như công cụ tạo hình ô theo thời gian (TTS) như phép biểu diễn phổ trong bộ giải mã được thực hiện trên tín hiệu dư TNS. Các hệ số dự báo TTS cần thiết được tính toán và áp dụng sử dụng phổ đầy đủ ở 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. Khi so sánh với TNS kế thừa, tần số dừng TTS được tăng đến tần số dừng của công cụ IGF, mà cao hơn $f_{IGFstart}$. Ở phía bộ giải mã, các hệ số TNS/TTS được áp dụng trên phổ đầy đủ một lần nữa, tức là, phổ lỗi 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 thời gian của phổ được tái tạo để phù hợp với đường bao của tín hiệu gốc một lần nữa. Vì vậy trước-tiếng vang đã thể hiện được giảm. Bên cạnh đó, nó vẫn tạo hình nhiều âm lượng tử hóa trong tín hiệu bên dưới $f_{IGFstart}$ như thường lệ với TNS.

Trong các bộ giải mã kế thừa, việc vá phổ trên tín hiệu âm thanh làm sai độ tương quan phổ tại các biên của bản vá và do đó làm suy yếu đường bao thời gian của tín hiệu âm thanh bằng việc đưa ra sự phân tán. Do đó, lợi ích khác của việc thực hiện điền đầy ô IGF trên tín hiệu dư là, sau khi áp dụng bộ lọc tạo hình, các biên ô có mối tương quan liên mạch, dẫn đến việc tái tạo tín hiệu theo thời gian trung thực hơn.

Trong bộ mã hóa theo sáng chế, phổ đã trải qua quá trình lọc TNS/TTS, xử lý màn chắn âm và ước lượng tham số IGF không có bất kỳ tín hiệu nào trên tần số bắt đầu IGF ngoại trừ các thành phần âm. Phổ tần số này hiện được mã hóa bởi bộ mã hóa lõi sử dụng các nguyên tắc mã hóa số học và mã hóa dự báo. Các thành phần được mã hóa này cùng với các bit tín hiệu tạo thành dòng bit âm thanh.

Fig.2a minh họa việc thực hiện bộ giải mã tương quan. Dòng bit trên Fig.2a tương ứng với tín hiệu âm thanh được mã hóa là đầu vào vào bộ tách kênh/bộ giải mã mà sẽ được kết nối, tham chiếu đến Fig.1b, với khối 112 và 114. Bộ tách kênh dòng bit phân tách tín hiệu âm thanh đầu vào thành phép biểu diễn được mã hóa thứ nhất 107 của Fig.1b và phép biểu diễn được mã hóa thứ hai 109 của Fig.1b. Phép biểu diễn được mã hóa thứ nhất có tập hợp thứ nhất của các phần phổ thứ nhất là đầu vào vào khối giải mã kênh ghép 204 tương ứng với bộ giải mã miền phổ 112 của Fig.1b. Phép biểu diễn được mã hóa thứ hai là đầu vào vào bộ giải mã theo tham số 114 không được minh họa trên Fig.2a sau đó đầu vào vào khối IGF 202 tương ứng với bộ tái tạo tần số 116 của Fig.1b. Tập hợp thứ nhất của các phần phổ thứ nhất được yêu cầu cho sự tái tạo tần số là đầu vào vào khối IGF 202 theo đường 203. Ngoài ra, sau việc giải mã

kênh ghép 204, việc giải mã lõi cụ thể được áp dụng trong khối màn chắn âm 206 sao cho đầu ra của màn chắn âm 206 tương ứng với đầu ra của bộ giải mã miền phổ 112. Sau đó, thực hiện việc kết hợp bởi bộ kết hợp 208, tức là, việc tạo khung tại đó đầu ra của bộ kết hợp 208 bây giờ có phổ khoảng đầy đủ, nhưng vẫn trong miền được lọc TNS/TTS. Sau đó, trong khối 210, hoạt động TNS/TTS nghịch đảo được thực hiện sử dụng thông tin bộ lọc TNS/TTS được cung cấp thông qua đường dẫn 109, tức là, thông tin phụ TTS tốt hơn là được chứa trong phép 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 phép 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 là tần số khoảng đầy đủ được xác định bởi tốc độ lấy mẫu của tín hiệu đầu vào gốc. Sau đó, việc chuyển đổi theo 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 phép biểu diễn dưới dạng sơ đồ của phổ. Phổ được chia nhỏ ra trong các dải hệ số tỷ lệ SCB tại đó có bảy dải hệ số tỷ lệ từ SCB1 tới SCB7 trong ví dụ minh họa của Fig.3a. Các dải hệ số tỷ lệ có thể là dải hệ số tỷ lệ AAC mà được xác định theo tiêu chuẩn AAC và có băng thông tăng dần tới tần số cao hơn như đã minh họa trên Fig.3a dưới dạng sơ đồ. Ưu tiên thực hiện điền đầy khoảng trống thông minh không phải từ đầu phổ, tức là ở các tần số thấp, mà bắt đầu hoạt động IGF ở tần số bắt đầu IGF được minh họa tại 309. Do đó, dải tần lõi mở rộng từ tần số thấp nhất tới 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 của các phần phổ thứ nhất) từ các thành phần có độ phân giải thấp được biểu diễn bởi tập hợp thứ hai của phần phổ thứ hai. Fig.3a minh họa phổ mà là ví dụ đầu vào vào bộ mã hóa miền phổ 106 hoặc bộ mã hóa kênh ghép 228, tức là, bộ mã hóa lõi hoạt động trong dải đầy đủ, nhưng mã hóa một lượng đáng kể giá trị phổ 0, tức là, các giá trị phổ 0 này được lượng tử hóa thành 0 hoặc là được đặt thành 0 trước khi lượng tử hóa hoặc sau khi lượng tử hóa. Dù thế nào đi nữa, bộ mã hóa lõi hoạt động trong khoảng đầy đủ, tức là, nếu khi phổ được minh họa, tức là, bộ giải mã lõi không nhất thiết phải biết về bất kỳ việc điền đầy khoảng trống thông minh hoặc mã hóa tập hợp thứ hai của các phần phổ thứ hai với độ phân giải phổ thấp hơn.

Tốt hơn là, độ phân giải cao được xác định bằng việc mã hóa theo đường của các đường phổ như dòng MDCT, trong khi độ phân giải thứ hai hoặc độ phân giải thấp được xác định, chẳng hạn như, bởi việc tính toán chỉ một giá trị phổ đơn trên dải hệ số tỷ lệ, tại đó dải hệ số tỷ lệ bao trùm các đường tần số khác nhau. Do đó, độ phân giải thấp thứ hai, về khía cạnh để nó là độ phân giải phổ, thấp hơn nhiều độ phân giải thứ nhất hoặc cao được xác định bởi việc mã hóa theo đường được ứng dụng bởi bộ mã hóa lỗi như AAC hoặc USAC.

Đối với hệ số tỷ lệ hoặc việc tính toán năng lượng, tình huống này được minh họa trên Fig.3b. Do bộ mã hóa là bộ mã hóa lỗi và do thực tế là có thể, nhưng không nhất thiết phải có, các thành phần của tập hợp thứ nhất của các phần phổ trong từng dải, bộ mã hóa lỗi tính toán hệ số tỷ lệ cho mỗi dải không chỉ trong khoảng lỗi dưới tần số bắt đầu IGF 309 mà còn trên tần số bắt đầu IGF cho đến khi tần số cực đại $f_{IGF_{stop}}$ 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 cùng với các hệ số tỷ lệ SB1 tới SB7 tương ứng với dữ liệu phổ độ phân giải cao. Dữ liệu phổ có độ phân giải thấp được tính toán bắt đầu từ tần số bắt đầu IGF và tương ứng với giá trị thông tin năng lượng E_1, E_2, E_3, E_4 , mà được truyền cùng với các hệ số tỷ lệ từ SF4 tới SF7.

Cụ thể, khi bộ mã hóa lỗi dưới điều kiện tốc độ bit thấp, hoạt động điền đầy nhiều âm bổ sung trong dải lỗi, tức là, thấp hơn về tần số so với tần số bắt đầu IGF, tức là, trong các dải hệ số tỷ lệ từ SB1 tới SB3 có thể được áp dụng bổ sung. Trong việc điền đầy nhiều âm, có các đường phổ liên kế khác nhau tồn tại mà đã lượng tử hóa thành 0. Ở phía bộ giải mã, các giá trị phổ đã lượng tử hóa thành 0 được tổng hợp lại và các giá trị phổ đã tổng hợp lại được điều chỉnh trong độ âm lượng của chúng sử dụng năng lượng điền đầy nhiều âm như NF2 đã mô tả ở 308 trên Fig.3b. Năng lượng điền đầy nhiều âm, có thể được đưa ra theo thuật ngữ tuyệt đối hoặc theo thuật ngữ tương đối đặc biệt liên quan đến yếu tố tỷ lệ như trong USAC tương ứng với năng lượng của tập hợp các giá trị phổ được lượng tử hóa thành 0. Các đường phổ điền đầy nhiều âm này có thể cũng được xem xét là tập hợp thứ ba của phần phổ thứ ba mà được tái tạo bằng cách tổng hợp điền đầy nhiều âm đơn giản mà không cần bất kỳ hoạt động IGF trên

việc tái tạo tần số sử dụng ô tần số từ các tần số khác cho ô tần số khôi phục sử dụng giá trị phổ từ phạm vi nguồn và thông tin năng lượng E_1, E_2, E_3, E_4 .

Tốt hơn là, các dải mà thông tin năng lượng được tính toán đồng thời với dải hệ số tỷ lệ. Trong phương án khác, việc tạo nhóm giá trị thông tin năng lượng được áp dụng sao cho, ví dụ, cho các dải hệ số tỷ lệ 4 và 5, chỉ có giá trị thông tin năng lượng đơn được truyền, nhưng thậm chí trong phương án này, các biên của các dải khôi phục được nhóm trùng với các biên của dải hệ số tỷ lệ. Nếu việc phân tách dải khác nhau được áp dụng, thì việc tính toán lại nhất định hoặc tính toán đồng bộ hóa có thể được áp dụng, và điều này có thể tạo cảnh 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 tâm thính học như được minh họa trên Fig.4a. Diễn hình là, như ví dụ đã mô tả theo tiêu chuẩn PEG2/4 AAC hoặc tiêu chuẩn MPEG1/2, tiêu chuẩn Lớp 3, tín hiệu âm thanh được mã hóa sau khi truyền vào trong khoảng phổ (401 trên Fig.4a) được truyền thẳng tới bộ tính toán hệ số tỷ lệ 400. Bộ tính toán hệ số tỷ lệ được điều khiển bởi mô hình tâm thính học nhận một cách bổ sung tín hiệu âm thanh đã lượng tử hóa hoặc nhận biểu diễn phổ phức hợp của tín hiệu âm thanh như theo tiêu chuẩn MPEG1/2 Lớp 3 hoặc tiêu chuẩn MPEG AAC phức hợp. Mô hình tâm thính học tính toán, cho từng dải hệ số tỷ lệ, hệ số tỷ lệ biểu diễn ngưỡng tâm thính học. Ngoài ra, sau đó các hệ số tỷ lệ được điều chỉnh sao cho các điều kiện tốc độ bit nhất định được đáp ứng bởi việc phối hợp của vòng lặp bên trong và bên ngoài đã biết hoặc bởi bất kỳ quy trình mã hóa thích hợp. Sau đó, một mặt các giá trị phổ đã lượng tử hóa và mặt khác hệ số tỷ lệ được tính toán là đầu vào vào bộ xử lý lượng tử hóa 404. Trong hoạt động của bộ mã hóa âm thanh đơn giản, các giá trị phổ lượng tử hóa được gán trọng số bởi hệ số tỷ lệ và, giá trị phổ được gán trọng số là đầu vào vào bộ lượng tử hóa cố định đặc trưng có chức năng nén tới khoảng biên độ cao hơn. Sau đó, tại đầu ra của bộ xử lý lượng tử hóa tại đó thực hiện lượng tử hóa các chỉ số mà sau đó được chuyển vào trong bộ mã hóa entropy đặc trưng có sự mã hóa cụ thể và rất hiệu quả cho bộ của các chỉ số lượng tử hóa 0 cho các giá trị tần số liền 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 về các phần phổ thứ hai từ bộ phân tích phổ. Do đó, bộ xử lý lượng tử hóa 404 đảm bảo rằng, trong đầu ra của bộ xử lý lượng tử hóa 404, các phần phổ thứ hai được nhận dạng bởi bộ phân tích phổ 102 bằng 0 hoặc có phép biểu diễn được xác nhận bởi bộ mã hóa hóa hoặc bộ giải mã là phép 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 việc thực hiện của bộ xử lý lượng tử hóa. Các giá trị phổ MDCT có thể là đầu vào vào khối thiết lập thành 0 410. Sau đó, các phần phổ thứ hai đã thiết lập thành 0 trước khi gán trọng số bởi các hệ số tỷ lệ trong khối 412 được thực hiện. Trong việc thực hiện bổ sung, khối 410 không được cung cấp, nhưng việc phối hợp thiết lập thành 0 được thực hiện trong khối 418 sau khối gán trọng số 412. Trong việc thực hiện khác, việc hoạt động thiết lập thành 0 có thể cũng được thực hiện trong khối thiết lập thành 0 422 sau việc lượng tử hóa trong khối lượng tử hóa 420. Trong việc thực hiện này, khối 410 và 418 sẽ không có mặt. Nói chung, ít nhất một trong số các khối 410, 418, 422 được cung cấp phụ thuộc vào từng việc thực hiện cụ thể.

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

Các khối thiết lập thành 0 410, 418, 422 được cung cấp thay thế cho nhau hoặc song song được kiểm soát bởi bộ phân tích phổ 424. Bộ phân tích phổ tốt hơn là bao gồm việc thực hiện bất kỳ của bộ phát hiện âm đã biết hoặc bao gồm dạng khác bất kỳ của bộ phát hiện hoạt động để phân tách phổ thành các thành phần được mã hóa với độ phân giải cao và các thành phần 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 nào khác quyết định, phụ thuộc vào thông tin phổ hoặc lý lịch 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 theo thời gian 100 của Fig.1a, ví dụ, như đã bổ sung trong AAC hoặc USAC. Bộ chuyển đổi

phổ theo thời gian 100 bao gồm bộ tạo cửa sổ 502 được điều khiển bởi bộ phát hiện chuyển tiếp 504. Khi bộ phát hiện chuyển tiếp 504 phát hiện chuyển tiếp, sau đó luân phiên từ các cửa sổ dài tới cửa sổ ngắn được tạo tín hiệu tới bộ tạo cửa sổ. Bộ tạo cửa sổ 502 sau đó tính toán, cho các khối chồng lấp, các khung cửa sổ, trong đó mỗi khung cửa sổ đặc trưng có N giá trị chẳng hạn như 2048 giá trị. Sau đó, việc biến đổi trong bộ biến đổi khối 506 được thực hiện, và bộ biến đổi khối điển hình này cung cấp bổ sung việc lấy thập phân sao cho việc lấy thập phân/biến đổi kết hợp được thực hiện để thu được khung phổ với N giá trị chẳng hạn như các giá trị phổ MDCT. Do đó, đối với hoạt động cửa sổ dài, khung tại đầu vào của khối 506 bao gồm N giá trị chẳng hạn như 2048 giá trị và khung phổ sau đó có 1024 giá trị. Tuy nhiên, sau đó, việc chuyển đổi được thực hiện cho các khối ngắn, khi tám khối ngắn được thực hiện tại từng khối ngắn có $1/8$ các giá trị miền thời gian tạo cửa sổ được so sánh với cửa sổ dài và mỗi khối phổ có $1/8$ giá trị phổ khi so sánh với khối dài. Do đó, việc lấy thập phân này được kết hợp với hoạt động chồng lấp 50% của bộ tạo cửa sổ, phổ là bản được lấy mẫu tới hạn của tín hiệu âm thanh miền thời gian 99.

Tiếp đó, tham chiếu đến Fig.5b minh họa việc thực hiện cụ thể của bộ tái tạo tần số 116 và bộ chuyển đổi phổ-thời gian 118 của Fig.1b, hoặc của hoạt động kết hợp của các khối 208, 212 của Fig.2a. Trên Fig.5b, dải khôi phục cụ thể được xem xét như dải hệ số tỷ lệ 6 của Fig.3a. Phần phổ thứ nhất trong dải khôi phục này, tức là, phần phổ thứ nhất 306 của Fig.3a là đầu vào vào khối xây dựng/điều chỉnh khung 510. Hơn nữa, phần phổ thứ hai được khôi phục cho dải hệ số tỷ lệ 6 cũng là đầu vào vào bộ xây dựng/điều chỉnh khung 510. Ngoài ra, thông tin năng lượng chẳng hạn như E_3 của Fig.3b cho dải hệ số tỷ lệ 6 cũng là đầu vào vào khối 510. Phần phổ thứ hai được khôi phục trong dải khôi phục được tạo ra bởi việc điền đầy ô tần số sử dụng phạm vi nguồn và dải khôi phục sau đó tương ứng với khoảng đích. Bây giờ, việc điều chỉnh năng lượng của khung được thực hiện để cuối cùng thu được khung được khôi phục hoàn chỉnh có N giá trị, ví dụ như được thu tại đầu ra của bộ kết hợp 208 của Fig.2a. Sau đó, trong khối 512, phép biến đổi/nội suy khối nghịch đảo được thực hiện để thu được các giá trị miền thời gian 248 cho, ví dụ, các giá trị phổ 124 tại đầu vào của khối 512. Sau đó, hoạt động tạo cửa sổ tổng hợp được thực hiện trong khối 514 lại được kiểm soát lần nữa sự biểu 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, hoạt động chồng lấp/bổ sung với khung thời gian đứng trước được thực hiện. Tốt hơn là, MDCT áp dụng việc chồng lấp 50% sao cho, với mỗi khung thời gian mới của 2N giá trị, giá trị miền thời gian N là đầu ra cuối cùng. Việc chồng lấp 50% được ưu tiên mạnh do thực tế là nó cung cấp việc lấy mẫu tới hạn và giao nhau liên tục từ một khung tới khung kế tiếp do hoạt động chồng lấp/bổ sung trong khối 516.

Như đã mô tả tại 301 trên Fig.3a, hoạt động 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 chẳng hạn như cho dải khôi phục đã dự tính trùng với dải hệ số tỷ lệ 6 của Fig.3a. Sau đó, các giá trị phổ điền đầy nhiễu âm có thể cũng là đầu vào vào bộ điều chỉnh/bộ xây dựng 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 bằng cách sử dụng năng lượng điền đầy nhiễu âm trước khi trở thành đầu vào vào bộ điều chỉnh/bộ xây dựng khung 510.

Tốt hơn là, hoạt động IGF, tức là, hoạt động 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 dải cao trên tần số bắt đầu IGF mà còn cũng được áp dụng trong cả dải thấp. Hơn nữa, việc điền đầy ô tần số không có nhiễu âm điền đầy 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 phát hiện ra rằng việc mã hóa âm thanh có chất lượng cao và hiệu quả cao có thể thu được khi hoạt động điền đầy nhiễu âm được giới hạn tới khoảng tần số dưới tần số bắt đầu IGF và khi hoạt động điền đầy ô tần số bị hạn chế tới khoảng tần số trên tần số bắt đầu IGF như được mô tả trên Fig.3a.

Tốt hơn là, các ô đích (TT) (có tần số lớn hơn tần số bắt đầu IGF) bị giới hạn tới các biên dải hệ số tỷ lệ của bộ mã hóa tỷ lệ đầy đủ. Các ô nguồn (ST), từ thông tin thu được, tức là, cho các tần số thấp hơn tần số bắt đầu IGF không được tạo giới hạn bởi các biên dải hệ số tỷ lệ. Kích thước của ST nên tương ứng với kích thước của TT đã kết hợp. Điều này được mô tả bằng cách sử dụng ví dụ dưới đây. TT[0] là độ dài của 10 ngăn MDCT. Điều này tương ứng chính xác với độ dài của hai SCB tiếp theo

(chẳng hạn như $4 + 6$). Sau đó, tất cả ST có thể mà tương quan với TT[0], cũng có độ dài là 10 ngăn. Ô đích thứ hai TT[1] liền kề với TT[0] có độ dài là 15 ngăn I (SCB có độ dài $7 + 8$). Sau đó, ST có độ dài là 15 ngăn chứ không phải là 10 ngăn như TT[0].

Nếu trường hợp phát sinh là không thể tìm thấy TT cho ST với độ dài của ô đích (ví dụ khi độ dài của TT lớn hơn phạm vi nguồn khả dụng), thì không tính được tương quan và phạm vi nguồn được sao chép một số lần vào TT này (việc sao chép được thực hiện lần lượt sao cho đường tần số cho tần số thấp nhất của việc sao chép thứ hai ngay sau đó - theo tần số - đường tần số cho tần số cao nhất của việc sao chép thứ nhất), cho đến khi ô đích TT được điền đầy hoàn toàn.

Tiếp đó, tham chiếu đến Fig.5c minh họa phương án được ưu tiên nữa của bộ tái tạo tần số 116 của Fig.1b hoặc khối IGF 202 của Fig.2a. Khối 522 là bộ tạo ô tần số nhận, không chỉ dải đích ID mà còn nhận thêm dải nguồn ID. Dẫn chứng là, được xác định ở phía bộ mã hóa rằng dải hệ số tỷ lệ 3 của Fig.3a rất phù hợp để khôi phục dải hệ số tỷ lệ 7. Do đó, dải nguồn ID sẽ là 2 và dải đích ID sẽ là 7. Dựa trên thông tin này, bộ tạo ô tần số 522 áp dụng việc sao chép tăng hoặc hoạt động điền đầy ô sóng hài hoặc hoạt động điền đầy ô khác bất kỳ để tạo ra phần thứ hai thô của thành phần phổ 523. Phần thứ hai thô của các thành phần phổ có độ phân giải tần số giống hệt với độ phân giải tần số được chứa trong tập hợp thứ nhất của các phần phổ thứ nhất.

Sau đó, phần phổ thứ nhất của dải khôi phục chẳng hạn như 307 của Fig.3a là đầu vào vào bộ xây dựng khung 524 và phần thứ hai thô 523 cũng là đầu vào vào bộ xây dựng khung 524. Sau đó, khung đã khôi phục được điều chỉnh bởi bộ điều chỉnh 526 bằng cách sử dụng hệ số khuếch đại cho dải khôi phục được tính toán bởi bộ tính toán hệ số khuếch đại 528. Tuy nhiên, quan trọng là, 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 bị ảnh hưởng bởi bộ điều chỉnh 526. Để đạt được điều này, bộ tính toán hệ số khuếch đại 528 phân tích dải nguồn hoặc phần thứ hai thô 523 và phân tích bổ sung phần phổ thứ nhất trong dải khôi phục để cuối cùng tìm ra hệ số khuếch đại chuẩn 527 sao cho năng lượng của đầu ra khung đã điều chỉnh xuất ra bởi bộ điều chỉnh 526 có năng lượng E_4 khi dải hệ số tỷ lệ 7 được dự tính.

Trong phạm vi này, quan trọng là ước tính độ chính xác khôi phục tần số cao của sáng chế so với HE-AAC. Điều này được giải thích đối với dải hệ số tỷ lệ 7 trên Fig.3a. Giả định rằng bộ mã hóa đã biết trong tình trạng kỹ thuật như mô tả trên Fig.13a sẽ phát hiện phần phổ 307 để 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 dải khôi phục như dải hệ số tỷ lệ 7 tới bộ giải mã. Sau đó, bộ giải mã sẽ tạo lại sóng hài khuyết. Tuy nhiên, giá trị phổ, tại sóng hài khuyết 307 sẽ được khôi phục bằng bộ giải mã đã biết trong tình trạng kỹ thuật của Fig.13b sẽ ở giữa dải 7 tại tần số được biểu thị bằng tần số khôi phục 390. Do đó, sáng chế tránh lỗi tần số 391 mà sẽ được đưa vào bởi bộ giải mã đã biết trong tình trạng kỹ thuật của Fig.13d.

Trong việc 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 điểm tương đồng đã tính toán, phần phổ thứ hai trong phạm vi khôi phục phần phổ thứ nhất phù hợp với phần phổ thứ hai nhất có thể. Sau đó, trong việc thực hiện phạm vi nguồn/khoảng đích có thể thay đổi này, bộ mã hóa theo tham số sẽ được đưa bổ sung vào trong phép biểu diễn được mã hóa thứ hai thông tin phù hợp biểu thị cho từng khoảng đích phạm vi nguồn phù hợp. Ở phía bộ giải mã, thông tin này sau đó sẽ được sử dụng bởi bộ tạo ô tần số 522 của Fig.5c minh họa việc tạo ra phần thứ hai thô 523 dựa trên dải nguồn ID và dải đích ID.

Ngoài ra, như được minh họa trên Fig.3a, bộ phân tích phổ được tạo cấu hình để phân tích phép 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 là ít nhất một phần tư của tần số lấy mẫu hoặc cao hơn.

Như đã 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 phép biểu diễn phổ có tần số Nyquist được xác định bởi tốc độ lấy mẫu của tín hiệu âm thanh đầu vào gốc.

Hơn nữa, như được minh họa trên Fig.3a, bộ phân tích phổ được tạo cấu hình để phân tích phép biểu diễn phổ bắt đầu với tần số bắt đầu điền đầy khoảng trống và kết

thức với tần số tối đa được biểu diễn bởi tần số tối đa được chứa trong phép 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 của các phần phổ và trong đó các phần phổ như 304, 305, 306, 307 có các giá trị tần số trên tần số điền đầy khoảng trống được chứa bổ sung trong tập hợp thứ nhất của các phần phổ thứ nhất.

Như đã phác họa, bộ giải mã âm thanh miền phổ 112 được tạo cấu hình để tần số tối đa được biểu diễn bởi giá trị phổ trong phép biểu diễn được giải mã thứ nhất bằng với tần số tối đa được chứa trong phép biểu diễn theo thời gian có tốc độ lấy mẫu trong đó giá trị phổ cho tần số tối đa trong tập hợp thứ nhất của các phần phổ thứ nhất là không hoặc khác không. Dù thế nào, để tần số tối đa này trong tập hợp thứ nhất của các thành phần phổ hệ số tỷ lệ cho dải hệ số tỷ lệ tồn tại, mà được tạo ra và được truyền bất kể có hay không tất cả giá trị phổ trong dải hệ số tỷ lệ này được thiết lập thành không hoặc không như đã thảo luận trong phạm vi của Fig.3a và Fig.3b.

Do đó, ưu điểm của sáng chế này đối với các kỹ thuật tham số khác là 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 nhiều âm như cảnh tín hiệu cục bộ) sáng chế cho phép tái tạo tần số chính xác của các thành phần âm. Đến nay, không kỹ thuật thuộc tình trạng kỹ thuật nào hướng đến việc biểu diễn tham số hiệu quả của lượng tín hiệu tùy ý bởi việc điền đầy khoảng trống phổ không hạn chế của sự phân chia ưu tiên được cố định trong dải thấp (LF) và dải cao (HF).

Các phương án của hệ thống theo sáng chế cải thiện cách tiếp cận trong 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 ngay cả với tốc độ bit thấp.

Hệ thống thông thường bao gồm:

- mã hóa lõi dải đầy đủ
- điền đầy khoảng trống thông minh (điền đầy ô hoặc điền đầy nhiều âm)
- các phần âm điệu thừa thớt trong lõi được chọn bởi màn chắn âm
- mã hóa cặp âm lập thể chung cho dải đầy đủ, bao gồm 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 cho biến đổi dữ liệu phổ thành miền biến đổi thứ hai khác với một trong số bộ mã hóa lỗi. Như đa số bộ mã hóa-giải mã âm thanh, chẳng hạn như AAC là ví dụ, sử dụng MDCT làm biến đổi căn bản, sẽ hữu ích để biến đổi BWE trong miền MDCT. Yêu cầu thứ hai với hệ thống BWE sẽ cần thiết để bảo toàn việc ghi 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 cho sơ đồ BWE, hệ thống mới theo dự kiến được gọi điền đầy khoảng trống thông minh (IGF). Fig.2b thể hiện sơ đồ khối của hệ thống đã dự kiến ở phía bộ mã hóa và Fig.2a thể hiện hệ thống ở phía bộ giải mã.

Fig.6a minh họa thiết bị giải mã tín hiệu âm thanh được mã hóa trong việc thực hiện khác của sáng chế. Thiết bị để giải mã bao gồm bộ giải mã âm thanh miền phổ 602 để tạo ra phép biểu diễn được giải mã thứ nhất của bộ thứ nhất của các phần phổ và bộ tái tạo tần số 604 được kết nối xuôi chiều với bộ giải mã âm thanh miền phổ 602 để tạo ra phần phổ thứ hai được khôi phục sử dụng phần phổ thứ nhất của tập hợp thứ nhất của các phần phổ thứ nhất. Như minh họa ở 603, các giá trị phổ trong phần phổ thứ nhất và phần phổ thứ hai là các giá trị dự dự báo phổ. Để biến đổi các giá trị dự dự báo phổ thành phép biểu diễn phổ đầy đủ, bộ lọc dự báo phổ 606 được đề xuất. Bộ lọc dự báo nghịch đảo được tạo cấu hình để thực hiện dự báo nghịch đảo trên tần số sử dụng các giá trị dự phổ cho tập hợp thứ nhất của tần số thứ nhất và các phần phổ thứ hai được khôi phục. Bộ lọc dự báo nghịch đảo phổ 606 được tạo cấu hình bởi thông tin bộ lọc có trong tín hiệu âm thanh được mã hóa. Fig.6b minh họa việc thực hiện chi tiết hơn của phương án Fig.6a. Các giá trị dự dự báo phổ 603 là đầu vào vào bộ tạo ra ô tần số 612 tạo ra các giá trị phổ thô cho dải khôi phục hoặc cho phần tần số thứ hai nhất định và dữ liệu thô này bây giờ có độ phân giải giống với độ phân giải cao của biểu diễn phổ thứ nhất là đầu vào vào trong bộ tạo hình phổ 614. Bộ tạo hình phổ bây giờ tạo hình phổ sử dụng thông tin đường bao được truyền trong dòng bit và dữ liệu đã tạo hình phổ sau đó được đặt vào bộ lọc dự báo phổ 616 cuối cùng tạo ra khung của

các giá trị phổ đầy đủ sử dụng thông tin thông tin bộ lọc 607 được truyền từ bộ mã hóa tới bộ giải mã thông qua dòng bit.

Trên Fig.6b, giả định rằng ở phía bộ mã hóa, việc tính toán thông tin bộ lọc được truyền thông qua dòng bit và sử dụng thông qua đường 607 được thực hiện tiếp theo tính toán thông tin đường bao. Do đó, nói cách khác, bộ mã hóa phù hợp với bộ giải mã của Fig.6b sẽ tính toán các giá trị dư phổ và sau đó sẽ tính toán thông tin đường bao với các giá trị dư phổ như, ví dụ, được minh họa trên Fig.7a. Tuy nhiên, việc thực hiện khác cũng hữu ích cho các thực hiện nhất định, tại đó thông tin đường bao được tính toán trước khi thực hiện lọc TNS hoặc TTS ở phía bộ mã hóa. Sau đó, bộ lọc dự báo phổ 622 được áp dụng trước khi thực hiện tạo hình phổ trong khối 624. Do vậy, nói cách khác, các giá trị phổ (đầy đủ) được tạo ra trước hoạt động tạo hình phổ 624 được áp dụng.

Tốt hơn là, bộ lọc TNS giá trị phức hợp hoặc bộ lọc TTS được tính toán. Điều này được minh họa trên Fig.7a. Tín hiệu âm thanh gốc là đầu vào vào khối MDCT phức hợp 702. Sau đó, việc tính toán bộ lọc TTS và việc lọc TTS được thực hiện trong miền phức hợp. Sau đó, trong khối 706, thông tin phụ IGF được tính toán và bất kỳ hoạt động khác như phân tích phổ cho mã hóa v.v cũng được tính toán. Tiếp theo, tập hợp thứ nhất của phần phổ thứ nhất được tạo ra bởi khối 706 được mã hóa với bộ mã hóa mô hình dẫn tâm thính học đã minh họa tại 708 để thu được tập hợp thứ nhất của các phần phổ thứ nhất được biểu thị tại $X(k)$ trên Fig.7a và tất cả dữ liệu này được đưa tới bộ trộn dòng bit 710.

Ở phía bộ giải mã, dữ liệu được mã hóa là đầu vào vào bộ tách kênh 720 để một mặt phân tách thông tin phụ IGF, mặt khác phân tách thông tin phụ TTS và việc biểu diễn được mã hóa của tập hợp thứ nhất của các phần phổ thứ nhất.

Sau đó, khối 724 được sử dụng để tính toán phổ phức hợp từ một hoặc nhiều phổ giá trị thực. Tiếp đó, cả phổ giá trị thực và phổ phức hợp là đầu vào vào khối 726 để tạo ra các giá trị tần số được khôi phục trong tập hợp thứ hai của các phần phổ thứ hai cho dải khôi phục. Sau đó, trên khung dải đầy đủ đã thu hoàn chỉnh hoặc khung dải đầy đủ đã điền đầy ô, hoạt động TTS nghịch đảo 728 được thực hiện và, ở phía bộ giải mã, việc hoạt động MDCT phức hợp nghịch đảo cuối cùng được thực hiện trong khối

730. Do đó, việc sử dụng thông tin bộ lọc TNS phức hợp cho phép, khi được áp dụng không chỉ trong dải lỗi hoặc trong dải ô phân tách mà còn được áp dụng trên các biên lỗi/ô hoặc biên ô/ô tạo ra một cách tự động xử lý biên ô, mà cuối cùng, lại đưa vào tương quan phổ giữa các ô. Tương quan phổ trên biên ô không thu được chỉ bằng cách tạo ra ô tần số và thực hiện điều chỉnh đường bao phổ trên dữ liệu thô của các ô tần số.

Fig.7c minh họa phép so sánh của tín hiệu gốc (bảng trái) và tín hiệu được mở rộng không có TTS. Điều này có thể nhận thấy rằng có các giả âm mạnh được minh họa bởi các phần được mở rộng trong khoảng tần số cao hơn được minh họa tại 750. Tuy nhiên, điều này không xuất hiện trên Fig.7e khi phân phổ tương tự tại 750 được so sánh với thành phần liên quan giả âm 750 của Fig.7c.

Các phương án hoặc hệ thống mã hóa âm thanh theo sáng chế sử dụng phần chính của tốc độ bit có sẵn để mã hóa dạng sóng chỉ kết cấu thích hợp nhất theo cảm giác của tín hiệu trong bộ mã hóa, và khoảng trống phổ thu được được điền đầy trong bộ giải mã với lượng tín hiệu mà bao gồm gần như xấp xỉ phổ gốc. Gói bit rất giới hạn được dùng để kiểm soát tham số được dẫn được gọi là điền đầy khoảng trống thông minh phổ (IGF) bằng thông tin phụ dành riêng được truyền từ bộ mã hóa tới bộ giải mã.

Việc lưu trữ hoặc truyền các tín hiệu âm thanh thường là đối tượng tuân theo các ràng buộc tốc độ bit chặt chẽ. Trước đây, bộ mã hóa buộc phải giảm đáng kể băng thông âm thanh được truyền khi mà chỉ tốc độ bit thấp có sẵn. Bộ mã hóa-giải mã âm thanh hiện đại ngày nay có khả năng mã hóa các tín hiệu dải rộng bằng việc sử dụng các phương pháp mở rộng băng thông (BWE) như Sao chép dải rộng phổ (Spectral Bandwidth Replication - SBR) [1]. Những thuật toán này dựa trên phép biểu diễn theo tham số của lượng tần số cao (HF) mà được tạo thành từ phần tần số thấp được mã hóa dạng sóng (LF) của tín hiệu được giải mã bằng cách chuyển vào vùng phổ HF ("vá lỗi") và ứng dụng xử lý sau được dẫn theo tham số. Trong sơ đồ BWE, việc khôi phục vùng phổ HF trên được gọi là tần số giao nhau có được thường dựa trên việc vá phổ. Điển hình, vùng HF gồm nhiều bản vá liên kế và mỗi bản vá này được bắt nguồn từ các vùng băng thông qua (BP) của phổ LF thấp hơn tần số giao nhau có được. Hệ

thống thuộc tình trạng kỹ thuật thực hiện hiệu quả với việc vá lỗi trong biểu diễn gián lọc bởi việc sao chép bộ hệ số băng con liên kề từ vùng nguồn tới vùng đích.

Nếu hệ thống BWE được thực hiện trong gián lọc hoặc miền biến đổi thời gian-tần số, đây là khả năng giới hạn để điều khiển tạo hình theo thời gian của tín hiệu được mở rộng băng thông. Điển hình, độ chi tiết theo thời gian bị giới hạn bởi kích thước bước nhảy được sử dụng giữa các cửa sổ biến đổi liên kề. Điều này có thể dẫn tới trước-tiếng vang hoặc sau-tiếng vang không mong muốn trong khoảng phổ BWE.

Từ việc mã hóa âm thanh theo cảm giác, người ta biết rằng hình dạng của đường bao thời gian của tín hiệu âm thanh có thể được khôi phục bằng việc sử dụng các kỹ thuật lọc phổ như Tạo hình đường bao theo thời gian (TNS) [14]. Tuy nhiên, bộ lọc TNS được biết tới trong tình trạng kỹ thuật là bộ lọc giá trị thực trên phổ giá trị thực. Bộ lọc giá trị thực như vậy trên phổ giá trị thực có thể bị suy yếu nghiêm trọng bởi giả âm sai số, đặc biệt nếu việc biến đổi thực tế bên dưới là Biến đổi cosin rời rạc được biến đổi (MDCT).

Việc tạo hình ô đường bao thời gian áp dụng lọc phức hợp trên phổ giá trị phức hợp thu được từ, ví dụ, Biến đổi cosin gián đoạn được biến đổi phức hợp (Modified Discrete Cosine Transform - CMDCT). Bằng cách đó, giả âm sai số được ngăn chặn.

Tạo hình ô theo thời gian bao gồm

- ước lượng và áp dụng hệ số bộ lọc phức hợp của bộ lọc làm phẳng trên phổ tín hiệu gốc tại bộ mã hóa.

- biến đổi các hệ số bộ lọc trong thông tin phụ

- áp dụng bộ lọc tạo hình trên phổ khôi phục được điền đầy ô trong bộ giải mã.

Sáng chế mở rộng kỹ thuật trong tình trạng kỹ thuật đã biết từ mã hóa biến đổi âm thanh, cụ thể, Tạo hình nhiều âm theo thời gian (TNS) bằng dự báo tuyến tính theo hướng tần số, để sử dụng cách thức thay đổi trong phạm vi mở rộng băng thông.

Ngoài ra, thuật toán mở rộng băng thông theo sáng chế dựa trên Điền đầy khoảng trống thông minh (IGF), nhưng sử dụng việc tăng tần số lấy mẫu, biến đổi có giá trị phức hợp (CMDCT), trái ngược với cấu hình tiêu chuẩn IGF dựa trên phép biểu diễn MDCT được lấy mẫu giá trị thực của tín hiệu. CMDCT có thể xem như tổ hợp của hệ

số MDCT trong phần thực và các hệ số MDST trong phần ảo của từng hệ số phở có giá trị phức hợp.

Mặc dù cách tiếp cận mới được mô tả trong phạm vi IGF, quy trình theo sáng chế có thể được sử dụng kết hợp với bất kỳ phương pháp BWE nào khác dựa trên phép biểu diễn giàn lọc của tín hiệu âm thanh.

Trong phạm vi mới này, dự báo tuyến tính theo hướng tần số không được sử dụng làm tạo hình nhiễu âm theo thời gian, mà là kỹ thuật tạo hình ô theo thời gian (TTS). Việc đổi tên được chứng minh bằng thực tế là các thành phần tín hiệu điền đầy ô được tạo hình theo thời gian bằng TTS trái ngược với tạo hình nhiễu âm lượng tử hóa bằng TNS trong bộ mã hóa-giải mã biến đổi cảm giác trong tình trạng kỹ thuật.

Fig.7a thể hiện sơ đồ khối của bộ mã hóa BWE sử dụng IGF và cách tiếp cận TTS mới.

Sơ đồ mã hóa cơ bản làm việc như sau:

- tính toán CMDCT của tín hiệu miền thời gian $x(n)$ để có tín hiệu miền tần số $X(k)$
- tính toán bộ lọc TTS có giá trị phức hợp
- có được thông tin phụ cho BWE và loại bỏ thông tin phở mà được tái tạo bằng bộ giải mã
- áp dụng lượng tử hóa sử dụng môđun tâm thính học (psycho acoustic module - PAM).
- lưu trữ/truyền dữ liệu, chỉ các hệ số MDCT có giá trị thực được truyền.

Fig.7b thể hiện bộ giải mã tương ứng. Nó đảo ngược chủ yếu các bước được thực hiện trong bộ mã hóa.

Ở đây, sơ đồ giải mã cơ bản làm việc như sau:

- ước lượng các hệ số MDST từ các giá trị MDCT (việc xử lý này thêm một độ trễ bộ giải mã khối) và kết hợp các hệ số MDCT với MDST thành các hệ số CMDCT có giá trị phức hợp.
- thực hiện điền đầy ô với việc xử lý sau của nó

- áp dụng việc lọc TTS nghịch đảo với hệ số bộ lọc TTS đã truyền
- tính toán CMDCT nghịch đảo

Lưu ý rằng, thay vào đó, để tổng hợp TTS và sau xử lý IGF có thể cũng bị nghịch đảo trong bộ giải mã nếu phân tích TTS và ước lượng tham số IGF được nghịch đảo phù hợp trong bộ mã hóa.

Với mã hóa biến đổi hệ số, tốt hơn là được gọi là "khối dài" gần 20ms được sử dụng để đạt được lợi ích biến đổi hợp lý. Nếu tín hiệu trong khối dài bao gồm các chuyển tiếp, có thể nghe thấy trước-tiếng vang và sau-tiếng vang xuất hiện trong dải phổ được khôi phục do điền đầy ô. Fig.7c thể hiện các hiệu ứng trước-tiếng vang và sau-tiếng vang điển hình mà làm suy yếu sự chuyển tiếp do IGF. Trên bảng trái của Fig.7c, thể hiện ảnh phổ của tín hiệu gốc, và trên bảng phải thể hiện ảnh phổ của tín hiệu đã điền đầy ô không có điền đầy TTS. Trong ví dụ này, tần số bắt đầu IGF $f_{IGFstart}$ hoặc f_{split} giữa dải lõi và dải đã điền đầy ô được chọn $f_s/4$. Trong bảng phải của Fig.7c, trước-tiếng vang và sau-tiếng vang rõ rệt bao quanh có thể thấy được các chuyển tiếp, đặc biệt đáng chú ý tại đầu phổ trên của vùng tần số tái tạo.

Nhiệm vụ chính của môđun TTS là để hạn chế các thành phần tín hiệu không mong muốn xung quanh lân cận vùng chuyển tiếp và do đó làm ẩn chúng trong vùng thời gian được không chế bởi hiệu quả chần theo thời gian của nhận thức con người. Do đó, các hệ số dự báo TTS cần thiết được tính toán và áp dụng bằng cách sử dụng "dự báo phía trước" trong miền CMDCT.

Trong phương án kết hợp TTS và IGF thành bộ mã hóa-giải mã là việc quan trọng để điều chỉnh thẳng hàng tham số TTS và tham số IGF nhất định như ô IGF hoặc là được điền đầy hoàn toàn bởi một bộ lọc TTS (làm phẳng hoặc lọc tạo hình) hoặc không. Do đó, tất cả tần số $TTSstart[.]$ hoặc $TTSstop[.]$ sẽ không có trong ô IGF, nhưng được làm thẳng hàng với tần số $f_{IGF...}$. Fig.7d thể hiện ví dụ của khu vực hoạt động TTS và IGF cho bộ ba bộ lọc TTS.

Tần số dừng TTS được điều chỉnh đến tần số dừng của công cụ IGF, mà cao hơn $f_{IGFstart}$. Nếu TTS sử dụng nhiều hơn một bộ lọc, nó được đảm bảo rằng tần số giao

nhau giữa hai bộ lọc TTS là phù hợp với tần số chia IGF. Mặt khác, một bộ lọc con TTS sẽ chạy trên $f_{IGFstart}$ dẫn đến các giả âm không mong muốn như tạo hình trên.

Trong việc thực hiện biến thể đã mô tả trên Fig.7a và Fig.7b, cần chú ý thêm là năng lượng trong bộ giải mã IGF được điều chỉnh chính xác. Điều này đặc biệt xảy ra nếu trong quá trình xử lý TTS và IGF, các bộ lọc TTS khác nhau có độ khuếch đại dự báo khác nhau được áp dụng cho vùng nguồn (như bộ lọc phẳng) và vùng phổ đích (như bộ lọc tạo hình không phải là bản sao chính xác của bộ lọc phẳng đã nêu) của một ô IGF. Trong trường hợp này, tỷ lệ khuếch đại dự báo của hai bộ lọc TTS đã áp dụng không bằng nhau nữa và do đó phải điều chỉnh năng lượng theo tỷ lệ này.

Trong biến thể thực hiện thay thế, thứ tự sau xử lý IGF và TTS bị đảo ngược. Trong bộ giải mã, điều này có nghĩa là việc điều chỉnh năng lượng bởi sau xử lý IGF được tính toán sau việc lọc TTS và do đó là bước xử lý cuối cùng trước khi biến đổi tổng hợp. Do đó, dù độ khuếch đại bộ lọc TTS khác nhau được áp dụng cho một ô trong suốt quá trình mã hóa, năng lượng cuối cùng luôn được điều chỉnh chính xác bằng cách xử lý IGF.

Ở phía bộ giải mã, các hệ số bộ lọc TTS được áp dụng lại trên phổ đầy đủ, tức là, phổ lỗi được mở rộng bằng phổ đã tái tạo. Việc áp dụng TTS là cần thiết để hình thành đường bao thời gian của phổ đã tái tạo cho phù hợp với đường bao của tín hiệu gốc. Vì trước-tiếng vang đã thể hiện được làm giảm. Ngoài ra, vẫn tạo hình theo thời gian nhiều âm lượng tử hóa trong tín hiệu bên dưới $f_{IGFstart}$ như với TNS kế thừa.

Trong các bộ mã hóa kế thừa, việc vá lỗi phổ trên tín hiệu âm thanh (ví dụ SBR) làm sai lệch sự tương quan phổ tại biên vá lỗi và theo đó làm suy yếu đường bao thời gian của tín hiệu âm thanh bởi sự phân tán được đưa vào. Do đó, lợi ích khác của việc thực hiện điền đầy ô IGF trên tín hiệu dư là, sau khi áp dụng bộ lọc tạo hình TTS, các biên ô được tương quan liền mạch, thu được nhiều hơn sự tái tạo theo thời gian trung thực của tín hiệu.

Kết quả của tín hiệu được xử lý phù hợp được thể hiện trên Fig.7e. Trong việc so sánh bản không được lọc (bảng phải Fig.7c) tín hiệu đã được lọc TTS thể hiện mức giảm tốt của trước-tiếng vang và sau-tiếng vang không mong muốn (bảng phải Fig.7e).

Ngoài ra, như đã thảo luận, Fig.7a minh họa bộ mã hóa phù hợp với bộ giải mã Fig.7b hoặc bộ giải mã Fig.6a. Về cơ bản, thiết bị để mã hóa tín hiệu âm thanh bao gồm bộ chuyển đổi thời gian-phổ chẳng hạn như 702 để chuyển đổi tín hiệu âm thanh thành phép biểu diễn phổ. Việc biểu diễn phổ có thể là việc biểu diễn phổ giá trị thực hoặc, như minh họa trong khối 702, biểu diễn phổ giá trị phức hợp. Hơn nữa, bộ lọc dự báo như 704 để thực hiện việc dự báo trên tần số được cung cấp để tạo ra giá trị dư phổ, trong đó bộ lọc dự báo 704 được xác định bằng thông tin bộ lọc dự báo nhận được từ tín hiệu âm thanh và được đưa tới bộ trộn kênh dòng bit 710 như minh họa ở 714 trên Fig.7a. Ngoài ra, bộ mã hóa âm thanh như bộ mã hóa âm thanh dẫn tâm thính học 704 được đề xuất. Bộ mã hóa âm thanh được tạo cấu hình để mã hóa tập hợp thứ nhất của phần phổ thứ nhất của giá trị dư phổ để thu được tập hợp thứ nhất được mã hóa của giá trị phổ thứ nhất. Thêm vào đó, bộ mã hóa theo tham số như bộ mã hóa minh họa ở 706 trên Fig.7a được cung cấp để mã hóa tập hợp thứ hai của phần phổ thứ hai. Tốt hơn là, tập hợp thứ nhất của phần phổ thứ nhất được mã hóa với độ phân giải phổ cao hơn tập hợp thứ hai của phần phổ thứ hai.

Cuối cùng, như minh họa trên Fig.7a, giao diện đầu ra được cung cấp để đưa ra tín hiệu được mã hóa bao gồm tập hợp thứ hai được mã hóa theo tham số của phần phổ thứ hai, tập hợp thứ nhất được mã hóa của phần phổ thứ nhất và thông tin bộ lọc được minh họa là "thông tin phụ TTS" tại 714 trên Fig.7a.

Tốt hơn là, bộ lọc dự báo 704 bao gồm bộ tính thông tin bộ lọc được tạo cấu hình để sử dụng các giá trị phổ của phép biểu diễn phổ cho việc tính toán thông tin bộ lọc. Hơn nữa, bộ lọc dự báo được tạo cấu hình để tính toán giá trị dư phổ sử dụng giá trị phổ tương tự của phép biểu diễn phổ đã sử dụng cho việc tính toán thông tin bộ lọc.

Tốt hơn là, bộ lọc TTS 704 được tạo cấu hình bằng cách tương tự như đã biết với bộ mã hóa âm thanh theo tình trạng kỹ thuật áp dụng công cụ TNS phù hợp với tiêu chuẩn AAC.

Tiếp theo, việc thực hiện khác sử dụng giải mã kênh đôi được thảo luận trong phạm vi các hình vẽ từ Fig.8a tới Fig.8e. Ngoài ra, tham chiếu được thực hiện để mô tả các yếu tố tương đương trong phạm vi của Fig.2a, Fig.2b (mã hóa kênh ghép 228 và giải mã kênh ghép 204).

Fig.8a minh họa bộ giải mã âm thanh để tạo ra tín hiệu kênh đôi được giải mã. Bộ giải mã âm thanh bao gồm bốn bộ giải mã âm thanh 802 để giải mã tín hiệu kênh đôi đã được mã hóa để thu tập hợp thứ nhất của các phần phổ thứ nhất và bộ giải mã theo tham số 804 bổ sung để cung cấp dữ liệu tham số cho tập hợp thứ hai của các phần phổ thứ hai và, ngoài ra, việc nhận dạng kênh đôi nhận dạng hoặc là phép biểu diễn kênh đôi khác nhau thứ nhất hoặc thứ hai cho các phần phổ thứ hai. Thêm vào đó, bộ tái tạo tần số 806 được cung cấp để tạo ra phần phổ thứ hai phụ thuộc vào phần phổ thứ nhất của tập hợp thứ nhất của phần phổ thứ nhất và dữ liệu tham số cho phần thứ hai và nhận dạng kênh đôi cho phần thứ hai. Fig.8b minh họa các tổ hợp khác nhau cho phép biểu diễn kênh đôi trong phạm vi nguồn và khoảng đích. Phạm vi nguồn có thể trong phép biểu diễn kênh đôi thứ nhất và khoảng đích có thể cũng là phép biểu diễn kênh đôi thứ nhất. Ngoài ra, phạm vi nguồn có thể là trong phép biểu diễn kênh đôi thứ nhất và khoảng đích có thể trong phép biểu diễn kênh đôi thứ hai. Ngoài ra, phạm vi nguồn có thể là trong phép biểu diễn kênh đôi thứ hai và khoảng đích có thể trong là phép biểu diễn kênh đôi thứ nhất như đã minh họa trong cột thứ ba của Fig.8b. Cuối cùng, cả phạm vi nguồn và khoảng đích có thể trong phép biểu diễn kênh đôi thứ hai. Trong phương án, phép biểu diễn kênh đôi thứ nhất là phép biểu diễn kênh đôi phân tách mà kênh đôi của tín hiệu kênh đôi được biểu diễn riêng rẽ. Sau đó, phép biểu diễn kênh đôi thứ nhất là phép biểu diễn ghép nơi mà kênh đôi của phép biểu diễn kênh đôi được biểu diễn ghép, tức là, trong đó việc xử lý thêm hoặc biến đổi biểu diễn được yêu cầu để tính toán lại phép biểu diễn kênh đôi phân tách như đã yêu cầu để xuất ra các loa tương ứng.

Trong việc thực hiện, phép biểu diễn kênh đôi thứ nhất có thể là phép biểu diễn trái/phải (L/R) và phép biểu diễn kênh đôi thứ hai là phép biểu diễn âm lập thể chung. Tuy nhiên, các phép biểu diễn kênh đôi khác từ trái/phải hoặc M/S hoặc dự báo âm lập thể có thể được áp dụng và sử dụng cho sáng chế.

Fig.8c minh họa lưu đồ cho các thao tác được thực hiện bởi bộ giải mã âm thanh của Fig.8a. Trong bước 812, bộ giải mã âm thanh 802 thực hiện việc giải mã phạm vi nguồn. Phạm vi nguồn có thể bao gồm, đối với Fig.3a, các dải hệ số tỷ lệ từ SCB1 tới SCB3. Ngoài ra, có thể là sự nhận dạng kênh đôi cho từng dải hệ số tỷ lệ và dải hệ số tỷ lệ 1, ví dụ, là trong phép biểu diễn thứ nhất (như L/R) và trong dải hệ số tỷ lệ thứ ba

có thể trong phép biểu diễn kênh đôi thứ hai như M/S hoặc dự báo trộn giảm/dư. Do đó, bước 812 có thể thu được những phép biểu diễn khác nhau cho các dải khác nhau. Tiếp đó, trong bước 814, bộ tái tạo tần số 806 được tạo cấu hình để lựa chọn phạm vi nguồn cho việc tái tạo tần số. Trong bước 816, bộ tái tạo tần số 806 sau đó kiểm tra phép biểu diễn của phạm vi nguồn và trong khối 818, bộ tái tạo tần số 806 so sánh phép biểu diễn kênh đôi của phạm vi nguồn với phép biểu diễn kênh đôi của khoảng đích. Nếu cả hai phép biểu diễn giống hệt nhau, bộ tái tạo tần số 806 cung cấp sự tái tạo tần số phân tách cho từng kênh của tín hiệu kênh đôi. Tuy nhiên, khi cả hai phép biểu diễn như đã phát hiện trong khối 818 không giống nhau, thì dòng tín hiệu 824 được giữ và khối 822 tính toán phép biểu diễn kênh đôi khác từ phạm vi nguồn và sử dụng phép biểu diễn kênh đôi khác được tính toán này cho sự tái tạo khoảng đích. Do đó, bộ giải mã theo Fig.8a tạo khả năng có thể tái tạo khoảng đích được biểu thị như có sự nhận dạng kênh đôi thứ hai sử dụng phạm vi nguồn trong phép biểu diễn kênh đôi thứ nhất. Tất nhiên, việc bổ sung theo sáng chế cho phép tái tạo khoảng đích sử dụng phạm vi nguồn có sự nhận dạng kênh đôi giống nhau. Và, thêm vào đó, sáng chế cho phép tái tạo khoảng đích có việc nhận dạng kênh đôi biểu thị phép biểu diễn kênh đôi ghép và sau đó biến đổi phép biểu diễn này thành phép biểu diễn kênh phân tách được yêu cầu để lưu trữ hoặc truyền tới loa phát tương ứng cho tín hiệu kênh đôi.

Cần nhấn mạnh rằng kênh đôi của phép biểu diễn kênh đôi có thể là kênh đôi âm lập thể chẳng hạn như kênh trái và kênh phải. Tuy nhiên, tín hiệu có thể cũng là tín hiệu đa kênh, ví dụ, năm kênh và kênh loa âm trầm phụ hoặc thậm chí nhiều kênh hơn nữa. Sau đó, việc xử lý kênh đôi theo cặp như đã thảo luận trong phạm vi của các hình vẽ từ Fig.8a đến Fig.8e có thể được thực hiện tại các cặp có thể, ví dụ, kênh trái và kênh phải, kênh bao quanh trái và kênh bao quanh phải, kênh trung tâm và kênh LFE (loa trầm phụ). Sự tạo cặp bất kỳ khác có thể sử dụng để biểu diễn, ví dụ, sáu kênh đầu vào bởi ba quy trình xử lý kênh đôi.

Fig.8d minh họa sơ đồ khối của bộ giải mã theo sáng chế tương ứng với Fig.8a. Phạm vi nguồn hoặc bộ giải mã lõi 830 có thể tương ứng với bộ giải mã âm thanh 802. Các khối khác 832, 834, 836, 838, 840, 842 và 846 có thể là các phần của bộ tái tạo tần số 806 của Fig.8a. Cụ thể, khối 832 là bộ biến đổi biểu diễn để biến đổi các phép biểu diễn phạm vi nguồn trong từng dải riêng biệt sao cho, tại đầu ra của khối 832, bộ

hoàn thiện của phạm vi nguồn một mặt là trong phép biểu diễn thứ nhất mặt khác là trong phép biểu diễn kênh đôi thứ hai. Hai phép biểu diễn phạm vi nguồn hoàn chỉnh này có thể được lưu trữ trong bộ lưu 834 cho cả hai phép biểu diễn của phạm vi nguồn.

Sau đó, khối 836 áp dụng việc tạo ra ô tần số sử dụng, như đầu vào, phạm vi nguồn ID và sử dụng bổ sung như đầu vào kênh đôi ID cho khoảng đích. Dựa trên kênh đôi ID cho khoảng đích, bộ tạo ra ô tần số truy cập vào bộ lưu trữ 834 và nhận phép biểu diễn kênh đôi của phạm vi nguồn phù hợp với kênh đôi ID để đầu vào khoảng đích vào bộ tạo ra ô tần số 835. Do đó, khi kênh đôi ID cho khoảng đích biểu thị việc xử lý âm lập thể chung, thì bộ tạo ra ô tần số 836 truy cập vào bộ lưu trữ 834 để thu được phép biểu diễn âm lập thể chung của phạm vi nguồn được biểu thị bởi phạm vi nguồn ID 833.

Bộ tạo ra ô tần số 836 thực hiện hoạt động này cho từng khoảng đích và đầu ra của bộ tạo ra ô tần số sao cho là từng kênh của phép biểu diễn kênh được biểu thị bởi sự nhận dạng kênh đôi được hiển thị. Tiếp đó, việc điều chỉnh đường bao bởi bộ điều chỉnh đường bao 838 được thực hiện. Việc điều chỉnh đường bao được thực hiện trong miền kênh đôi được nhận dạng bởi sự nhận dạng kênh đôi. Để đạt được điều này, cần các tham số điều chỉnh đường bao và các tham số này được truyền từ bộ mã hóa tới bộ giải mã trong phép biểu diễn kênh đôi tương tự như đã mô tả. Khi sự nhận dạng kênh đôi trong khoảng đích được xử lý bởi bộ điều chỉnh đường bao có sự nhận dạng kênh đôi biểu thị phép biểu diễn kênh đôi khác với dữ liệu đường bao của khoảng đích này, thì bộ biến đổi theo tham số 840 biến đổi tham số đường bao thành phép biểu diễn kênh đôi cần thiết. Ví dụ, khi sự nhận dạng kênh đôi cho một dải biểu thị mã hóa âm lập thể chung và khi các tham số cho khoảng đích này được truyền dưới dạng tham số đường bao L/R, thì bộ biến đổi theo tham số tính toán các tham số đường bao âm lập thể chung từ tham số đường bao L/R như đã mô tả sao cho phép biểu diễn theo tham số chính xác được sử dụng để điều chỉnh đường bao phổ của khoảng đích.

Trong phương pháp ưu tiên khác các tham số đường bao được truyền như tham số âm lập thể chung khi âm lập thể chung được sử dụng trong dải đích.

Khi giả định rằng đầu vào vào bộ điều chỉnh đường bao 838 là tập hợp các khoảng đích có các phép biểu diễn kênh đôi khác nhau, thì đầu ra của bộ điều chỉnh

đường bao 838 là tập hợp các khoảng đích cũng trong các phép biểu diễn kênh đôi khác nhau. Khi khoảng đích có phép biểu diễn được ghép như M/S, thì khoảng đích này được xử lý bởi bộ biến đổi phép biểu diễn 842 để tính toán phép biểu diễn phân tách được yêu cầu cho việc lưu trữ hoặc truyền tới loa phát. Tuy nhiên, khi khoảng đích có phép biểu diễn phân tách, dòng tín hiệu 844 được giữ và bộ biến đổi 842 được bỏ qua. Tại đầu ra của khối 842, phép biểu diễn phổ kênh đôi trở thành phép biểu diễn kênh đôi phân tách thu được mà có thể sau đó được xử lý thêm như được biểu thị bởi khối 846, việc xử lý thêm này có thể, ví dụ là sự chuyển đổi tần số/thời gian hoặc là các xử lý cần thiết khác.

Tốt hơn là, các phần phổ thứ hai tương ứng với các dải tần, và việc nhận dạng kênh đôi được cung cấp như dãy tín hiệu tương ứng với bảng của Fig.8b, tại đó một tín hiệu cho mỗi dải tần tồn tại. Sau đó, bộ giải mã theo tham số được tạo cấu hình để kiểm tra có hay không dấu hiệu được thiết lập và để điều khiển bộ tái tạo tần số 106 phù hợp với dấu hiệu để sử dụng hoặc là phép biểu diễn thứ nhất hoặc phép biểu diễn thứ hai của phần phổ thứ nhất.

Trong phương án, chỉ phạm vi khôi phục bắt đầu với tần số bắt đầu IGF 309 của Fig.3a có các nhận dạng kênh đôi cho các dải khôi phục khác nhau. Trong phương án khác, điều này cũng được áp dụng dưới cho khoảng tần số dưới tần số bắt đầu IGF 309.

Trong phương án khác, việc nhận dạng dải nguồn và nhận dạng dải đích có thể được xác định thích ứng bằng phân tích tương tự. Tuy nhiên, việc xử lý kênh đôi theo sáng chế có thể luôn được áp dụng khi có sự kết hợp cố định của phạm vi nguồn tới khoảng đích. Phạm vi nguồn có thể được sử dụng để tái tạo lại với tần số, khoảng đích lớn hơn hoặc là bởi hoạt động điền đầy ô tần số sóng hài hoặc hoạt động điền đầy ô tần số sao chép tăng sử dụng hai hoặc nhiều hơn hoạt động điền đầy ô tần số tương tự với việc xử lý nhiều và lỗi đã biết từ việc xử lý AAC có hiệu quả cao.

Fig.8e minh họa bộ mã hóa âm thanh để mã hóa tín hiệu âm thanh kênh đôi. Bộ mã hóa bao gồm bộ chuyển đổi thời gian - phổ 860 để chuyển đổi tín hiệu âm thanh kênh đôi thành phép biểu diễn phổ. Hơn nữa, bộ phân tích phổ 866 để chuyển đổi tín hiệu âm thanh kênh đôi thành phép biểu diễn phổ. Ngoài ra, bộ phân

tích phổ 866 được cung cấp để thực hiện việc phân tích để xác định, các phần phổ được mã hóa với độ phân giải cao, cụ thể, để tìm ra tập hợp thứ nhất của phần phổ thứ nhất và tìm bổ sung tập hợp thứ hai của các phần phổ thứ hai.

Ngoài ra, bộ phân tích kênh đôi 864 được cung cấp để phân tích tập hợp thứ hai của các phần phổ thứ hai để xác định việc nhận dạng kênh đôi nhận dạng hoặc phép biểu diễn kênh đôi thứ nhất hoặc phép biểu diễn kênh đôi thứ hai.

Tùy thuộc vào kết quả của bộ phân tích kênh đôi, dải trong phép biểu diễn phổ thứ hai hoặc là được tham số hóa sử dụng phép biểu diễn kênh đôi thứ nhất hoặc phép biểu diễn kênh đôi thứ hai, và điều này được thực hiện bằng bộ mã hóa theo tham số 868. Khoảng tần số lõi, tức là, dải tần dưới tần số bắt đầu IGF 309 của Fig.3a được mã hóa bởi bộ mã hóa lõi 870. Kết quả của khối 868 và 870 là đầu vào vào giao diện đầu ra 872. Như đã biểu thị, bộ phân tích kênh đôi cung cấp sự nhận dạng kênh đôi cho mỗi dải trên tần số bắt đầu IGF hoặc cho toàn bộ khoảng tần số, và việc nhận dạng kênh đôi này luôn chuyển tới giao diện đầu ra 872 sao cho dữ liệu này cũng được chứa trong đầu ra tín hiệu được mã hóa 873 bởi giao diện đầu ra 872.

Ngoài ra, tốt hơn là bộ mã hóa âm thanh bao gồm bộ biến đổi theo dải 862. Dựa trên sự quyết định của bộ phân tích kênh đôi 862, tín hiệu đầu ra của bộ chuyển đổi thời gian phổ 862 được biến đổi thành phép biểu diễn được biểu thị bởi bộ phân tích kênh đôi và, cụ thể, bởi ID kênh đôi 835. Do đó, đầu ra của bộ biến đổi theo dải 862 là bộ của các dải tần nơi mà từng dải tần có thể hoặc là trong phép biểu diễn kênh đôi thứ nhất hoặc là phép biểu diễn kênh đôi khác nhau thứ hai. Khi sáng chế được áp dụng trong toàn dải, tức là, khi phạm vi nguồn và phạm vi khôi phục đều được xử lý bởi bộ biến đổi theo dải, bộ phân tích phổ 860 có thể phân tích phép biểu diễn này. Tuy nhiên, ngoài ra, bộ phân tích phổ 860 có thể cũng phân tích tín hiệu xuất ra bởi bộ biến đổi thời gian phổ như đã biểu thị bằng đường điều khiển 861. Do đó, bộ phân tích phổ 860 có thể hoặc là áp dụng việc phân tích âm được ưu tiên trên đầu ra của bộ biến đổi theo dải 862 hoặc đầu ra của bộ chuyển đổi thời gian phổ 860 trước khi được xử lý bởi bộ biến đổi theo dải 862. Ngoài ra, bộ phân tích phổ có thể áp dụng việc nhận dạng của phạm vi nguồn phù hợp nhất cho khoảng đích nhất định hoặc là trên kết quả của bộ biến đổi theo dải 862 hoặc trên kết quả của bộ biến đổi thời gian-phổ 860.

Tiếp theo, tham chiếu tới các hình vẽ từ Fig.9a đến Fig.9d minh họa việc tính toán được ưu tiên của các giá trị thông tin năng lượng đã thảo luận trong phạm vi của Fig.3a và Fig.3b.

Tình trạng hiện tại của bộ mã hóa âm thanh trong kỹ thuật áp dụng các kỹ thuật khác nhau để tối thiểu hóa lượng dữ liệu biểu diễn tín hiệu âm thanh đã có. Bộ mã hóa âm thanh như USAC [1] áp dụng sự biến đổi thời gian thành tần số như MDCT để có phép biểu diễn phổ của tín hiệu âm thanh đã có. Các hệ số MDCT này được lượng tử hóa khai thác phương diện tâm thính học của hệ thống thính giác con người. Nếu tốc độ bit có sẵn được giảm sự lượng tử hóa nhận được số lượng lớn đưa vào thô hơn của giá trị phổ được đưa về không dẫn tới thành phần âm thanh lạ tại phía bộ giải mã. Để cải thiện chất lượng cảm giác, bộ giải mã theo tình trạng kỹ thuật điền đầy các phần phổ đã được đưa về không này với nhiễu âm ngẫu nhiên. Phương pháp IGF thu được các ô từ tín hiệu không phải 0 còn lại để điền đầy khoảng trống trong phổ. Điều này quyết định cho chất lượng cảm giác của tín hiệu âm thanh được giải mã để đường bao phổ và việc phân bổ năng lượng của hệ số phổ được bảo toàn. Phương pháp điều chỉnh năng lượng được duy trì ở đây sử dụng thông tin phụ đã biến đổi để khôi phục đường bao phổ MDCT của tín hiệu âm thanh.

Trong eSBR [15] tín hiệu âm thanh được lấy mẫu giảm ít nhất bởi lũy thừa hai và phần tần số cao của phổ được đưa về không một cách hoàn chỉnh [1, 17]. Phần được xóa này được thay thế bởi kỹ thuật theo tham số, eSBR, ở phía bộ giải mã, eSBR bao hàm việc sử dụng của biến đổi bổ sung, việc biến đổi QMF được sử dụng để thay thế phần tần số cao rộng và để lấy lại mẫu tín hiệu âm thanh [17]. Việc này bổ sung sự phức tạp tính toán và mức tiêu hao bộ nhớ cho bộ mã hóa âm thanh.

Bộ mã hóa USAC [15] thường có khả năng điền đầy các lỗ trống phổ (các đường phổ đã đưa về không) với nhiễu âm ngẫu nhiên nhưng lại có các bất lợi sau đây: nhiễu âm ngẫu nhiên không thể bảo toàn cấu trúc tốt theo thời gian của tín hiệu chuyển tiếp và nó không thể bảo toàn cấu trúc sóng hài của tín hiệu âm.

Khu vực mà eSBR hoạt động ở phía bộ giải mã được xóa hoàn toàn bởi bộ mã hóa [1]. Do đó eSBR dễ xóa các đường âm trong vùng tần số cao hoặc cấu trúc sóng hài biến dạng của tín hiệu gốc. Vì độ phân giải tần số QMF của eSBR rất thấp và độ

phục hồi của thành phần dạng sin chỉ có khả năng trong độ phân giải thô của giàn lọc ở dưới, việc tái tạo các thành phần âm trong eSBR trong khoảng tần số lặp lại có độ chính xác rất thấp.

eSBR sử dụng các kỹ thuật để điều chỉnh năng lượng của khu vực được vá lỗi, việc điều chỉnh đường bao phổ [1]. Kỹ thuật này sử dụng giá trị năng lượng đã truyền trên lưới thời gian tần số QMF để tái tạo hình đường bao phổ. Tình trạng kỹ thuật không điều khiển một phần bằng tay được xóa phổ và do độ phân giải thời gian cao hoặc có thể cần lượng dữ liệu tương đối lớn của bit để truyền giá trị năng lượng thích hợp hoặc áp dụng việc lượng tử hóa thô cho giá trị năng lượng.

Phương pháp IGF không cần sự biến đổi bổ sung vì nó sử dụng việc sự biến đổi MDCT kế thừa mà được tính toán như mô tả trong [15].

Phương pháp điều chỉnh năng lượng được đưa ra ở đây sử dụng thông tin phụ được tạo ra bởi bộ mã hóa để khôi phục đường bao phổ của tín hiệu âm thanh. Thông tin phụ này được tạo ra bởi bộ mã hóa như phác họa dưới đây:

a) Áp dụng sự biến đổi MDCT được tạo cửa sổ cho tín hiệu âm thanh đầu vào [16, phần 4.6], tùy ý tính toán MDCT được tạo cửa sổ, hoặc ước tính MDCT được tạo cửa sổ từ MDCT được tính toán.

b) Áp dụng TNS/TTS trên hệ số MDCT [15, phần 7.8].

c) Tính toán năng lượng trung bình cho mọi dải hệ số tỷ lệ phía trên tần số bắt đầu IGF ($f_{IGFstart}$) tới tần số dừng IGF ($f_{IGFstop}$)

d) Lượng tử hóa các giá trị năng lượng trung bình

$f_{IGFstart}$ và $f_{IGFstop}$ là các tham số đã cho người sử dụng.

Các giá trị đã tính toán từ bước c) và d) được mã hóa không mất dữ liệu và được truyền dưới dạng thông tin phụ với dòng bit đến bộ giải mã.

Bộ giải mã nhận giá trị đã truyền và sử dụng chúng để điều chỉnh đường bao phổ.

a) Khử lượng tử hóa các giá trị MDCT đã truyền

b) Áp dụng điền đầy nhiễu âm USAC kế thừa nếu được tín hiệu hóa

- c) Áp dụng điền đầy ô IGF
- d) Khử lượng tử hóa các giá trị năng lượng đã truyền
- e) Điều chỉnh theo dải hệ số tỷ lệ đường bao phổ
- f) Áp dụng TNS/TTS nếu được tín hiệu hóa

Để $\hat{x} \in \mathbb{R}^N$ là MDCT được biến đổi, phép biểu diễn phổ giá trị thực của tín hiệu âm thanh được tạo cửa sổ của chiều dài cửa sổ $2N$. Việc biến đổi này được mô tả trong [16]. Bộ mã hóa áp dụng tùy ý TNS trên $\hat{x}$.

Trong các tài liệu [16, 4.6.2] sự phân bố của $\hat{x}$ trong các dải hệ số tỷ lệ được mô tả. Các dải hệ số tỷ lệ là tập hợp của bộ các hệ số và được biểu thị và kí hiệu trong bản mô tả này là scb_k .

Giới hạn của mỗi scb_k with $k = 0, 1, 2, \dots, max_sfb$ được xác định bởi dãy sub_offset (16, 4.6.2), trong đó $swb_offset[k]$ và $swb_offset[k+1]-1$ xác định chỉ số thứ nhất và cuối cùng cho đường hệ số phổ thấp nhất và cao nhất có trong scb_k . Biểu thị dải hệ số tỷ lệ

$$scb_k := swb_offset[k], 1+swb_offset[k], 2+swb_offset[k], \dots, swb_offset[k+1]-1\}$$

Nếu công cụ IGF được sử dụng bởi bộ mã hóa, người sử dụng xác định tần số bắt đầu IGF và tần số dừng IGF. Hai giá trị này ánh xạ lên chỉ số dải hệ số tỷ lệ phù hợp nhất $igfStartSfb$ và $igfStopSfb$. Cả hai được tạo tín hiệu trong dòng truyền dữ liệu tới bộ giải mã.

Tài liệu [16] mô tả cả sự biến đổi khối dài và khối ngắn. Với khối dài chỉ duy nhất một bộ hệ số phổ cùng với một bộ hệ số tỷ lệ được truyền tới bộ giải mã. Để khối ngắn tám cửa sổ ngắn với tám bộ khác nhau của hệ số phổ được tính toán. Để lưu tốc độ bit, các hệ số tỷ lệ của tám cửa sổ khối ngắn được nhóm lại bằng bộ mã hóa.

Trong trường hợp của IGF phương pháp được thể hiện ở đây sử dụng các dải hệ số tỷ lệ kế thừa để nhóm các giá trị phổ mà được truyền tới bộ giải mã.

$$E_k = \sqrt{\frac{1}{|scb_k|} \sum_{i \in scb_k} \hat{x}_i^2}$$

Trong đó $k = igfStartSfb, 1 + igfStartSfb, 2 + igfStartSfb, \dots, igfEndSfb$.

Với việc lượng tử hóa

$$\tilde{E}_k = nINT(4\log_2(E_k))$$

được tính toán. Tất cả giá trị $\tilde{E}_k$ được truyền tới bộ giải mã.

Giả định rằng bộ mã hóa này quyết định để nhóm num_window_group các tập hợp hệ số tỷ lệ.

Ta biểu thị với w phần nhóm này của tập hợp $\{0, 1, 2, \dots, 7\}$ là kí hiệu của tám cửa sổ ngắn. w_l biểu thị tập con của w thứ l , trong đó l biểu thị chỉ số của nhóm cửa sổ, $0 \leq l < num_window_group$.

Để việc tính toán khối ngắn người sử dụng xác định tần số bắt đầu/dừng IGF được ánh xạ lên dải hệ số tỷ lệ thích hợp. Tuy nhiên, để đơn giản, cũng biểu thị cho các khối ngắn $k = igfStartSfb, 1 + igfStartSfb, 2 + igfStartSfb, \dots, igfEndSfb$.

Việc tính toán năng lượng IGF sử dụng thông tin nhóm để nhóm giá trị $E_{k,l}$

$$E_{k,l} = \sqrt{\frac{1}{|w_l|} \sum_{j \in w_l} \frac{1}{|scb_k|} \sum_{i \in scb_k} \hat{x}_{j,i}^2}$$

Với việc lượng tử hóa

$$\tilde{E}_{k,l} = nINT(4\log_2(E_{k,l}))$$

được tính toán. Tất cả giá trị $\tilde{E}_{k,l}$ được truyền tới bộ giải mã.

Các công thức mã hóa được đề cập ở trên hoạt động chỉ sử dụng các hệ số MDCT có giá trị thực $\hat{x}$. Để thu được sự phân bố năng lượng ổn định hơn trong khoảng IGF, nghĩa là, để giảm dao động biên độ theo thời gian, một phương pháp thay thế có thể được sử dụng để tính toán các giá trị $\tilde{E}_k$:

Để $\hat{x}_r \in \mathbb{R}^N$ được biến đổi MDCT, phép biểu diễn phổ có giá trị thực của tín hiệu âm thanh được tạo cửa sổ của cửa sổ dài $2N$, và $\hat{x}_i \in \mathbb{R}^N$ phép biểu diễn phổ được biến đổi MDCT có giá trị thực của phần tương tự của tín hiệu âm thanh. Phép biểu diễn phổ MDCT $\hat{x}_i$ có thể hoặc là được tính toán chính xác hoặc được ước tính từ $\hat{x}_r$.

$\hat{c} = (\hat{x}_r, \hat{x}_i) \in \mathbb{C}^N$ biểu thị phép biểu diễn phổ phức hợp của tín hiệu âm thanh cửa sổ, có $\hat{x}_r$ làm phần thực và $\hat{x}_i$ làm phần ảo. Bộ mã hóa áp dụng tùy ý TNS trên $\hat{x}_r$ và $\hat{x}_i$.

Lúc này năng lượng của tín hiệu gốc trong khoảng IGF có thể được đo bởi

$$E_{ok} = \frac{1}{|scb_k|} \sum_{i \in scb_k} \hat{c}_i^2$$

Năng lượng có giá trị thực và năng lượng có giá trị phức hợp của dải khôi phục, nghĩa là, ô sẽ được sử dụng ở phía bộ giải mã trong việc khôi phục khoảng IGF scb_k được tính toán là:

$$E_{tk} = \frac{1}{|scb_k|} \sum_{i \in tr_k} \hat{c}_i^2, \quad E_{rk} = \frac{1}{|scb_k|} \sum_{i \in tr_k} \hat{x}_{r_i}^2$$

Trong đó tr_k là tập hợp biểu thị phạm vi ô nguồn được kết hợp phụ thuộc vào scb_k . Trong hai công thức trên, thay vì tập hợp chỉ số scb_k , tập hợp $\overline{scb_k}$ (được xác định sau) có thể được sử dụng để tạo tr_k để đạt được các giá trị chính xác hơn E_r và E_r .

Tính toán

$$f_k = \frac{E_{ok}}{E_{rk}}$$

nếu $E_{rk} > 0$, ngược lại $f_k = 0$.

Với

$$E_k = \sqrt{f_k E_{rk}}$$

ngay sau đây bản ổn định hơn của E_k sẽ được tính toán, do sự tính toán của E_k với các giá trị MDCT chỉ được làm giảm bởi thực tế là các giá trị MDCT không tuân theo định lý của Parseval, và do đó chúng không phản ánh thông tin năng lượng hoàn chỉnh của các giá trị phổ. $\tilde{E}_k$ được tính toán như ở trên.

Như lưu ý gần đây, với các khối ngăn giả sử rằng bộ mã hóa quyết định để nhóm các bộ hệ số tỷ lệ *num_window_group*. Như nêu ở trên, w_l biểu thị là tập con của w thứ l , trong đó l biểu thị chỉ số của nhóm cửa sổ, $0 \leq l < \text{num_window_group}$.

Một lần nữa, bản thay thế đã phác thảo ở trên để tính toán bản ổn định hơn của $E_{k,l}$ có thể được tính toán. Với việc xác định của $\hat{c} = (\hat{x}_r, \hat{x}_i) \in \mathbb{C}^N$, $\hat{x}_r \in \mathbb{R}^N$ là MCDT

được biến đổi và $\hat{x}_e \in \mathbb{R}^N$ là tín hiệu âm thanh của số MDCT đã biến đổi của độ dài $2N$, tính toán

$$E_{ok,l} = \frac{1}{|w_l|} \sum_{i \in w_l} \frac{1}{|scb_k|} \sum_{i \in scb_k} \hat{c}_{i,l}^2$$

Tương tự tính

$$E_{rk,l} = \frac{1}{|w_l|} \sum_{i \in w_l} \frac{1}{|scb_k|} \sum_{i \in tr_k} \hat{c}_{i,l}^2, \quad E_{rk,l} = \frac{1}{|w_l|} \sum_{i \in w_l} \frac{1}{|scb_k|} \sum_{i \in tr_k} \hat{x}_{r,l}^2$$

và tiếp tục với hệ số $f_{k,l}$

$$f_{k,l} = \frac{E_{ok,l}}{E_{rk,l}}$$

mà được sử dụng để điều chỉnh các tính toán trước đó $E_{rk,l}$:

$$E_{k,l} = \sqrt{f_{k,l} E_{rk,l}}$$

$\tilde{E}_{k,l}$ được tính toán như ở trên.

Quy trình không chỉ sử dụng năng lượng của dải khôi phục có nguồn gốc từ dải khôi phục phức hợp hoặc từ các giá trị MDCT, mà còn sử dụng thông tin năng lượng từ phạm vi nguồn cung cấp sự khôi phục năng lượng cải thiện hơn.

Đặc biệt, bộ tính toán theo tham số 1006 được tạo cấu hình để tính toán thông tin năng lượng cho dải khôi phục bằng cách sử dụng thông tin về năng lượng của dải khôi phục và sử dụng thêm thông tin về năng lượng của phạm vi nguồn để sử dụng cho việc khôi phục dải khôi phục.

Hơn nữa, bộ tính toán theo tham số 1006 được tạo cấu hình để tính toán thông tin năng lượng (E_{ok}) trên dải khôi phục của phổ phức hợp của tín hiệu gốc, để tính toán thông tin năng lượng thêm (E_{rk}) trên phạm vi nguồn của phần giá trị thực của phổ phức hợp trong tín hiệu gốc được sử dụng cho việc khôi phục dải khôi phục, và trong đó bộ tính toán theo tham số được tạo cấu hình để tính toán thông tin năng lượng cho dải khôi phục sử dụng thông tin năng lượng (E_{ok}) và thông tin năng lượng thêm (E_{rk}).

Ngoài ra, bộ tính toán theo tham số 1006 được tạo cấu hình để xác định thông tin năng lượng thứ nhất (E_{ok}) về dải hệ số tỷ lệ đã khôi phục của phổ phức hợp của trong

tần số gốc, để xác định thông tin năng lượng thứ hai (E_{tk}) về phạm vi nguồn của phổ phức hợp của tín hiệu gốc được sử dụng để khôi phục dải hệ số tỷ lệ đã khôi phục, để xác định thông tin năng lượng thứ ba (E_{rk}) về phạm vi nguồn của phần giá trị thực của phổ phức hợp của tín hiệu gốc được sử dụng để khôi phục dải hệ số tỷ lệ đã khôi phục, để xác định thông tin trọng số dựa trên mối liên hệ giữa ít nhất hai phần của thông tin năng lượng thứ nhất, thông tin năng lượng thứ hai, và thông tin năng lượng thứ ba, và để gán trọng số một trong số thông tin năng lượng thứ nhất và thông tin năng lượng thứ ba sử dụng thông tin gán trọng số để thu được thông tin năng lượng có trọng số và sử dụng thông tin năng lượng có trọng số làm thông tin năng lượng dải khôi phục.

Các ví dụ cho việc tính toán như sau, nhưng nhiều cách khác có thể xuất hiện với người có kỹ năng trong lĩnh vực kỹ thuật được thể hiện của nguyên lý thông thường ở trên:

A)

$$f_k = E_{ok}/E_{tk};$$

$$E_k = \text{sqrt}(f_k * E_{rk});$$

B)

$$f_k = E_{tk}/E_{ok};$$

$$E_k = \text{sqrt}((1/f_k) * E_{rk});$$

C)

$$f_k = E_{rk}/E_{tk};$$

$$E_k = \text{sqrt}(f_k * E_{ok})$$

D)

$$f_k = E_{tk}/E_{rk};$$

$$E_k = \text{sqrt}((1/f_k) * E_{ok})$$

Tất cả các ví dụ này xác nhận thực tế là mặc dù chỉ các giá trị MDCT thực được xử lý ở phía bộ giải mã, việc tính toán thực tế do sự phù hợp và bổ sung của miền thời gian tức là xóa bỏ hoàn toàn cách thức được tạo sử dụng các số phức hợp. Tuy nhiên,

cụ thể, việc xác định 918 của thông tin năng lượng ô của các phần phổ thêm 922, 923 của dải khôi phục 920 cho các giá trị tần số khác phần phổ thứ nhất 921 có các tần số trong dải khôi phục 920 dựa vào giá trị MDCT thực. Do đó, thông tin năng lượng đã truyền tới bộ giải mã sẽ nhỏ hơn thông tin năng lượng E_{ok} về dải khôi phục của phổ phức hợp của tín hiệu gốc. Ví dụ cho trường hợp C ở trên, điều này có nghĩa là yếu tố f_k (thông tin trọng số) có thể nhỏ hơn 1.

Ở phía bộ giải mã, nếu công cụ IGF được tạo tín hiệu bất, giá trị thông tin $\hat{E}_k$ thu được từ đường truyền dữ liệu và sẽ bị khử lượng tử hóa với

$$E_k = 2^{\frac{1}{4}\hat{E}_k}$$

đối với mọi $k = igfStartSfb, 1 + igfStartSfb, 2 + igfStartSfb, \dots, igfEndSfb$.

Bộ giải mã khử lượng tử hóa các giá trị MDCT đã truyền tới $x \in \mathbb{R}^N$ và tính toán năng lượng tồn tại còn lại :

$$sE_k := \sum_{i \in scb_k} x_i^2$$

Trong đó k nằm trong khoảng như được xác định ở trên.

Ký hiệu $\overline{scb_k} = \{i | i \in scb_k \wedge x_i = 0\}$. Tập hợp này bao gồm tất cả các chỉ số của dải hệ số tỷ lệ scb_k mà hệ số được lượng tử hóa thành không bởi bộ mã hóa.

IGF có phương pháp băng con (không được mô tả ở đây) được sử dụng để điền đầy khoảng trống phổ thu được từ sự lượng tử hóa thô của các giá trị phổ MDCT tại phía bộ mã hóa bằng việc sử dụng giá trị khác không của MDCT đã truyền. x sẽ bao gồm thêm các giá trị mà thay thế tất cả các giá trị đến không trước. Năng lượng ô được tính toán bởi:

$$tE_k := \sum_{i \in \overline{scb_k}} x_i^2$$

trong đó k nằm trong khoảng như được xác định ở trên.

Năng lượng bị thiếu trong dải khôi phục được tính bằng:

$$mE_k := |scb_k| E_k^2 - sE_k$$

Và hệ số khuếch đại cho việc điều chỉnh thu được bởi:

$$g := \begin{cases} \sqrt{\frac{mE_k}{tE_k}} & \text{if } (mE_k > 0 \wedge tE_k > 0) \\ 0 & \text{else} \end{cases}$$

Với

$$g' = \min(g, 10)$$

Việc điều chỉnh đường bao phổ sử dụng hệ số khuếch đại là:

$$x_i := g' x_i$$

đối với mọi $i \in \overline{scb_k}$ và k nằm trong khoảng như đã xác định ở trên.

Điều này khôi phục hình dạng đường bao phổ của x thành tạo hình đường bao phổ ban đầu x .

Với chuỗi cửa sổ ngắn tất cả các tính toán như đã phác họa ở trên với cùng nguyên lý, nhưng việc nhóm các dải hệ số tỷ lệ được xem xét. Kí hiệu $E_{k,l}$ là các giá trị năng lượng đã lượng tử hóa, đã được nhóm thu được từ dòng bit. Tính toán

$$sE_{k,l} := \frac{1}{|w_l|} \sum_{j \in w_l} \sum_{i \in scb_{j,k}} x_{j,i}^2$$

và

$$sE_{k,l} := \frac{1}{|w_l|} \sum_{j \in w_l} \sum_{i \in scb_{j,k}} x_{j,i}^2$$

Chỉ số j mô tả chỉ số cửa sổ của chuỗi khối ngắn.

Tính toán

$$mE_{k,l} := |scb_k| E_{k,l}^2 - sE_{k,l}$$

Và

$$g := \begin{cases} \sqrt{\frac{mE_{k,l}}{pE_{k,l}}} & \text{if } (mE_{k,l} > 0 \wedge pE_{k,l} > 0) \\ 0 & \text{else} \end{cases}$$

Với

$$g' = \min(g, 10)$$

Áp dụng

$$x_{j,i} = g^i x_{j,i}$$

đối với mọi $i \in \overline{scb_{k,l}}$.

Với việc áp dụng tốc độ bit thấp nhóm từng cặp giá trị E_k có thể không mất quá nhiều độ chính xác. Phương pháp này chỉ được áp dụng với các khối dài:

$$E_{k \gg 1} = \sqrt{\frac{1}{|scb_k \cup scb_{k+1}|} \sum_{i \in scb_k \cup scb_{k+1}} \hat{x}_i^2}$$

trong đó $k = igfStartSfb, 2 + igfStartSfb, 4 + igfStartSfb, \dots, igfEndSfb$.

Một lần nữa, sau khi lượng tử hóa, tất cả tín hiệu $E_{k \gg 1}$ được truyền tới bộ giải mã.

Fig.9a minh họa thiết bị giải mã tín hiệu âm thanh được mã hóa bao gồm phép biểu diễn được mã hóa của tập hợp thứ nhất của phần phổ thứ nhất và phép biểu diễn được mã hóa của dữ liệu tham số biểu thị năng lượng phổ cho bộ thứ hai của phần phổ thứ hai. Tập hợp thứ nhất của phần phổ thứ nhất đã xác định tại 901a trên Fig.9a, và phép biểu diễn được mã hóa của dữ liệu tham số được biểu thị tại 901b trên Fig.9a. Bộ giải mã âm thanh 900 được cung cấp cho việc giải mã phép biểu diễn được mã hóa 901a của tập hợp thứ nhất của phần phổ thứ nhất để thu tập hợp thứ nhất được giải mã của phần phổ thứ nhất 904 và để giải mã phép biểu diễn được mã hóa của dữ liệu tham số để thu được dữ liệu tham số được giải mã 902 cho tập hợp thứ hai của phần phổ thứ hai biểu thị năng lượng riêng cho dải khôi phục riêng, trong đó phần phổ thứ hai được đặt trong dải khôi phục. Ngoài ra, bộ tái tạo tần số 906 được cung cấp để khôi phục giá trị phổ của dải khôi phục bao gồm phần phổ thứ hai. Bộ tái tạo tần số 906 sử dụng phần phổ thứ nhất của tập hợp thứ nhất của các phần phổ thứ nhất và thông tin năng lượng riêng cho dải khôi phục, trong đó dải khôi phục bao gồm phần phổ thứ nhất và phần phổ thứ hai. Bộ tái tạo tần số 906 bao gồm bộ tính toán 912 để xác định thông tin năng lượng tồn tại bao gồm năng lượng đã tích lũy của phần phổ thứ nhất có tần số trong dải khôi phục. Ngoài ra, bộ tái tạo tần số 906 bao gồm bộ tính toán 918 để xác

định thông tin năng lượng ô của phần phổ khác nữa của dải khôi phục và giá trị tần số khác nhau từ phần phổ thứ nhất, trong đó, giá trị tần số này có tần số trong dải khôi phục, trong đó phần phổ khác nữa được tạo ra bởi việc tái tạo tần số sử dụng phần phổ thứ nhất khác phần phổ thứ nhất trong dải khôi phục.

Bộ tái tạo tần số 906 còn bao gồm bộ tính toán 914 cho năng lượng thiếu trong dải khôi phục, và bộ tính toán 914 hoạt động sử dụng năng lượng riêng cho dải khôi phục và năng lượng còn lại được tạo ra bởi khối 912. Hơn nữa, bộ tái tạo tần số 906 bao gồm bộ điều chỉnh đường bao phổ 916 để điều chỉnh phần phổ khác trong dải khôi phục dựa trên thông tin năng lượng thiếu và thông tin năng lượng ô được tạo ra bởi khối 918.

Tham chiếu tới Fig.9c minh họa dải khôi phục nhất định 920. Dải khôi phục bao gồm phần phổ thứ nhất trong dải khôi phục như phần phổ thứ nhất 306 trên Fig.3a minh họa dưới dạng biểu đồ ở 912. Hơn nữa, phần còn lại của giá trị phổ trong dải khôi phục 920 được tạo ra sử dụng vùng nguồn, ví dụ từ dải hệ số tỷ lệ 1,2,3 dưới tần số bắt đầu điền đầy khoảng trống thông minh 309 của Fig.3a. Bộ tái tạo tần số 906 được tạo cấu hình để tạo ra giá trị phổ thô cho phần phổ thứ hai 922 và 923. Sau đó, hệ số khuếch đại g được tính toán như minh họa trên Fig.9c để cuối cùng điều chỉnh giá trị phổ thô trong dải phổ 922, 923 để thu được phần phổ thứ hai đã khôi phục và điều chỉnh trong dải khôi phục 920 mà giờ đây có độ phân giải phổ tương tự, tức là khoảng cách đường giống nhau như phần phổ thứ nhất 921. Đây là điều quan trọng để hiểu rõ rằng phần phổ thứ nhất trong dải khôi phục đã minh họa ở 921 trên Fig.9c được giải mã bởi bộ giải mã âm thanh 900 và không bị ảnh hưởng bởi khối đã thể hiện điều chỉnh đường bao 916 của Fig.9b. Thay vào đó, phần phổ thứ nhất trong dải khôi phục đã cho biết ở 921 là phía trái như nó có, từ phần phổ thứ nhất này là đầu ra bởi băng thông đầy đủ hoặc bộ giải mã âm thanh tốc độ đầy đủ 900 thông qua đường 904.

Tiếp theo, ví dụ nhất định với số thực được thảo luận. Năng lượng tồn tại như tính toán bởi khối 912 là, ví dụ năm đơn vị năng lượng và năng lượng này là năng lượng của ví dụ đã biểu thị bốn đường phổ trong phần phổ thứ nhất 921.

Ngoài ra, giá trị năng lượng E3 cho dải khôi phục tương ứng với dải hệ số tỷ lệ 6 của Fig.3b hoặc 3a là bằng với 10 đơn vị. Quan trọng là, giá trị năng lượng không chỉ

bao gồm năng lượng của phần phổ 922, 923 mà còn năng lượng đầy đủ của dải khôi phục 920 như đã tính toán ở phía bộ mã hóa, tức là, trước khi thực hiện việc sử dụng phân tích phổ, ví dụ như màn chắn âm. Do đó, mười đơn vị năng lượng che phần phổ thứ nhất và thứ hai trong dải khôi phục. Sau đó, giả sử rằng năng lượng của dữ liệu phạm vi nguồn cho khối 922, 923 hoặc cho dữ liệu khoảng đích thô cho khối 922, 923 là bằng với tám đơn vị năng lượng. Do đó, năng lượng thiếu của năm đơn vị được tính toán.

Dựa trên năng lượng thiếu bị tách ra bởi năng lượng ô tEk, hệ số khuếch đại 0,79 được tính toán. Sau đó, đường phổ thô cho phần phổ thứ hai 922, 923 được nhân lên với hệ số khuếch đại tính toán. Theo đó, chỉ các giá trị phổ cho phần phổ thứ hai 922, 923 được điều chỉnh và đường phổ cho phần phổ thứ nhất 921 không bị ảnh hưởng bởi việc điều chỉnh đường bao này. Tiếp theo việc nhân lên giá trị phổ thô cho phần phổ thứ hai 922, 923, dải khôi phục hoàn chỉnh được tính toán bao gồm phần phổ thứ nhất trong dải khôi phục, và gồm đường phổ trong phần phổ thứ hai 922, 923 trong dải khôi phục 920.

Tốt hơn là, phạm vi nguồn để tạo ra dữ liệu phổ thô trong dải 922, 923 là, với tần số, dưới tần số bắt đầu IGF 309 và dải khôi phục 920 là phía trên tần số bắt đầu IGF 309.

Ngoài ra, ưu tiên các biên dải khôi phục trùng với các biên dải hệ số tỷ lệ. Do đó, dải khôi phục có, trong một phương án, kích thước của sự tương đồng dải hệ số tỷ lệ của bộ giải mã âm thanh lõi có kích thước để, khi việc tạo cặp năng lượng được áp dụng, giá trị năng lượng cho dải khôi phục cung cấp năng lượng của hai hoặc nhiều số nguyên cao hơn của dải hệ số tỷ lệ. Do đó, khi được giả sử rằng việc tích lũy năng lượng được thực hiện cho dải hệ số tỷ lệ 4, dải hệ số tỷ lệ 5 và dải hệ số tỷ lệ 6, sau đó biên tần số thấp hơn của dải khôi phục 920 là bằng với biên thấp hơn của dải hệ số tỷ lệ 4 và biên tần số cao hơn của dải khôi phục 920 trùng với biên của dải cao hơn của hệ số tỷ lệ 6.

Tiếp theo, Fig.9d được thảo luận để thể hiện chức năng khác của bộ giải mã của Fig.9a. Bộ giải mã âm thanh 900 nhận giá trị phổ khử lượng tử hóa tương ứng với phần phổ thứ nhất của tập hợp thứ nhất của phần phổ và, bổ sung thêm, hệ số tỷ lệ cho

dải hệ số tỷ lệ như đã minh họa trên Fig.3b được cung cấp cho khối tỷ lệ nghịch đảo 940. Khối tỷ lệ nghịch đảo 940 cung cấp tất cả các tập hợp thứ nhất của phần phổ thứ nhất dưới tần số bắt đầu IGF 309 của Fig.3a và, bổ sung thêm, phần phổ thứ nhất trên tần số bắt đầu IGF, tức là phần phổ thứ nhất 304, 305, 306, 307 của Fig.3a được đặt tất cả trong dải khôi phục như minh họa ở 941 trên Fig.9d. Ngoài ra, phần phổ thứ nhất trong dải nguồn được sử dụng để điền đầy ô tần số trong dải khôi phục được cung cấp cho bộ điều chỉnh/tính toán đường bao 942 và khối này nhận bổ sung thông tin năng lượng cho dải khôi phục được cung cấp như thông tin phụ tham số tới tín hiệu âm thanh được mã hóa như minh họa ở 943 của Fig.9d. Sau đó, bộ điều chỉnh/tính toán đường bao 942 cung cấp chức năng của Fig.9b và 9c và giá trị phổ đã điều chỉnh đầu ra cuối cùng cho phần phổ thứ hai trong dải khôi phục. Các giá trị đã điều chỉnh này 922, 923 cho phần phổ thứ hai trong dải khôi phục và phần phổ thứ nhất 921 trong dải khôi phục biểu thị đường 941 trên Fig.9d biểu diễn phép biểu diễn phổ hoàn chỉnh của dải khôi phục.

Tiếp theo, tham chiếu tới các hình vẽ từ Fig.10a đến Fig.10b để giải thích các phương án được ưu tiên của bộ mã hóa âm thanh để mã hóa tín hiệu âm thanh để cung cấp hoặc tạo ra tín hiệu âm thanh được giải mã. Bộ giải mã bao gồm bộ chuyển đổi thời gian/phổ 1002 cung cấp bộ phân tích phổ 1004, và bộ phân tích phổ 1004 một mặt được nối với bộ tính toán và mặt khác được nối với bộ mã hóa âm thanh 1008. Bộ mã hóa âm thanh 1008 cung cấp phép biểu diễn được mã hóa của tập hợp thứ nhất của phần phổ thứ nhất và bao gồm tập hợp thứ hai của phần phổ thứ hai. Mặt khác, bộ tính toán theo tham số 1006 cung cấp thông tin năng lượng cho dải khôi phục bao gồm phần phổ thứ nhất và thứ hai. Ngoài ra, bộ mã hóa âm thanh 1008 được tạo cấu hình để tạo ra phép biểu diễn được mã hóa thứ nhất của tập hợp thứ nhất của phần phổ thứ nhất có độ phân giải phổ thứ nhất trong đó bộ mã hóa âm thanh 1008 cung cấp hệ số tỷ lệ cho tất cả các dải của phép biểu diễn phổ được tạo ra bởi khối 1002. Và như minh họa trên Fig.3b, bộ mã hóa cung cấp thông tin năng lượng ít nhất cho dải khôi phục được định vị, đối với tần số, phía trên tần số bắt đầu IGF 309 như minh họa trên Fig.3a. Do đó, để dải khôi phục tốt hơn là trùng với dải hệ số tỷ lệ hoặc với các nhóm hệ số tỷ lệ, hai giá trị được đưa ra,... hệ số tỷ lệ tương ứng từ bộ mã hóa âm thanh 1008 và, bổ sung thêm, thông tin năng lượng cung cấp bởi bộ tính toán theo tham số 1006.

Bộ mã hóa âm thanh tốt hơn là có dải hệ số tỷ lệ với các băng thông tần số khác nhau, tức là, với số lượng khác nhau của giá trị phổ. Theo đó, bộ tính toán theo tham số bao gồm bộ chuẩn hóa 1012 để chuẩn hóa năng lượng cho băng thông khác nhau đối với băng thông của dải khôi phục cụ thể. Để đạt được điều này, bộ chuẩn hóa 1012 nhận, như đầu vào, năng lượng trong dải và số lượng các giá trị phổ trong dải và bộ chuẩn hóa 1012 sau đó cung cấp năng lượng đã chuẩn hóa mỗi dải khôi phục/hệ số tỷ lệ.

Ngoài ra, bộ tính toán theo tham số 1006 của Fig.10a bao gồm bộ tính toán giá trị năng lượng nhận thông tin điều khiển lỗi hoặc bộ mã hóa âm thanh 1008 như minh họa bằng đường 1007 trên Fig.10a. Thông tin điều khiển này có thể bao gồm thông tin về khối dài/ngắn được sử dụng bởi bộ mã hóa âm thanh và/hoặc việc nhóm thông tin. Do đó, trong khi thông tin về các khối dài/ngắn và việc nhóm thông tin về cửa sổ ngắn liên quan với việc nhóm "thời gian", việc nhóm thông tin có thể có liên quan bổ sung với việc nhóm phổ, tức là, việc nhóm hai dải hệ số tỷ lệ thành dải khôi phục đơn. Do đó, bộ tính toán giá trị năng lượng 1014 cung cấp giá trị năng lượng đơn cho từng dải đã được nhóm bao gồm phần phổ thứ nhất và thứ hai khi chỉ phần phổ được nhóm.

Fig.10d minh họa phương án khác để thực hiện việc nhóm phổ. Để đạt được điều này, khối 1016 được tạo cấu hình để tính toán giá trị năng lượng cho hai dải liền kề. Sau đó, trong khối 1018, giá trị năng lượng cho dải liền kề được so sánh và, khi giá trị năng lượng không khác biệt nhiều hơn hoặc khác biệt ít hơn đã xác định bởi, ví dụ, ngưỡng, sau đó giá trị (được chuẩn hóa) đơn cho cả các dải được tạo ra như đã xác định trong khối 1020. Như minh họa bởi đường 1019, khối 1018 có thể được bỏ qua. Ngoài ra, việc tạo ra giá trị đơn cho hai hoặc nhiều dải được thể hiện bởi khối 1020 có thể được điều khiển bởi bộ mã hóa tốc độ bit điều khiển 1024. Do đó, khi tốc độ bit được làm giảm, bộ điều khiển tốc độ bit được mã hóa 1024 điều khiển khối 1020 để tạo ra giá trị chuẩn hóa đơn cho hai hoặc nhiều dải thậm chí dù việc so sánh trong khối 1018 không được cho phép để nhóm giá trị thông tin năng lượng.

Trong trường hợp bộ mã hóa âm thanh thực hiện việc nhóm hai hoặc nhiều cửa sổ, việc nhóm này cũng được áp dụng cho thông tin năng lượng. Khi bộ mã hóa lỗi thực hiện việc nhóm hai hoặc nhiều khối ngắn, thì đối với hai hoặc nhiều khối này, chỉ

tập hợp đơn của các hệ số tỷ lệ được tính toán và được truyền. Ở phía bộ giải mã, bộ giải mã âm thanh sau đó áp dụng tập hợp các hệ số tỷ lệ tương tự cho cả hai cửa sổ được nhóm.

Đối với việc tính toán thông tin năng lượng, giá trị phổ trong dải khôi phục được tích lũy trên hai hoặc nhiều cửa sổ ngắn. Nói cách khác, có nghĩa là giá trị phổ trong dải khôi phục nhất định cho khối ngắn và cho khối ngắn tiếp theo được tích lũy với nhau và chỉ thông tin năng lượng đơn được truyền cho dải khôi phục này chứa khối ngắn. Sau đó, ở phía bộ giải mã, việc điều chỉnh đường bao được thảo luận đối với các hình vẽ từ Fig.9a đến Fig.9d không được thực hiện riêng rẽ cho từng khối ngắn mà được thực hiện cùng nhau cho tập hợp các cửa sổ ngắn đã nhóm.

Việc chuẩn hóa tương ứng sau đó được áp dụng một lần nữa để mặc dù đã thực hiện bất kỳ việc phân nhóm tần số hoặc nhóm thời gian nào, việc chuẩn hóa dễ dàng cho phép, đối với việc tính toán thông tin giá trị năng lượng ở phía bộ giải mã, chỉ một mặt giá trị thông tin năng lượng và số lượng các vạch phổ trong dải khôi phục hoặc trong tập hợp các dải khôi phục được nhóm phải được biết đến.

Trong các sơ đồ BWE thuộc tình trạng kỹ thuật, việc khôi phục vùng phổ HF trên tần số được gọi là tần số giao nhau thường dựa trên việc vá phổ. Thông thường, vùng HF gồm nhiều bản vá liền kề và mỗi bản vá này có nguồn gốc từ các vùng băng thông (BP) của phổ LF bên dưới tần số giao nhau đã cho. Trong phép biểu diễn giàn lọc của tín hiệu, các hệ thống như vậy sao chép tập hợp các hệ số băng con liền kề ra khỏi phổ LF vào vùng đích. Đường biên của các tập hợp được chọn là hệ thống điển hình phụ thuộc và không phụ thuộc tín hiệu. Đối với lượng tín hiệu, việc lựa chọn vá lỗi ổn định này dẫn tới âm sắc và màu sắc khó chịu của tín hiệu được khôi phục.

Các cách tiếp cận khác chuyển tín hiệu LF tới HF thông qua sự biến điệu Dải Phụ đơn thích ứng tín hiệu (SSB). Các cách tiếp cận như vậy có độ phức tạp tính toán cao so với [1] vì chúng hoạt động với tốc độ lấy mẫu cao trên các mẫu miền thời gian. Ngoài ra, việc vá có thể trở nên không ổn định, đặc biệt là đối với các tín hiệu không âm (ví dụ đoạn vô thanh), và do đó, việc vá thích ứng tín hiệu thuộc tình trạng kỹ thuật có thể gây ra các khiếm khuyết cho tín hiệu.

Cách tiếp cận theo sáng chế được gọi là điền đầy khoảng trống thông minh (IGF) và, trong cấu hình được ưu tiên của nó, nó được áp dụng trong hệ thống BWE dựa trên việc biến đổi thời gian-tần số, như ví dụ biến đổi Cosin rời rạc được biến đổi (MDCT). Tuy nhiên, bài học của sáng chế thường được áp dụng, ví dụ tương tự trong hệ thống dựa trên giàn lọc gương cầu phương (QMF).

Lợi thế của cấu hình IGF dựa trên MDCT là tích hợp liền mạch vào các bộ mã hóa âm thanh dựa trên MDCT, ví dụ mã hóa âm thanh nâng cao MPEG (AAC). Việc phân bổ phép biến đổi tương tự cho mã hóa âm thanh dạng sóng và cho BWE làm giảm đáng kể độ phức tạp tính toán cho mã hóa-giải mã âm thanh.

Ngoài ra, sáng chế đề xuất giải pháp cho các vấn đề ổn định vốn có trong các sơ đồ vá lỗi thích ứng trong tình trạng kỹ thuật.

Hệ thống được đề xuất dựa trên việc quan sát rằng đối với một số tín hiệu, việc lựa chọn bản vá lỗi ngẫu nhiên có thể dẫn tới thay đổi âm sắc và nhuộm màu tín hiệu. Nếu tín hiệu có âm trong vùng nguồn phổ (SSR) nhưng giống như nhiều âm trong vùng đích phổ (STR), việc vá lỗi nhiều âm STR bởi âm SSR có thể dẫn tới âm sắc không tự nhiên. Âm sắc của tín hiệu có thể cũng thay đổi từ vì cấu trúc âm của tín hiệu có thể bị lệch đồng chỉnh hoặc thậm chí bị phá hủy bởi quy trình vá lỗi.

Hệ thống IGF đã đưa ra thực hiện việc lựa chọn ô thông minh sử dụng tính tương quan chéo làm phép đo độ giống nhau giữa SSR cụ thể và STR cụ thể. Tính tương quan chéo của hai tín hiệu cung cấp phép đo độ tương tự của các tín hiệu này và độ trễ của tính tương quan tối đa và ký hiệu của nó. Do đó, cách tiếp cận của tính tương quan dựa trên việc lựa chọn ô có thể cũng được sử dụng để điều chỉnh một cách chính xác sự dịch chuyển phổ của phổ được sao chép để trở nên gần nhất với cấu trúc phổ ban đầu.

Đóng góp cơ bản của hệ thống được đề xuất là lựa chọn phép đo tương tự phù hợp, cũng như các kỹ thuật để ổn định quy trình lựa chọn ô. Kỹ thuật đã đưa ra cung cấp sự cân bằng tối ưu giữa việc thích ứng tín hiệu tức thời và đồng thời có tính ổn định theo thời gian. Việc cung cấp tính ổn định theo thời gian là đặc biệt quan trọng đối với các tín hiệu có ít sự tương đồng của SSR và STR và do đó thể hiện các giá trị tương quan chéo thấp hoặc nếu phép đo tương tự được sử dụng không rõ ràng. Trong

những trường hợp như vậy, tính ổn định ngăn chặn trạng thái ngẫu nhiên giả của việc lựa chọn ô thích ứng.

Ví dụ, lớp tín hiệu mà thường đặt ra các vấn đề của BWE thuộc tình trạng kỹ thuật được đặc trưng bởi đặc tính tập trung riêng của năng lượng cho các vùng phổ bất kỳ như được thể hiện trên Fig.12a (trái). Mặc dù đây là các phương pháp có sẵn để điều chỉnh đường bao phổ và tính âm điệu của phổ đã khôi phục trong vùng đích, đối với một số tín hiệu các phương pháp này không có khả năng bảo toàn âm sắc tốt như phương pháp được thể hiện trên Fig.12a (phải). Trong ví dụ được thể hiện trên Fig.12a, độ lớn của phổ trong vùng đích của tín hiệu gốc như trên được gọi là tần số giao nhau $f_{crossover}$ (Fig.12a trái) giảm gần như tuyến tính. Ngược lại, trong phổ đã khôi phục (Fig.12a phải) tập hợp riêng biệt của độ nghiêng và các đỉnh có mặt được coi là giả âm màu hóa âm sắc.

Bước quan trọng của cách tiếp cận mới là xác định tập hợp các ô mà trong đó lựa chọn dựa trên sự tương tự tiếp theo có thể diễn ra. Trước tiên, đường biên ô của cả vùng nguồn và vùng đích phải được xác định phù hợp với nhau. Do đó, vùng đích giữa tần số bắt đầu IGF của bộ mã hóa lõi $f_{IGFstart}$ và tần số khả dụng cao nhất $f_{IGFstop}$ được chia thành số nguyên tùy ý $nTar$ của các ô, mỗi ô này có kích thước được xác định riêng biệt. Sau đó, đối với mỗi ô đích $tar[idx_tar]$, tập hợp các ô nguồn có kích thước bằng nhau $src[idx_src]$ được tạo ra. Nhờ đó, mức độ cơ bản của tính tự do của hệ thống IGF được xác định. Tổng số ô nguồn $nSrc$ được xác định bởi bảng thông của vùng nguồn.

$$bw_{src} = (f_{IGFstart} - f_{IGFmin})$$

trong đó f_{IGFmin} là tần số khả dụng thấp nhất cho việc lựa chọn ô chẳng hạn như số nguyên $nSrc$ của việc làm phù hợp các ô nguồn thành bw_{src} . Số lượng tối thiểu của các ô nguồn là 0.

Để làm tăng hơn nữa mức độ tự do cho việc lựa chọn và điều chỉnh, các ô nguồn có thể được xác định để chồng lấp lên nhau bởi hệ số chồng lấp nằm trong khoảng giữa 0 và 1, trong đó 0 có nghĩa là không chồng lấp và 1 có nghĩa là chồng lấp 100%. Trường hợp chồng lấp 100% biểu thị rằng chỉ có một hoặc không có ô nguồn nào khả dụng.

Fig.12b thể hiện ví dụ của các đường biên ô của tập hợp các ô. Trong trường hợp này, tất cả ô đích được tương quan với từng ô nguồn. Trong ví dụ này, các ô nguồn chồng lấp 50%.

Đối với ô đích, sự tương quan chéo được tính toán với các ô nguồn khác nhau ở độ trễ lên đến các ngăn *xcorr_maxLag*. Đối với ô đích *idx_tar* và ô nguồn *idx_src*, *xcorr_val[idx_tar][idx_src]* đã cho đưa ra giá trị tối đa của sự tương quan chéo tuyệt đối giữa các ô, trong khi *xcorr_lag[idx_tar][idx_src]* đưa ra độ trễ mà giá trị tối đa này xuất hiện và *xcorr_sign[idx_tar][idx_src]* đưa ra dấu hiệu của sự tương quan chéo tại *xcorr_lag[idx_tar][idx_src]*.

Tham số *xcorr_lag* được sử dụng để điều khiển tính chặt chẽ của sự phù hợp giữa ô nguồn và ô đích. Tham số này dẫn tới các giả âm được giảm bớt và giúp bảo toàn tốt hơn âm sắc và sắc thái của tín hiệu.

Một số tình huống có thể xảy ra đó là kích thước của ô đích cụ thể lớn hơn kích thước của các ô nguồn khả dụng. Trong trường hợp này, ô nguồn khả dụng được lặp lại thường xuyên khi cần để điền đầy ô đích cụ thể hoàn toàn. Việc này vẫn có thể thể hiện sự tương quan chéo giữa ô đích lớn và ô nguồn nhỏ hơn để có vị trí tốt nhất của ô nguồn trong ô đích về mặt độ trễ *xcorr_lag* và ký hiệu *corr_sign* tương quan chéo.

Sự tương quan chéo của ô phổ thô và tín hiệu gốc có thể không là phép đo tương tự phù hợp nhất được áp dụng cho phổ âm thanh có kết cấu focman mạnh. Việc làm trắng phổ loại bỏ thông tin đường bao thô và do đó nhấn mạnh kết cấu mịn của phổ, mà được quan tâm hàng đầu để đánh giá độ tương tự của ô. Việc làm trắng cũng hỗ trợ trong việc tạo hình đường bao một cách dễ dàng của STR tại bộ giải mã cho các vùng được xử lý bởi IGF. Do đó, một cách tùy ý, ô và tín hiệu nguồn được làm trắng trước khi tính toán sự tương quan chéo.

Trong cấu hình khác, chỉ ô được làm trắng sử dụng quy trình được xác định trước. Dấu hiệu "làm trắng" được truyền biểu thị đến bộ giải mã rằng quy trình làm trắng được xác định trước tương tự sẽ được áp dụng cho ô trong IGF.

Để làm trắng tín hiệu, trước hết ước lượng đường bao phổ được tính toán. Sau đó, phổ MDCT được chia ra bởi đường bao phổ. Việc ước lượng đường bao phổ có thể

được ước lượng trên phổ MDCT, năng lượng phổ MDCT, phổ công suất phức hợp dựa trên MDCT hoặc sự ước lượng phổ công suất. Tín hiệu mà trên đó đường bao được ước lượng từ bây giờ sẽ được gọi là tín hiệu cơ bản.

Các đường bao được tính toán trên phổ công suất phức hợp dựa trên MDCT hoặc sự ước lượng phổ công suất như tín hiệu cơ bản có ưu điểm là không có sự dao động theo thời gian trên các thành phần âm.

Nếu tín hiệu cơ bản là miền năng lượng, phổ MDCT phải được lấy căn bậc hai của đường bao để làm trắng tín hiệu một cách chính xác.

Có các phương pháp khác nhau để tính toán đường bao:

- biến đổi tín hiệu cơ bản với việc biến đổi cosin rời rạc (DCT), chỉ giữ lại các hệ số DCT thấp hơn (thiết lập tầng trên cùng thành 0) và sau đó tính toán DCT nghịch đảo.
- tính toán đường bao phổ của tập hợp của các hệ số dự báo tuyến tính (LPC) được tính toán trên khung âm thanh miền thời gian.
- lọc tín hiệu cơ bản với bộ lọc thông thấp

Tốt hơn là, việc tiếp cận cuối cùng được chọn. Đối với các ứng dụng mà yêu cầu độ phức tạp tính toán thấp, một số sự đơn giản hóa có thể được thực hiện để làm trắng phổ MDCT: Trước tiên đường bao được tính toán bởi giá trị trung bình dịch chuyển. Điều này chỉ cần hai chu trình bộ xử lý trên mỗi ngăn MDCT. Sau đó để tránh việc tính toán phép chia và căn bậc hai, đường bao phổ được tính gần đúng bằng 2^n trong đó n là logarit số nguyên của đường bao. Trong miền này hoạt động căn bậc hai đơn giản trở thành hoạt động dịch chuyển và hơn nữa việc chia bởi đường bao có thể được thể hiện bởi hoạt động dịch chuyển khác.

Sau khi tính toán sự tương quan của từng ô nguồn với từng ô đích, đối với tất cả n_{Tar} ô đích, ô nguồn với sự tương quan cao nhất được chọn để thay thế chúng. Để phù hợp nhất với kết cấu phổ gốc, độ trễ của sự tương quan được sử dụng để điều biến phổ được thay thế bởi số lượng nguyên của các ngăn biến đổi. Trong trường hợp độ trễ lẻ, ô được điều biến bổ sung thông qua phép nhân bởi chuỗi theo thời gian thay thế của $-1/1$ để bù vào cho phép biểu diễn nghịch đảo tần số của mọi dải khác trong MDCT.

Fig.12c thể hiện ví dụ về sự tương quan giữa ô nguồn và ô đích. Trong ví dụ này, độ trễ của sự tương quan là 5, vì ô nguồn phải được điều biến bởi 5 ngăn về phía các ngăn tần số cao hơn trong giai đoạn sao chép tăng của thuật toán BWE. Ngoài ra, ký hiệu của ô được lật như giá trị tương quan tối đa là âm và việc điều biến bổ sung như mô tả ở trên tính cho độ trễ lẻ

Vì tổng lượng thông tin phụ để truyền từ bộ mã hóa tới bộ giải mã sẽ bao gồm dữ liệu dưới đây:

- $\text{tileNum}[nTar]$: chỉ số ô nguồn được lựa chọn trên ô đích
- $\text{tileSign}[nTar]$: ký hiệu của ô đích
- $\text{tileMod}[nTar]$: độ trễ của sự tương quan trên ô đích

Việc lược bỏ và làm ổn định ô là bước quan trọng trong IGF. Sự cần thiết và lợi ích của nó được giải thích với ví dụ, giả định tín hiệu âm thanh có âm ổn định như, ví dụ, nốt ống sáo để điều chỉnh âm điệu ổn định. Các mệnh lệnh có tính logic là các giả âm nhỏ nhất được đưa vào nếu, với vùng đích đã cho, các ô nguồn luôn được lựa chọn từ các khung chéo vùng nguồn tương tự. Thậm chí tín hiệu được giả định là ổn định, điều kiện này không được giữ tốt trong mọi khung vì phép đo tính giống nhau (ví dụ tính tương quan) của vùng nguồn giống nhau khác có thể trội hơn kết quả giống nhau (ví dụ sự tương quan chéo). Điều này dẫn tới $\text{tileNum}[nTar]$ giữa các khung liên kế để dao động giữa hai hoặc ba phép lựa chọn rất giống nhau. Nó có thể là nguồn gốc của nhiều âm có tính nhạc khó chịu giống giả âm.

Để loại bỏ loại giả âm này, tập hợp các ô nguồn sẽ được lược bớt sao cho các chi tiết còn lại của tập hợp nguồn không giống nhau một cách tối đa. Điều này đạt được trên tập hợp các ô nguồn

$$S = \{s_1, s_2, \dots, s_n\}$$

như sau. Với ô nguồn s_i bất kỳ, chúng ta tương quan nó với tất cả các ô nguồn khác, tìm ra sự tương quan nhất giữa s_i và s_j và lưu trữ nó trong ma trận S_x . Ở đây $S_x[i][j]$ bao gồm giá trị tương quan chéo tuyệt đối tối đa giữa s_i và s_j . Bổ sung ma trận S_x dọc theo các cột, cho chúng ta tổng sự tương quan chéo của ô nguồn s_i với tất cả các ô nguồn T khác.

$$T[i] = S_x[i][1] + S_x[i][2] \dots + S_x[i][n]$$

Ở đây T biểu diễn phép đo nguồn giống với các ô nguồn khác như thế nào. Nếu đối với ô nguồn i bất kỳ,

$$T > \text{ngưỡng}$$

ô nguồn i có thể bị loại bỏ khỏi tập hợp các nguồn tiềm năng vì nó tương quan cao với nguồn khác. Ô với độ tương quan thấp nhất từ tập hợp của các ô mà thỏa mãn điều kiện trong phương trình 1 được chọn làm ô biểu diễn cho tập con này. Với cách này, đảm bảo rằng các ô nguồn không được giống nhau một cách tối đa.

Phương pháp lược bớt ô cũng bao gồm bộ nhớ của tập hợp ô được lược bớt được sử dụng trong khung đứng trước. Các ô mà có hiệu quả trong khung trước cũng được giữ lại trong khung kế tiếp nếu các yếu tố thay thế để lược bớt tồn tại.

Cho các ô s_3 , s_4 và s_5 hoạt động ngoài các ô $\{s_1, s_2, \dots, s_5\}$ trong khung k , sau đó trong khung $k+1$ thậm chí nếu các ô s_1 , s_3 và s_2 cạnh tranh để được lược bớt với s_3 được tương quan tối đa với các ô khác, s_3 được giữ lại vì nó là ô nguồn hữu ích trong khung trước, và do đó việc giữ lại ô này trong tập hợp của ô nguồn là có ích để bắt buộc tính liên tục theo thời gian trong việc lựa chọn ô. Phương pháp này tốt hơn là được áp dụng nếu tính tương quan chéo giữa nguồn i và mục tiêu j , được biểu diễn là $T_x[i][j]$ là cao.

Phương pháp bổ sung cho việc làm ổn định ô là để giữ lại thứ tự ô từ khung trước $k-1$ nếu không có ô nguồn nào trong khung hiện thời k tương quan tốt với ô đích. Điều này có thể xảy ra nếu tính tương quan chéo giữa nguồn i và đích j , được biểu diễn dưới dạng $T_x[i][j]$ rất thấp cho mọi i, j

Ví dụ nếu

$$T_x[i][j] < 0,6$$

ngưỡng tạm thời được sử dụng bây giờ, sau đó

$$\text{tile Num}[n\text{Tar}]_k = \text{tileNum}[n\text{Tar}]_{k-1}$$

cho mọi $n\text{Tar}$ của khung k này.

Hai kỹ thuật trên làm giảm đáng kể các giả âm mà xuất hiện từ sự thay đổi nhanh chóng việc thiết lập số lượng ô khung chéo. Một ưu điểm khác của việc lược bớt ô và làm ổn định ô này là không cần gửi thêm thông tin đến bộ giải mã cũng như không cần thay đổi cấu trúc bộ giải mã. Việc lược bớt ô được đề xuất này là cách nhẹ nhàng để làm giảm nhiều âm có tính nhạc tiềm năng giống các giả âm hoặc nhiều âm quá mức trong vùng phổ được tạo ô.

Fig.11a minh họa bộ giải mã âm thanh cho việc giải mã tín hiệu âm thanh được mã hóa. Bộ giải mã âm thanh bao gồm bộ giải mã (lỗi) âm thanh 1102 để tạo ra phép biểu diễn được giải mã thứ nhất của tập hợp thứ nhất của phần phổ thứ nhất, phép biểu diễn được giải mã có độ phân giải phổ thứ nhất.

Ngoài ra, bộ giải mã âm thanh bao gồm bộ giải mã theo tham số 1104 để tạo ra phép biểu diễn được giải mã thứ hai của tập hợp thứ hai của phần phổ thứ hai có độ phân giải phổ thứ hai thấp hơn độ phân giải phổ thứ nhất. Ngoài ra, bộ tạo tần số 1106 được cung cấp mà nhận, như đầu vào thứ nhất 1101, phần phổ thứ nhất được giải mã và như đầu vào thứ hai tại 1103 thông tin tham số bao gồm, cho mỗi ô tần số mục tiêu hoặc dải khôi phục mục tiêu thông tin phạm vi nguồn. Bộ tạo tần số 1106 sau đó sử dụng sự tái tạo tần số bằng việc sử dụng giá trị phổ từ phạm vi nguồn được nhận dạng bởi thông tin phù hợp để tạo ra dữ liệu phổ cho khoảng đích. Sau đó, phần phổ thứ nhất 1101 và đầu ra của bộ tái tạo tần số 1107 đều là đầu vào vào bộ chuyển đổi phổ-thời gian 1108 để cuối cùng tạo ra tín hiệu âm thanh được giải mã.

Tốt hơn là, bộ giải mã âm thanh 1102 là bộ giải mã âm thanh miền phổ, mặc dù bộ giải mã âm thanh có thể cũng được thực hiện như bộ giải mã âm thanh bất kỳ khác như miền thời gian hoặc bộ giải mã âm thanh theo tham số.

Như được biểu thị trên Fig.11b, bộ tái tạo tần số 1106 có thể bao gồm các chức năng của khối 1120 minh họa bộ điều biến lựa chọn ô phạm vi nguồn cho độ trễ l, bộ lọc đã làm trắng 1122, khi mà dấu hiệu làm trắng 1123 được cung cấp, và ngoài ra, đường bao phổ với các chức năng điều chỉnh đã thực hiện được minh họa trong khối 1128 sử dụng dữ liệu phổ thô được tạo ra bởi hoặc là khối 1120 hoặc 1122 hoặc phối hợp cả hai khối. Dù thể nào, bộ tái tạo tần số 1106 có thể bao gồm bộ phận chuyển mạch 1124 kích hoạt lại dấu hiệu làm trắng đã nhận 1123. Khi dấu hiệu làm trắng

được thiết lập, đầu ra của bộ lựa chọn phạm vi nguồn/bộ điều biến ô cho các độ trễ lẻ là đầu vào vào trong bộ lọc làm trắng 1122. Sau đó, tuy nhiên dấu hiệu làm trắng 1123 không được thiết lập cho dải khôi phục nhất định, sau đó đường dẫn thông 1126 được kích hoạt vì đầu ra của khối 1120 được cung cấp cho khối điều chỉnh đường bao phổ 1128 mà không có bất kỳ việc làm trắng nào.

Có thể có nhiều hơn một mức làm trắng (1123) được ký hiệu trong dòng bit và các mức này có thể được ký hiệu trên mỗi ô. Trong trường hợp có ba mức được tạo tín hiệu trên ô, chúng sẽ được mã hóa bằng cách sau đây:

```

bit = readBit(1);
if(bit == 1) {
    for(tile_index = 0..nT)
        /*same levels as last frame*/
        whitening_level[tile_index] =
whitening_level_prev_frame[tile_index];
} else {
    /*first tile:*/
    tile_index = 0;
    bit = readBit(1);
    if(bit == 1) {
        whitening_level[tile_index] = MID_WHITENING;
    } else {
        bit = readBit(1);
        if(bit == 1) {
            whitening_level[tile_index] = STRONG_WHITENING;
        } else {
            whitening_level[tile_index] = OFF; /*no-whitening*/
        }
    }
}
/*remaining tiles:*/
bit = readBit(1);
if(bit == 1) {
    /*flattening levels for remaining tiles same as first.*/
    /*No further bits have to be read*/
    for(tile_index = 1..nT)
        whitening_level[tile_index] = whitening_level[0];
}

```

```

    } else {
        /*read bits for remaining tiles as for first tile*/
        for(tile_index = 1..nT) {
            bit = readBit(1);
            if(bit == 1) {
                whitening_level[tile_index] = MID_WHITENING;
            } else {
                bit = readBit(1);
                if(bit == 1) {
                    whitening_level[tile_index] = STRONG_WHITENING;
                } else {
                    whitening_level[tile_index] = OFF; /*no-whitening*/
                }
            }
        }
    }
}

```

MID_WHITENING and STRONG_WHITENING đề cập đến bộ lọc làm trắng khác nhau (1122) mà có thể khác nhau theo cách mà đường bao được tính toán (như mô tả trước đó).

Bộ tái tạo tần số phía bộ giải mã có thể được điều khiển bởi phạm vi nguồn ID 1121 khi chỉ sơ đồ lựa chọn ô phổ thô được sử dụng. Tuy nhiên, khi sơ đồ lựa chọn ô phổ đã điều chỉnh chính xác được sử dụng, thì thêm vào đó, độ trễ phạm vi nguồn 1119 được cung cấp. Ngoài ra, được đề xuất rằng việc tính toán sự tương quan cung cấp kết quả âm, thì thêm vào đó, ký hiệu của sự tương quan có thể cũng được sử dụng cho khối 1120 sao cho đường phổ dữ liệu trang được nhân lên cho từng đường với “-1” để tính toán cho ký hiệu âm.

Do đó, sáng chế này như được thảo luận trên Fig.11a, 11b đảm bảo rằng chất lượng âm thanh tốt nhất thu được do thực tế là phạm vi nguồn phù hợp nhất cho phạm vi dự định hoặc khoảng đích nhất định được tính toán ở phía bộ mã hóa và được sử dụng ở phía bộ giải mã.

Fig.11c là bộ mã hóa âm thanh nhất định để mã hóa tín hiệu âm thanh bao gồm bộ chuyển đổi thời gian-phổ 1130, bộ phân tích phổ được nối sau đó 1132 và ngoài ra,

bộ tính toán tham số 1134 và bộ mã hóa lỗi 1136. Bộ mã hóa lỗi 1136 xuất ra phạm vi nguồn được mã hóa và bộ tính toán tham số 1134 cung cấp thông tin phù hợp cho phạm vi đích.

Phạm vi nguồn được mã hóa được truyền tới bộ giải mã cùng với thông tin phù hợp cho các phạm vi đích sao cho bộ giải mã được minh họa trên Fig.11a là ở vị trí để thực hiện việc tái tạo tần số.

Bộ tính toán tham số 1134 được tạo cấu hình để tính toán sự tương đồng giữa phần phổ thứ nhất và phần phổ thứ hai và để xác định, dựa trên sự tương đồng đã tính toán, để phần phổ thứ hai phù hợp phần phổ thứ nhất phù hợp với phần phổ thứ hai. Tốt hơn là, kết quả phù hợp cho phạm vi nguồn khác và phạm vi đích khác nhau như minh họa trên Fig.12a, Fig.12b để xác định cặp phù hợp đã chọn gồm phần phổ thứ hai, và bộ tính toán tham số được tạo cấu hình để cung cấp thông tin phù hợp này nhận dạng cặp phù hợp thành tín hiệu âm thanh được mã hóa. Tốt hơn là, bộ tính toán tham số 1134 được tạo cấu hình để sử dụng các vùng đích được xác định trước trong tập hợp thứ hai của phần phổ thứ hai hoặc phạm vi nguồn được xác định trước trong tập hợp thứ nhất của phần phổ thứ nhất như minh họa, ví dụ trên Fig.12b. Tốt hơn là, các vùng đích đã xác định trước là không chồng lấp hoặc phạm vi nguồn được xác định trước là chồng lấp. Khi phạm vi nguồn được xác định trước là tập con của tập hợp thứ nhất của phần phổ thứ nhất dưới tần số bắt đầu điền đầy khoảng trống 309 của Fig.3a, và tốt hơn là, vùng đích được xác định trước bao trùm vùng phổ thấp hơn chồng lấp với tần số thấp hơn với tần số bắt đầu điền đầy khoảng trống để bất kỳ phạm vi mục tiêu nào được đặt phía trên tần số bắt đầu điền đầy khoảng trống và phạm vi nguồn được đặt dưới tần số bắt đầu điền đầy khoảng trống.

Như đã thảo luận, độ chi tiết mịn thu được bằng cách so sánh vùng đích và vùng nguồn mà không có bất kỳ độ trễ nào với vùng nguồn và vùng nguồn tương tự, nhưng với độ trễ nhất định. Độ trễ này được sử dụng trong bộ tính toán tương quan chéo 1140 của Fig.11d và việc lựa chọn cặp phù hợp được thực hiện cuối cùng bởi bộ lựa chọn 1144.

Ngoài ra, ưu tiên thực hiện việc làm làm trắng các phạm vi nguồn và/hoặc đích được minh họa ở khối 1142. Khối 1142 này sau đó cung cấp dấu hiệu làm trắng tới

đường truyền dữ liệu mà được sử dụng để điều khiển việc chuyển mạch phía bộ giải mã 1123 của Fig.11b. Ngoài ra, nếu bộ tính toán tương quan chéo 1140 cung cấp kết quả âm, sau đó kết quả âm này được tạo tín hiệu tới bộ giải mã. Do vậy, trong phương án ưu tiên, bộ lựa chọn ô đưa ra phạm vi nguồn ID cho phạm vi đích, độ trễ, dấu hiệu và khối 1142 cung cấp bổ sung dấu hiệu làm trắng.

Ngoài ra, bộ tính toán tham số 1134 được tạo cấu hình để thực hiện việc chia nhỏ ô nguồn 1146 bằng việc làm giảm số lượng các phạm vi nguồn tiềm năng mà trong đó vá lỗi nguồn bị hạ từ tập hợp của ô nguồn tiềm năng dựa trên ngưỡng giống nhau. Do vậy, khi hai ô nguồn là giống hơn hoặc bằng với ngưỡng tương tự, thì một trong số hai ô nguồn bị loại bỏ khỏi tập hợp của nguồn tiềm năng và ô nguồn đã loại bỏ không được sử dụng nhiều hơn cho việc xử lý khác và, cụ thể, không thể được lựa chọn bởi bộ lựa chọn ô 1144 hoặc không được sử dụng cho việc tính toán sự tương quan chéo giữa các phạm vi nguồn khác nhau và các phạm vi đích khác nhau được thể hiện trong khối 1140.

Các phép triển khai khác nhau được mô tả với các hình minh họa khác nhau. Các hình vẽ từ Fig.1a đến Fig.5c đề cập đến tốc độ đầy đủ hoặc sơ đồ bộ mã hóa/giải mã bằng thông đầy đủ. Các hình vẽ từ Fig.6a đến Fig.7e đề cập đến sơ đồ bộ mã hóa/giải mã với việc xử lý TNS hoặc TTS. Các hình vẽ từ Fig.8a đến Fig.8e đề cập đến sơ đồ bộ mã hóa/bộ giải mã với việc xử lý kênh đôi cụ thể. Các hình vẽ từ Fig.9a đến Fig.10d đề cập đến việc tính toán và áp dụng thông tin năng lượng cụ thể, và các hình vẽ từ Fig.11a đến Fig.12c đề cập đến cách thức cụ thể của việc lựa chọn ô.

Tất cả các khía cạnh khác nhau này có thể được sử dụng sáng tạo độc lập với nhau, nhưng, ngoài ra, cũng có thể được áp dụng cùng nhau như được minh họa một cách cơ bản trên Fig.2a và Fig.2b. Tuy nhiên, việc xử lý kênh đôi cụ thể có thể được áp dụng cho sơ đồ bộ mã hóa/bộ giải mã được minh họa trên Fig.13, và điều này cũng đúng với xử lý TNS/TTS, việc tính toán và áp dụng thông tin năng lượng đường bao trong dải khôi phục hoặc việc nhận dạng phạm vi nguồn thích ứng và áp dụng tính tương quan ở phía bộ giải mã. Mặt khác, phương diện tốc độ đầy đủ có thể được áp dụng có hoặc không có việc xử lý TNS/TTS, có hoặc không có việc xử lý kênh đôi, có hoặc không có sự nhận dạng phạm vi nguồn thích ứng hoặc với dạng khác của việc

tính toán năng lượng cho việc biểu diễn đường bao phổ. Do đó, đây là các đặc điểm rõ ràng của một trong các khía cạnh riêng này có thể được áp dụng tốt trong các khía cạnh khác.

Mặc dù một vài khía cạnh đã được mô tả trong phạm vi của thiết bị để mã hóa hoặc giải mã, rõ ràng rằng các khía cạnh này cũng thể hiện mô tả của phương pháp tương ứng, trong đó khối hoặc thiết bị tương ứng với bước xử lý của phương pháp hoặc đặc điểm của bước xử lý của phương pháp. Tương tự, các khía cạnh được mô tả trong phạm vi của bước xử lý của phương pháp cũng thể hiện mô tả của khối hoặc mục tương ứng hoặc đặc điểm của thiết bị tương ứng. Một số hoặc tất cả các bước xử lý của phương pháp có thể được thực thi bằng (hoặc sử dụng) thiết bị phần cứng, như ví dụ, bộ vi xử lý, máy tính có thể lập trình được hoặc mạch điện tử. Trong một số phương án, một số hoặc nhiều bước xử lý phương pháp quan trọng có thể có thực thi bằng các thiết bị như vậy.

Tùy 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 chuyển tiếp như vật ghi kỹ thuật số, ví dụ, đĩa mềm, DVD, CD, ROM, PROM, 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 lập trình được sao cho phương pháp tương ứng được thực hiện. Do đó, vật ghi lưu trữ kỹ thuật 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 nhờ điện tử, chúng có khả năng kết hợp với hệ thống máy tính lập trình được, để mà một trong các phương pháp được mô tả ở đây được thực hiện.

Thông thường, các phương án của sáng chế có thể được thực hiện như một sản phẩm chương trình máy tính với mã chương trình, mã chương trình hoạt động để thực hiện một trong các phương pháp khi sản phẩm chương trình máy tính chạy trên máy tính. Mã chương trình có thể ví dụ được lưu trữ trên vật mang có thể đọc được bằng máy.

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

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

Phương án khác của các 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 dạng hữu hình điển hình và/hoặc không chuyển tiếp.

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

Phương án khác bao gồm 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 lắp trên máy tính để thực hiện một trong các phương pháp được mô tả ở đây.

Phương án khác theo sáng chế bao gồm thiết bị hoặc hệ thống được tạo cấu hình để chuyển (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 nhận. Bộ nhận có thể, ví dụ, là máy tính, thiết bị di động, thiết bị bộ nhớ hoặc tương tự. Thiết bị hoặc hệ thống có thể, ví dụ, bao gồm máy chủ tập tin để chuyển chương trình máy tính đến bộ nhận.

Theo một số phương án, thiết bị logic lập trình được (ví dụ, mảng cổng 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 các phương pháp

pháp được mô tả ở đây. Thông thường, các phương pháp ưu tiên được thực hiện bởi thiết bị phần cứng bất kỳ.

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

YÊU CẦU BẢO HỘ

1. Thiết bị giải mã tín hiệu âm thanh được mã hóa bao gồm:

bộ giải mã âm thanh miền phổ (602) để tạo ra phép biểu diễn được giải mã thứ nhất của tập hợp thứ nhất của các phần phổ thứ nhất là các giá trị dự dự báo phổ;

bộ tái tạo tần số (604) để tạo ra phần phổ thứ hai được khôi phục sử dụng phần phổ thứ nhất của tập hợp thứ nhất của các phần phổ thứ nhất, trong đó phần phổ thứ hai được khôi phục và tập hợp thứ nhất của các phần phổ thứ nhất bao gồm các giá trị dự dự báo phổ; và

bộ lọc dự báo nghịch đảo (606, 616, 626) để thực hiện dự báo nghịch đảo trên tần số bằng cách sử dụng các giá trị dự dự báo phổ cho tập hợp thứ nhất của các phần phổ thứ nhất và phần phổ thứ hai được khôi phục sử dụng thông tin bộ lọc dự báo (607) có trong tín hiệu âm thanh được mã hóa.

2. Thiết bị theo điểm 1,

còn bao gồm bộ tạo hình đường bao phổ (614) để tạo hình đường bao phổ của tín hiệu đầu vào hoặc tín hiệu đầu ra của bộ lọc dự báo nghịch đảo (606).

3. Thiết bị theo điểm 2,

trong đó tín hiệu âm thanh được mã hóa bao gồm thông tin đường bao phổ cho phần phổ thứ hai, thông tin đường bao phổ có độ phân giải phổ thứ hai, độ phân giải phổ thứ hai thấp hơn độ phân giải phổ thứ nhất liên quan đến phép biểu diễn được giải mã thứ nhất,

trong đó bộ tạo hình đường bao phổ (624) được tạo cấu hình để áp dụng thao tác tạo hình đường bao phổ trên đầu ra của bộ lọc dự báo nghịch đảo (622), trong đó thông tin bộ lọc dự báo (607) được xác định bằng cách sử dụng tín hiệu âm thanh trước khi lọc dự báo, hoặc

trong đó bộ tạo hình đường bao phổ (614) được tạo cấu hình để áp dụng thao tác tạo hình đường bao phổ trên đầu vào của bộ lọc dự báo nghịch đảo (616), trong đó thông tin bộ lọc dự báo (607) được xác định bằng cách sử dụng tín hiệu âm thanh sau khi lọc dự báo trong bộ mã hóa.

4. Thiết bị theo một điểm bất kỳ trong số các điểm trên,
còn bao gồm bộ chuyển đổi tần số thời gian (212) để chuyển đổi đầu ra của bộ lọc dự báo nghịch đảo (210) hoặc đầu ra được tạo hình đường bao của bộ lọc dự báo nghịch đảo (210) thành phép biểu diễn thời gian.
5. Thiết bị theo một điểm bất kỳ trong số các điểm trên,
trong đó bộ lọc dự báo nghịch đảo (606) là bộ lọc phức hợp được xác định bởi thông tin bộ lọc dự báo (607).
6. Thiết bị theo một điểm bất kỳ trong số các điểm trên,
trong đó bộ giải mã âm thanh miền phổ (602) được tạo cấu hình để tạo ra phép biểu diễn được giải mã thứ nhất sao cho phép biểu diễn được giải mã thứ nhất có tần số Nyquist bằng với tốc độ lấy mẫu của tín hiệu miền thời gian được tạo ra bởi sự chuyển đổi thời gian-tần số của đầu ra của bộ lọc dự báo nghịch đảo (606).
7. Thiết bị theo một điểm bất kỳ trong số các điểm trên,
trong đó bộ giải mã âm thanh miền phổ (602) được tạo cấu hình sao cho tần số tối đa được biểu diễn bởi giá trị phổ cho tần số tối đa trong phép biểu diễn được giải mã thứ nhất bằng với tần số tối đa có trong phép biểu diễn thời gian được tạo ra bởi sự chuyển đổi tần số-thời gian đầu ra của bộ lọc dự báo (606), trong đó giá trị phổ cho tần số tối đa trong phép biểu diễn được giải mã thứ nhất là không hoặc khác không.
8. Thiết bị theo một điểm bất kỳ trong số các điểm trên,
trong đó phép biểu diễn được giải mã thứ nhất của tập hợp thứ nhất của các phần phổ thứ nhất bao gồm các giá trị phổ thực.
trong đó thiết bị còn bao gồm bộ ước lượng (724) để ước lượng giá trị ảo cho tập hợp thứ nhất của các phần phổ thứ nhất từ tập hợp thứ nhất giá trị thực của các phần phổ thứ nhất, và trong đó bộ lọc dự báo nghịch đảo (728) là bộ lọc dự báo nghịch đảo phức hợp được xác định bởi thông tin bộ lọc dự báo giá trị phức hợp (714), và trong đó thiết bị còn bao gồm bộ chuyển đổi tần số-thời gian (730) được tạo cấu hình để thực hiện việc chuyển đổi phổ có giá trị phức hợp thành tín hiệu âm thanh miền thời gian.

9. Thiết bị theo một điểm bất kỳ trong số các điểm trên,

trong đó bộ lọc dự báo nghịch đảo (606, 728) được tạo cấu hình để áp dụng nhiều bộ lọc con, trong đó biên tần số của mỗi bộ lọc con trùng với biên tần số của băng khôi phục trùng với ô tần số.

10. Thiết bị mã hóa tín hiệu âm thanh, bao gồm:

bộ chuyển đổi thời gian-phổ (100, 702) để chuyển đổi tín hiệu âm thanh (99) thành phép biểu diễn phổ;

bộ lọc dự báo (704) để thực hiện việc dự báo trên tần số trên phép biểu diễn phổ để tạo ra các giá trị dư phổ, bộ lọc dự báo được xác định bởi thông tin bộ lọc được suy ra từ tín hiệu âm thanh;

bộ mã hóa âm thanh (707) để mã hóa tập hợp thứ nhất của các phần phổ thứ nhất của các giá trị dư phổ để thu được tập hợp thứ nhất được mã hóa của các giá trị phổ thứ nhất có độ phân giải phổ thứ nhất;

bộ mã hóa theo tham số (706) để mã hóa theo tham số tập hợp thứ hai của các phần phổ thứ hai của các giá trị dư phổ hoặc của các giá trị của phép biểu diễn phổ với độ phân giải phổ thứ hai thấp hơn độ phân giải phổ thứ nhất; và

giao diện đầu ra (710) để xuất ra tín hiệu được mã hóa bao gồm tập hợp thứ hai được mã hóa, tập hợp thứ nhất được mã hóa và thông tin bộ lọc (714).

11. Thiết bị theo điểm 10,

trong đó bộ chuyển đổi thời gian-phổ được tạo cấu hình để thực hiện phép biến đổi cosin rời rạc cải biên, và trong đó các giá trị dư phổ là các giá trị dư phổ biến đổi cosin rời rạc cải biên.

12. Thiết bị theo một điểm trong số các điểm 10 và 11,

trong đó bộ lọc dự báo (704) bao gồm bộ tính thông tin bộ lọc, bộ tính thông tin bộ lọc được tạo cấu hình để sử dụng các giá trị phổ của phép biểu diễn phổ để tính toán thông tin bộ lọc và trong đó bộ lọc dự báo được tạo cấu hình để tính toán các giá trị dư phổ sử dụng các giá trị phổ của phép biểu diễn phổ, trong đó các giá trị phổ để

tính toán thông tin bộ lọc và đầu vào các giá trị phổ vào bộ lọc dự báo được nhận từ tín hiệu âm thanh tương tự.

13. Thiết bị theo một điểm bất kỳ trong số các điểm từ 10 đến 12

trong đó bộ lọc dự báo bao gồm bộ tính toán bộ lọc để tính toán thông tin bộ lọc sử dụng các giá trị phổ từ tần số bắt đầu TNS tới tần số dừng TNS, trong đó tần số bắt đầu TNS thấp hơn 4kHz và tần số dừng TNS cao hơn 9kHz.

14. Thiết bị theo một điểm bất kỳ trong số các điểm từ 10 đến 13 còn bao gồm bộ phân tích (102, 706) để xác định tập hợp thứ nhất của các phần phổ thứ nhất được mã hóa bởi bộ mã hóa âm thanh (708), bộ phân tích sử dụng tần số bắt đầu điền đầy khoảng trống, trong đó các phần phổ dưới tần số bắt đầu điền đầy khoảng trống là các phần phổ thứ nhất, và

trong đó tần số dừng TNS lớn hơn tần số bắt đầu điền đầy khoảng trống.

15. Thiết bị theo một điểm bất kỳ trong số các điểm từ 10 đến 14,

trong đó bộ chuyển đổi thời gian-tần số (702) được tạo cấu hình để cung cấp phép biểu diễn phổ phức hợp,

trong đó bộ lọc dự báo được tạo cấu hình để thực hiện việc dự báo trên tần số với phép biểu diễn phổ phức hợp, và

trong đó thông tin bộ lọc (714) được tạo cấu hình để xác định bộ lọc dự báo nghịch đảo phức hợp.

16. Phương pháp giải mã tín hiệu âm thanh được mã hóa, phương pháp bao gồm các bước:

tạo ra (602) phép biểu diễn được giải mã thứ nhất của tập hợp thứ nhất của các phần phổ thứ nhất là các giá trị dư dự báo phổ;

tái tạo (604) phần phổ thứ hai được khôi phục sử dụng phần phổ thứ nhất của tập hợp thứ nhất của các phần phổ thứ nhất, trong đó phần phổ thứ hai được khôi phục và tập hợp thứ nhất của các phần phổ thứ nhất bao gồm các giá trị dư dự báo phổ; và

thực hiện (606, 616, 626) phép dự báo nghịch đảo trên tần số sử dụng các giá trị dư dự báo phổ cho tập hợp thứ nhất của các phần phổ thứ nhất và phần phổ thứ hai

được khôi phục sử dụng thông tin bộ lọc dự báo (607) có trong tín hiệu âm thanh được mã hóa.

17. Phương pháp theo điểm 16,

trong đó tín hiệu âm thanh được mã hóa bao gồm thông tin đường bao phổ cho phần phổ thứ hai, thông tin đường bao phổ có độ phân giải phổ thứ hai, độ phân giải phổ thứ hai thấp hơn độ phân giải phổ thứ nhất liên quan đến phép biểu diễn được giải mã thứ nhất,

trong đó việc tái tạo (604) bao gồm việc tạo hình đường bao phổ (624) bao gồm việc áp dụng thao tác tạo hình đường bao phổ trên đầu ra của bước thực hiện (606, 616, 626) dự báo nghịch đảo trên tần số, trong đó thông tin bộ lọc dự báo (607) được xác định bằng cách sử dụng tín hiệu âm thanh trước khi lọc dự báo, hoặc

trong đó việc tái tạo (604) bao gồm việc tạo hình đường bao phổ (624) bao gồm việc áp dụng hoạt động tạo hình đường bao phổ trên đầu vào của bước thực hiện (606, 616, 626) dự báo nghịch đảo trên tần số, trong đó thông tin bộ lọc dự báo (607) được xác định bằng cách sử dụng tín hiệu âm thanh sau khi lọc dự báo trong bộ mã hóa.

18. Phương pháp mã hóa tín hiệu âm thanh, bao gồm:

chuyển đổi (100, 702) tín hiệu âm thanh (99) thành phép biểu diễn phổ;

thực hiện (704) dự báo trên tần số trên phép biểu diễn phổ để tạo ra các giá trị dư phổ, bộ lọc dự báo được xác định bởi thông tin bộ lọc nhận được từ tín hiệu âm thanh;

mã hóa (708) tập hợp thứ nhất của các phần phổ thứ nhất của các giá trị dư phổ để thu được tập hợp thứ nhất được mã hóa của các giá trị phổ thứ nhất có độ phân giải phổ thứ nhất;

mã hóa theo tham số (706) tập hợp thứ hai của các phần phổ thứ hai của các giá trị dư phổ hoặc của các giá trị của phép biểu diễn phổ với độ phân giải phổ thứ hai thấp hơn độ phân giải phổ thứ nhất; và

xuất ra (710) tín hiệu được mã hóa bao gồm tập hợp thứ hai được mã hóa, tập hợp thứ nhất được mã hóa và thông tin bộ lọc (714).

19. Vật ghi đọc được bằng máy tính chứa chương trình máy tính để thực hiện phương pháp theo điểm 16 khi chạy trên máy tính hoặc bộ xử lý.
20. Vật ghi đọc được bằng máy tính chứa chương trình máy tính để thực hiện phương pháp theo điểm 18 khi chạy trên máy tính hoặc bộ xử lý.

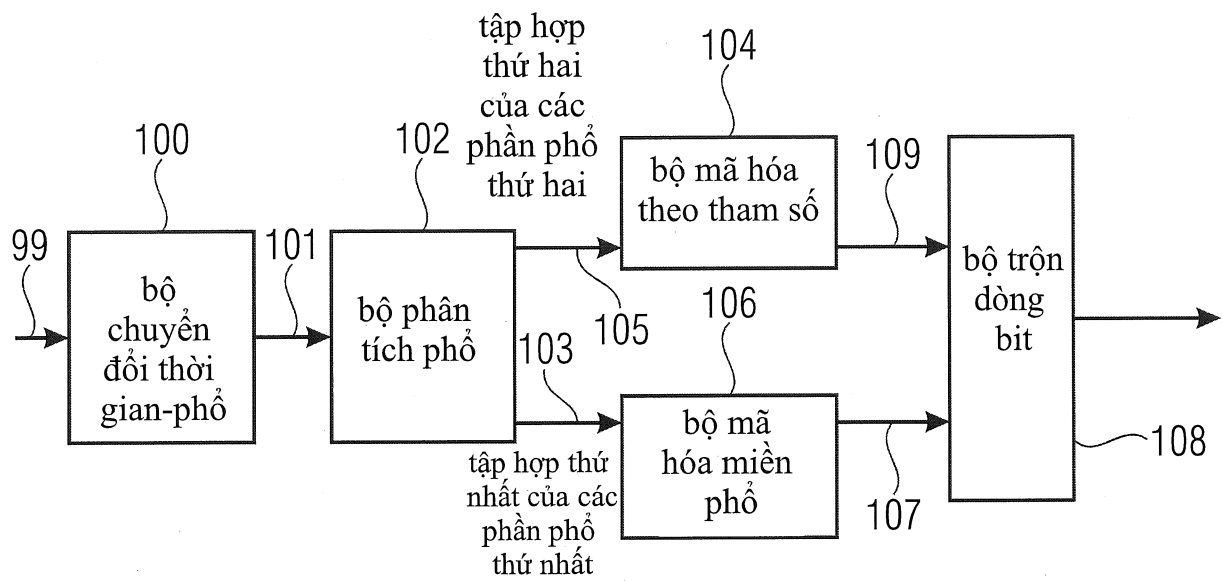


FIG 1A

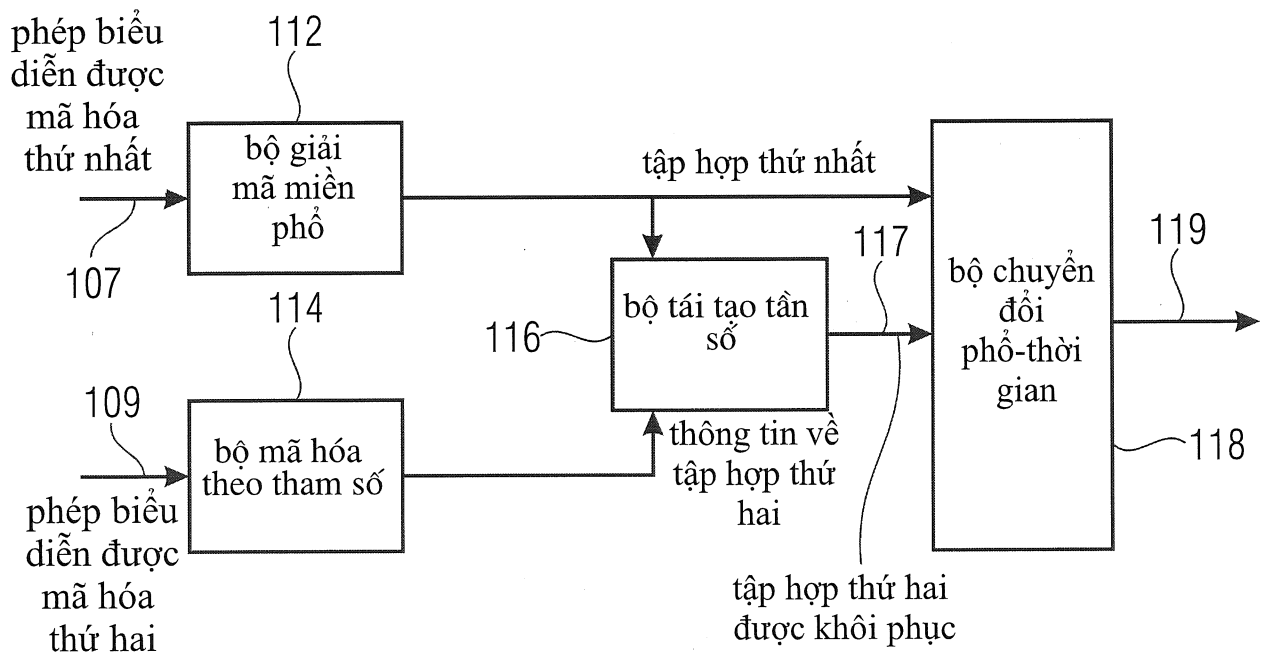


FIG 1B

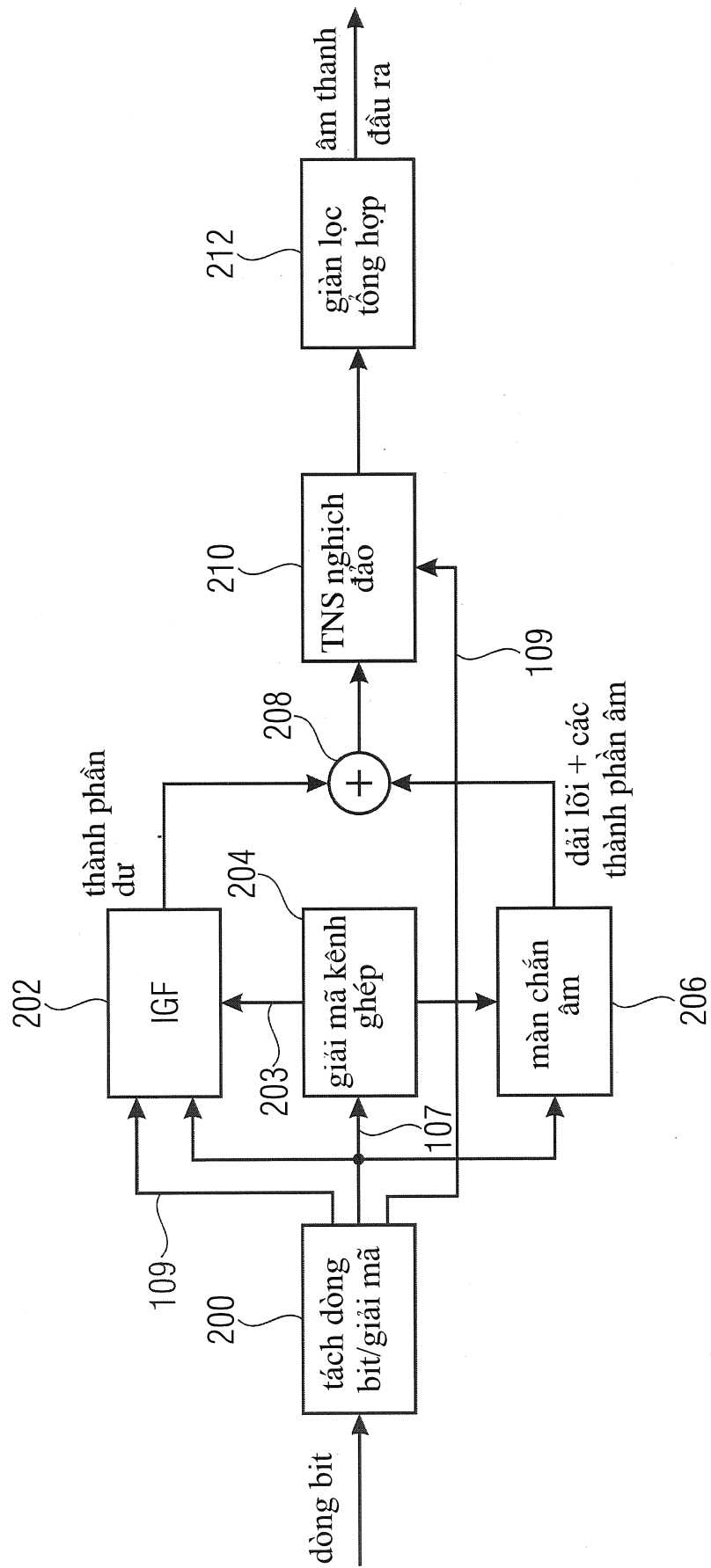


FIG 2A

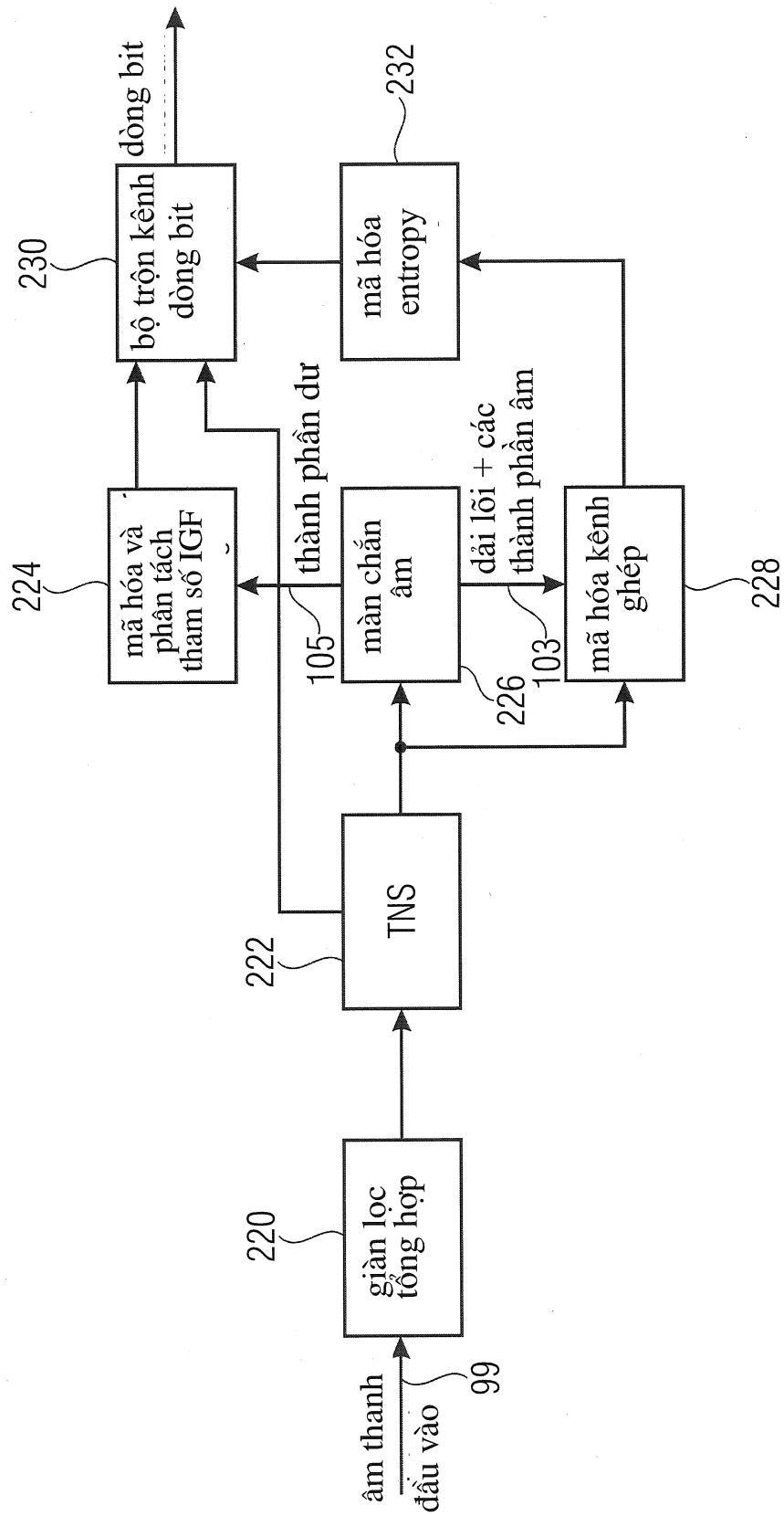


FIG 2B

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

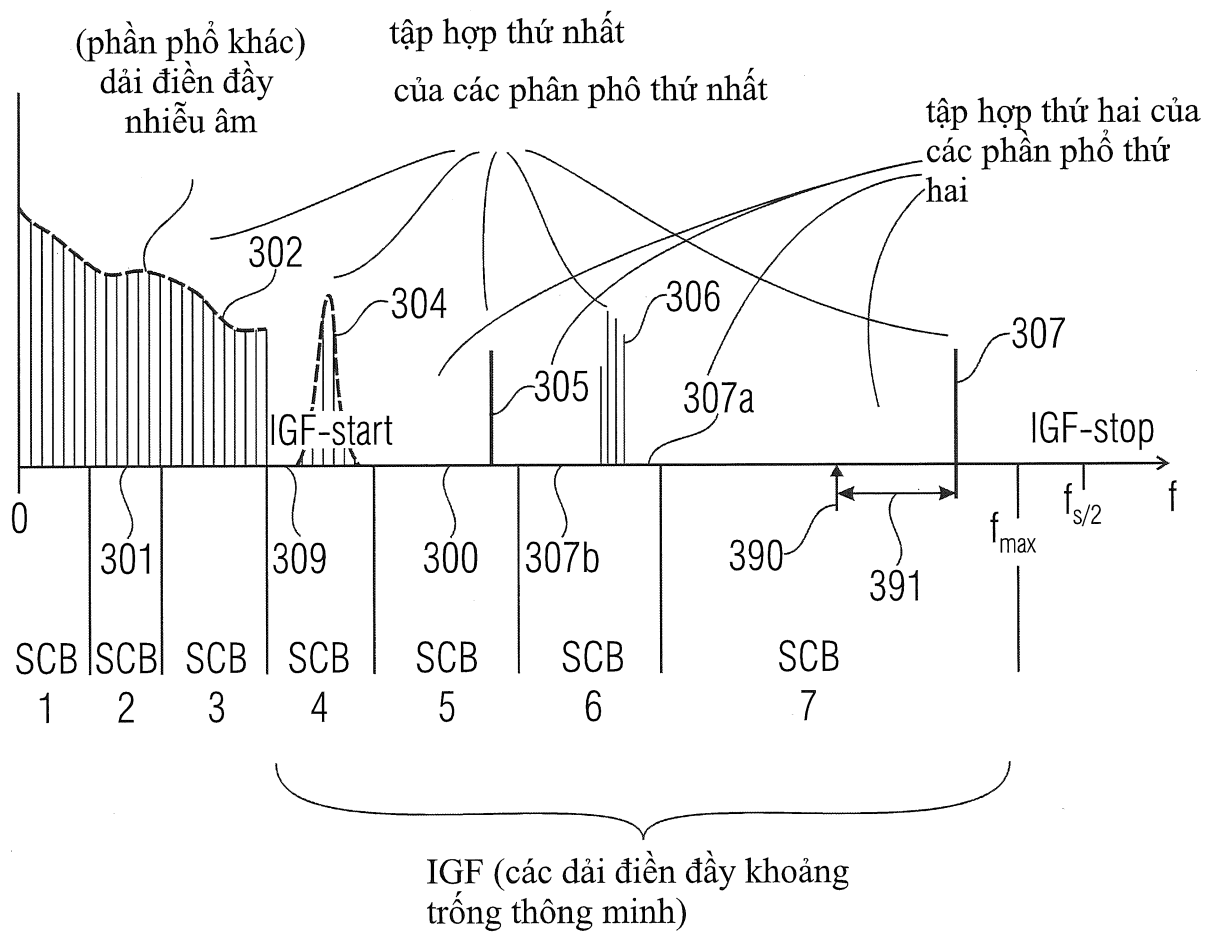


FIG 3A

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

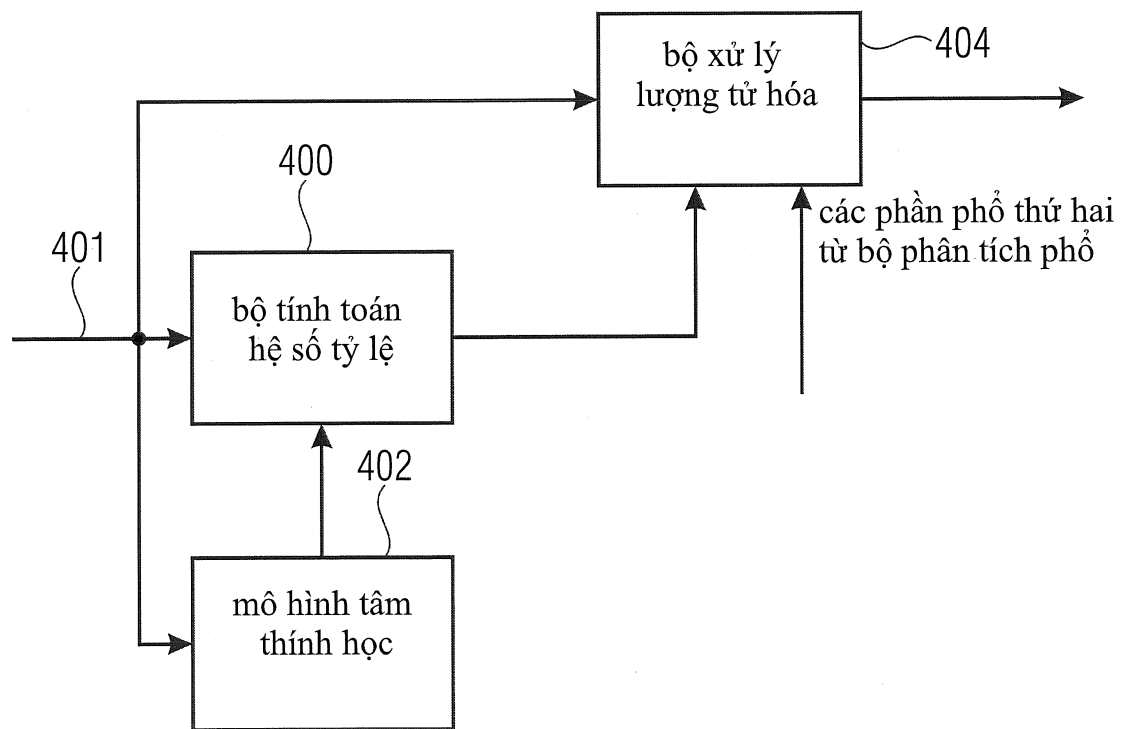


FIG 4A

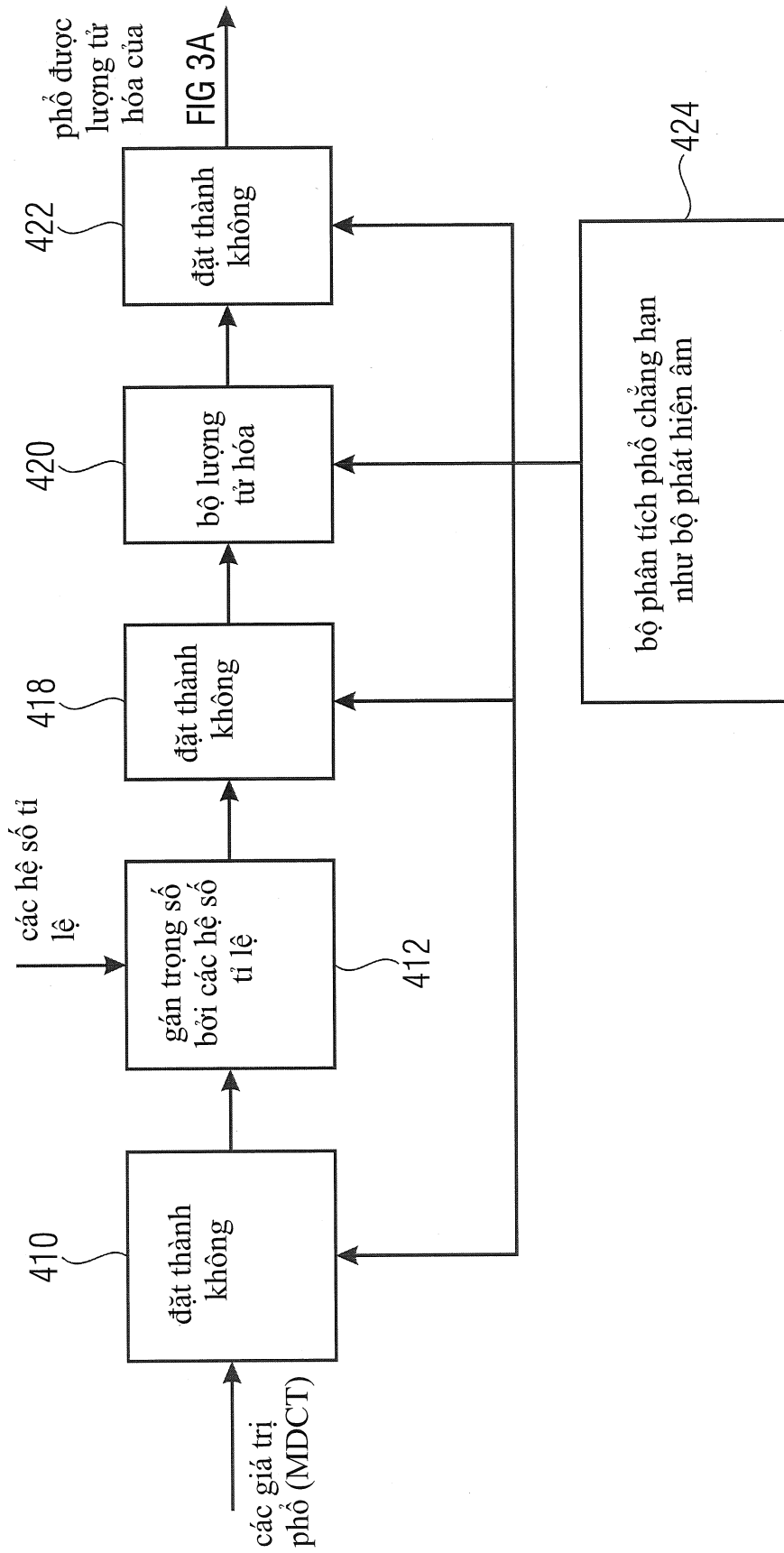


FIG 4B

(Bộ xử lý lượng tử hóa)

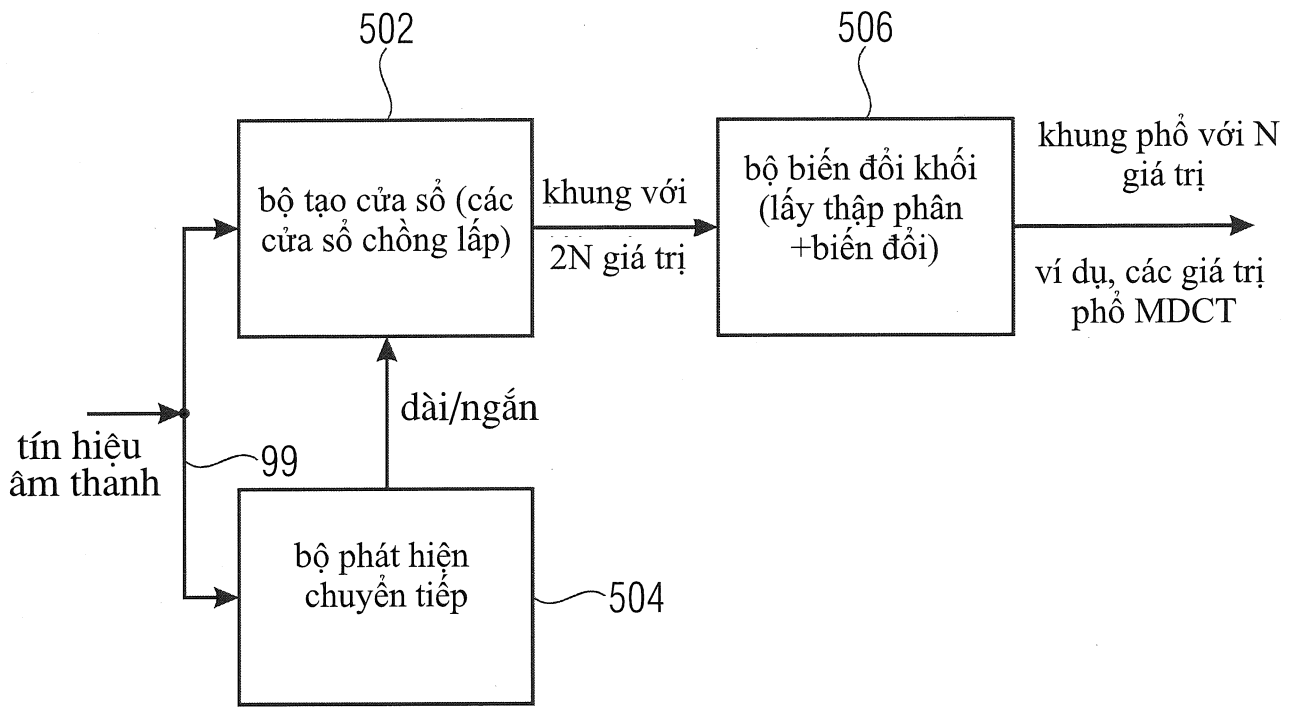


FIG 5A

(Các phần phổ khác)

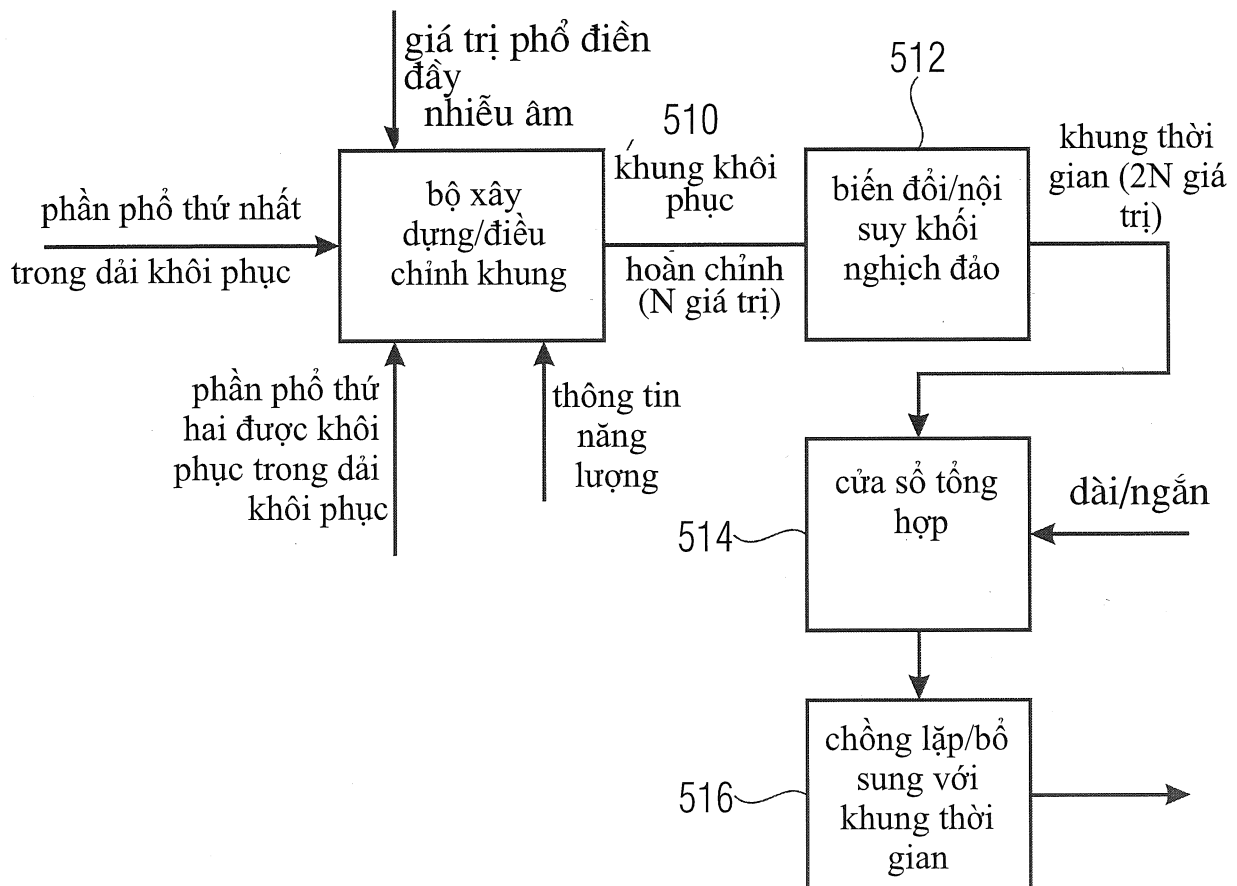


FIG 5B

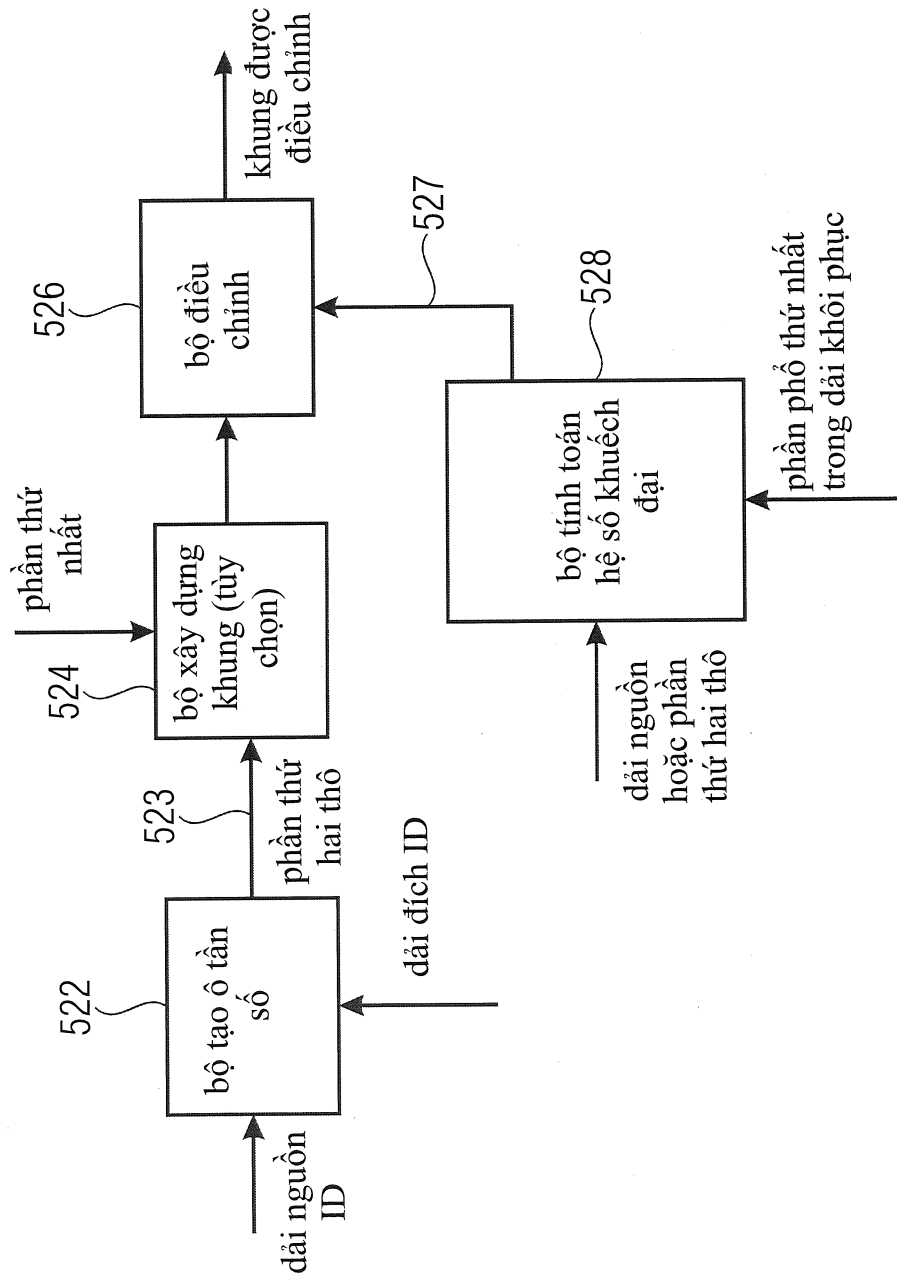


FIG 5C

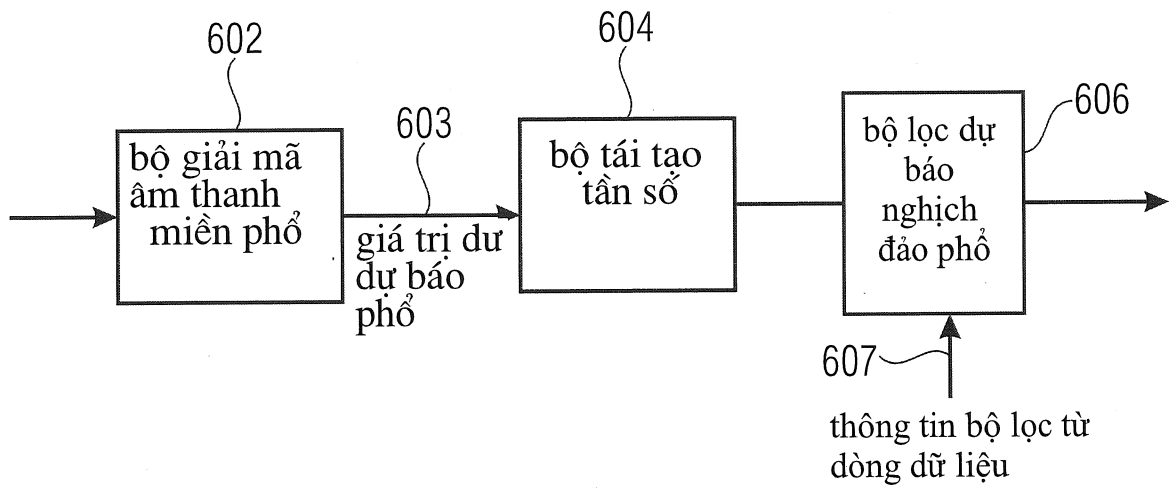


FIG 6A

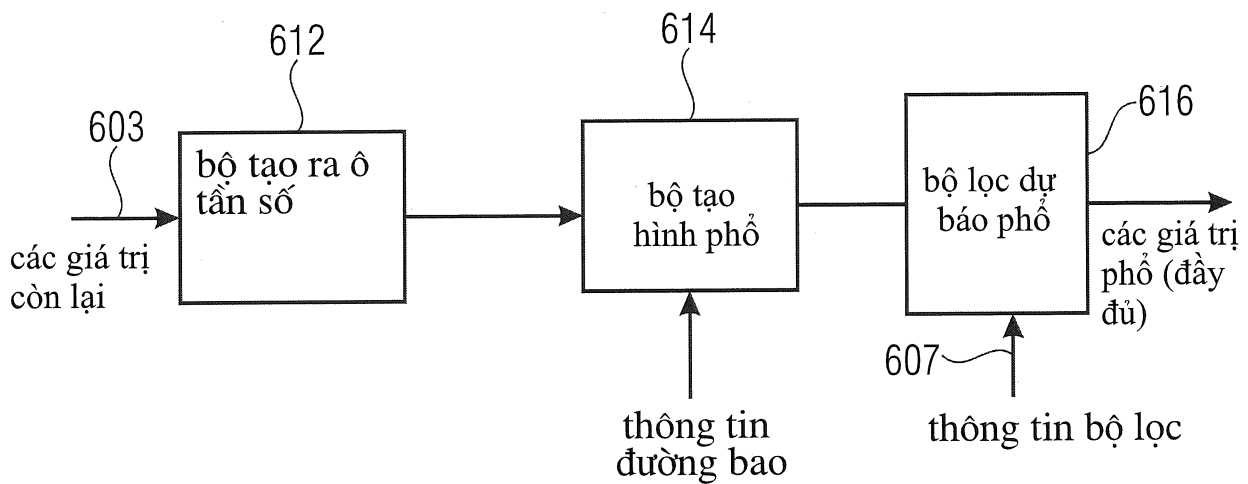


FIG 6B

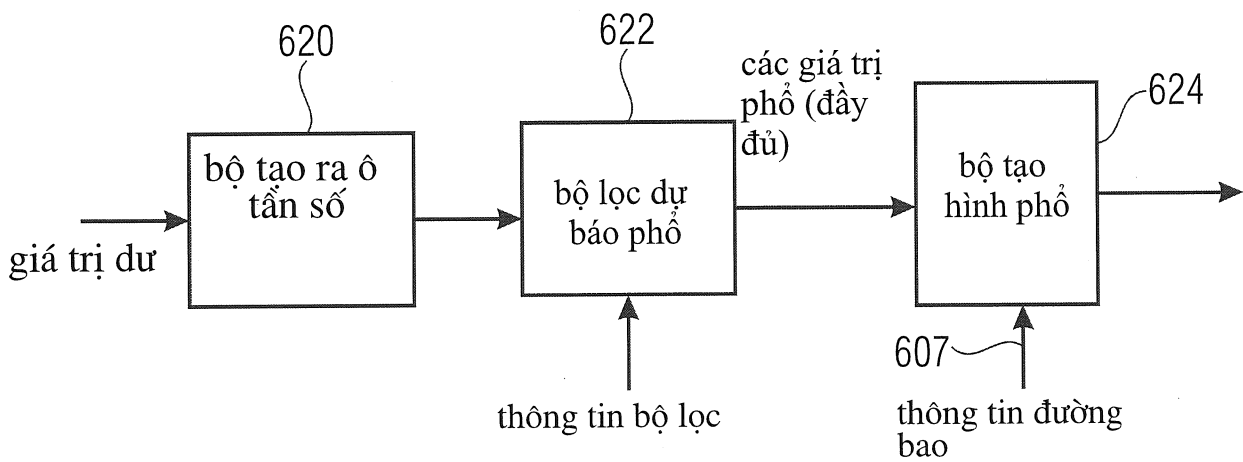


FIG 6C

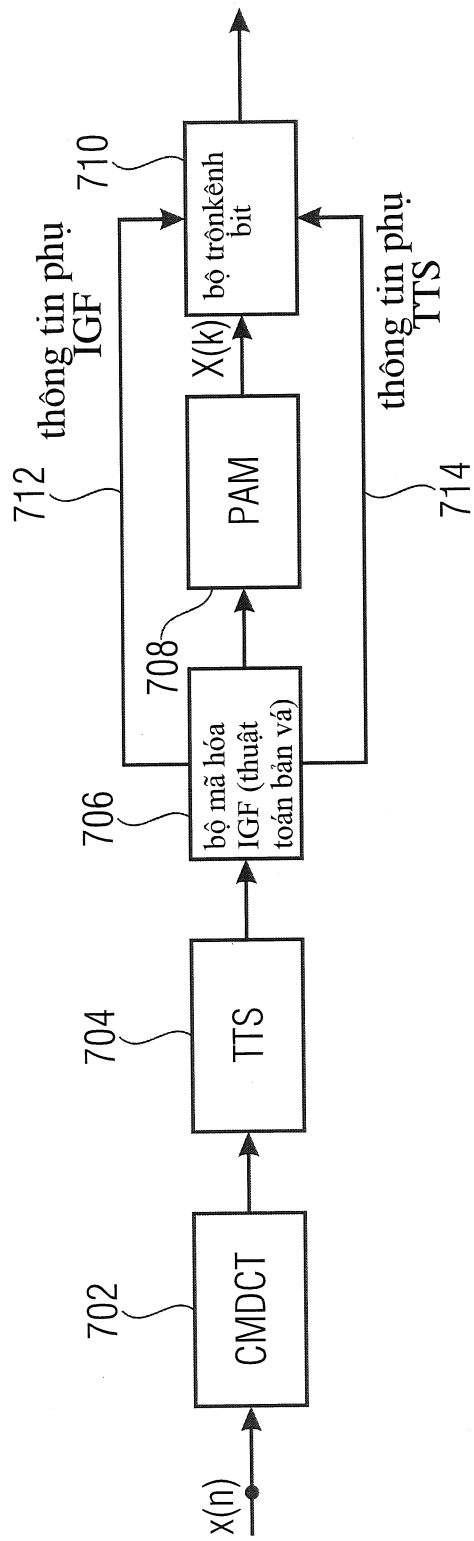


FIG 7A
(Bộ mã hóa)

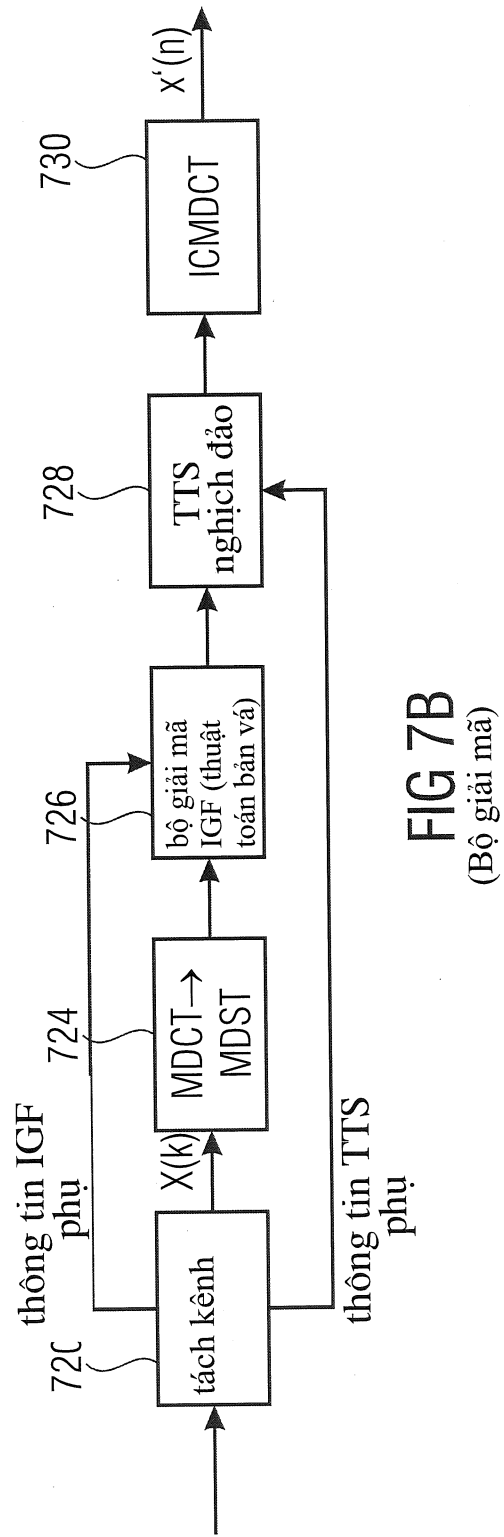
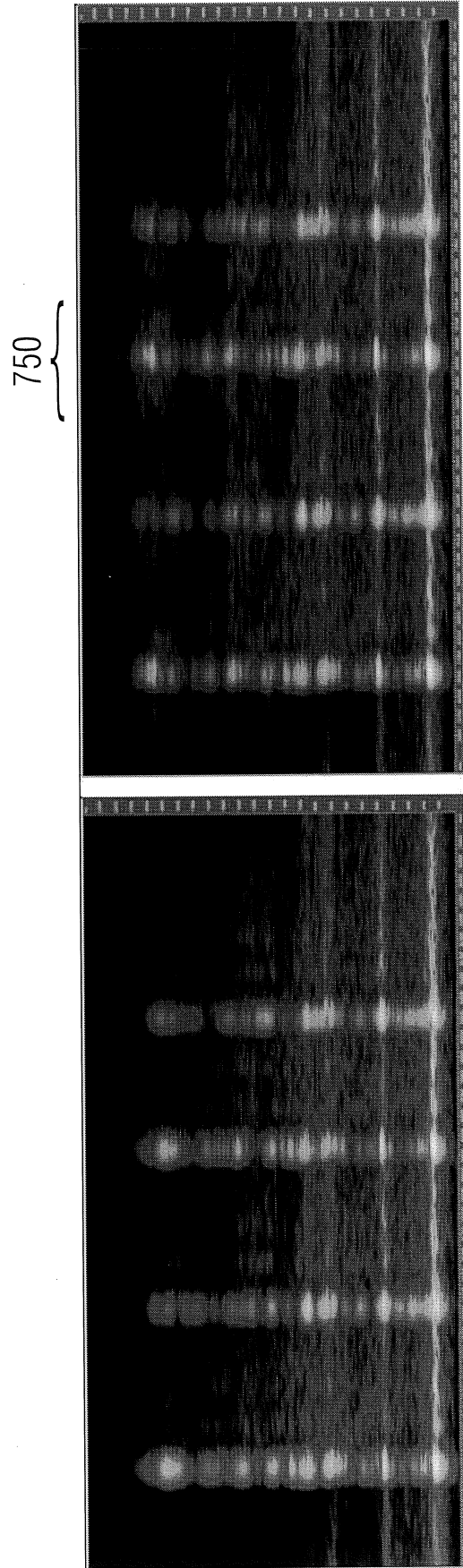
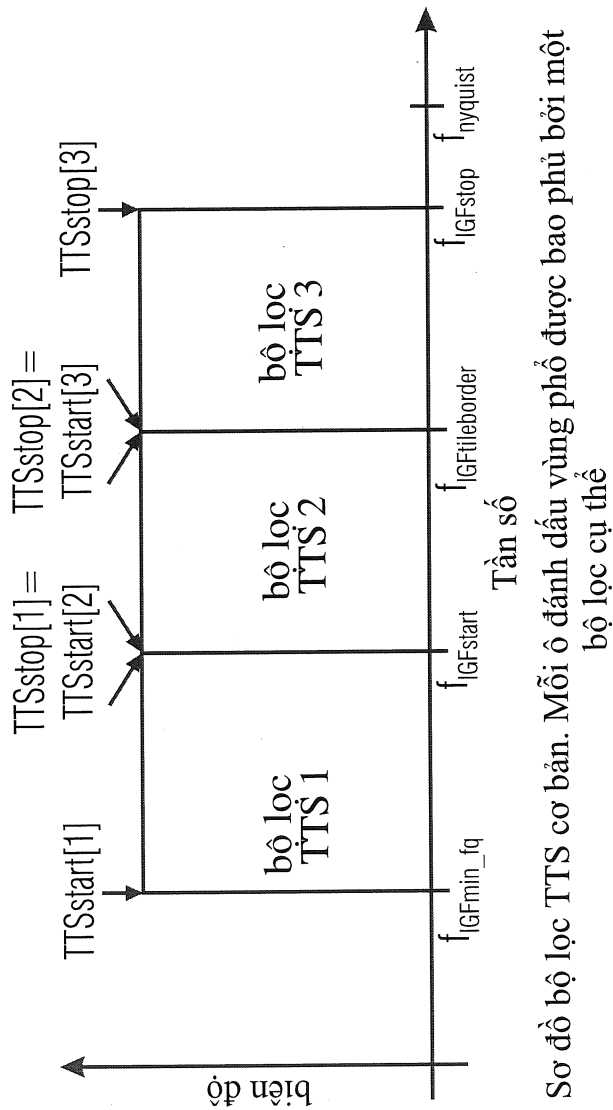


FIG 7B
(Bộ giải mã)



Ảnh phổ DFT của tín hiệu ban đầu (hình bên trái);
ảnh phổ DFT của tín hiệu mở rộng không có TTS (hình bên phải)

FIG 7C



Sơ đồ bộ lọc TTS cơ bản. Mỗi ô đánh dấu vùng phổ được bao phủ bởi một bộ lọc cụ thể

FIG 7D

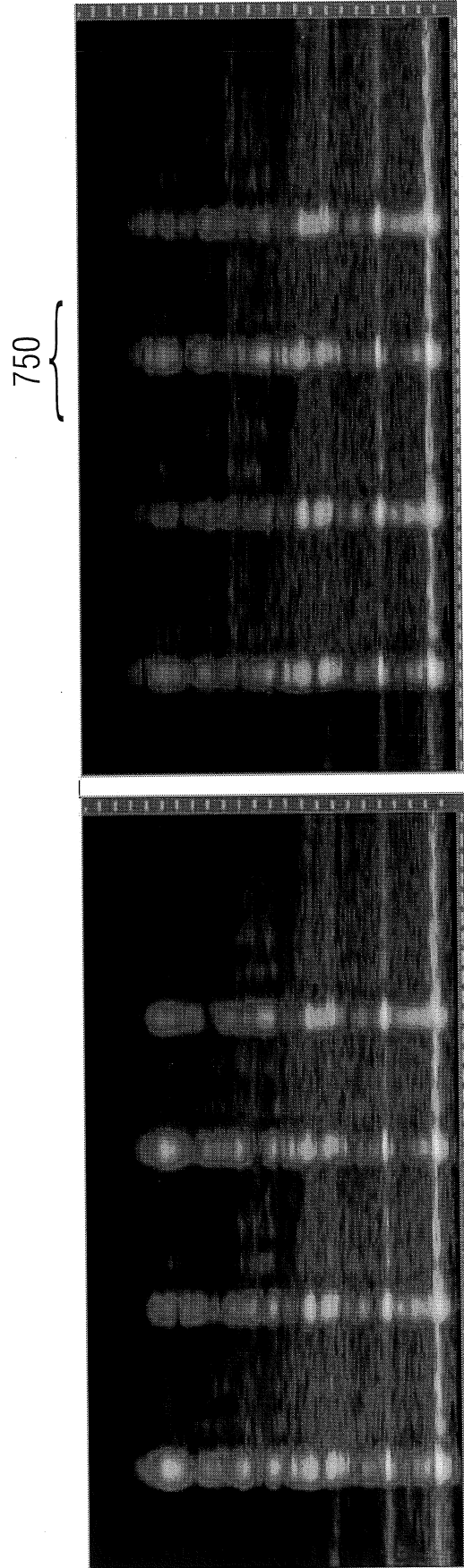


FIG 7E

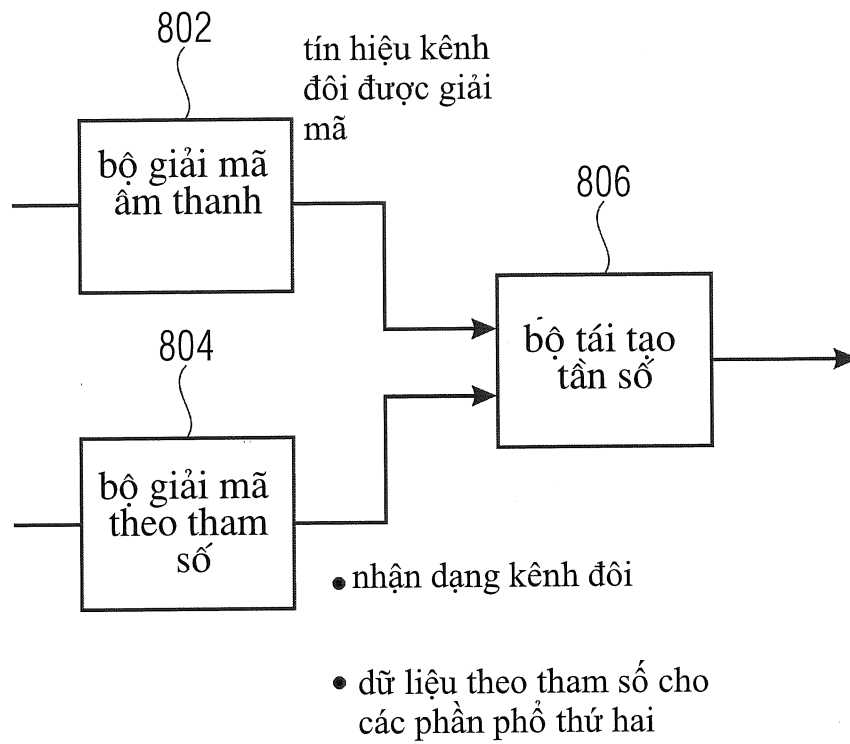


FIG 8A

phạm vi nguồn	1	1	2	2
phạm vi đích	1	2	1	2

1: Phép biểu diễn kênh đôi thứ nhất (ví dụ L/R)

2: việc biểu diễn kênh đôi thứ hai (vd. M/S)

FIG 8B

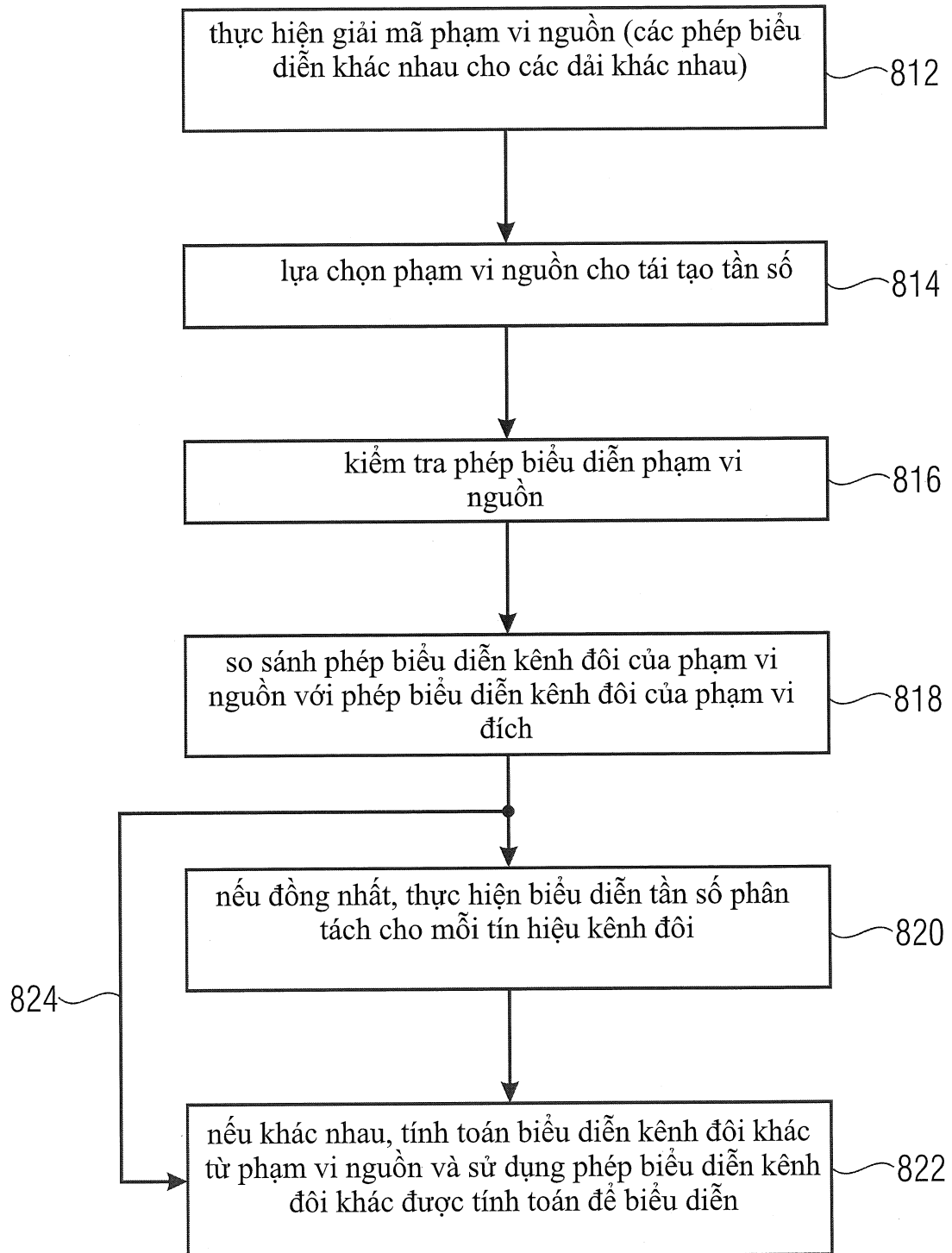


FIG 8C

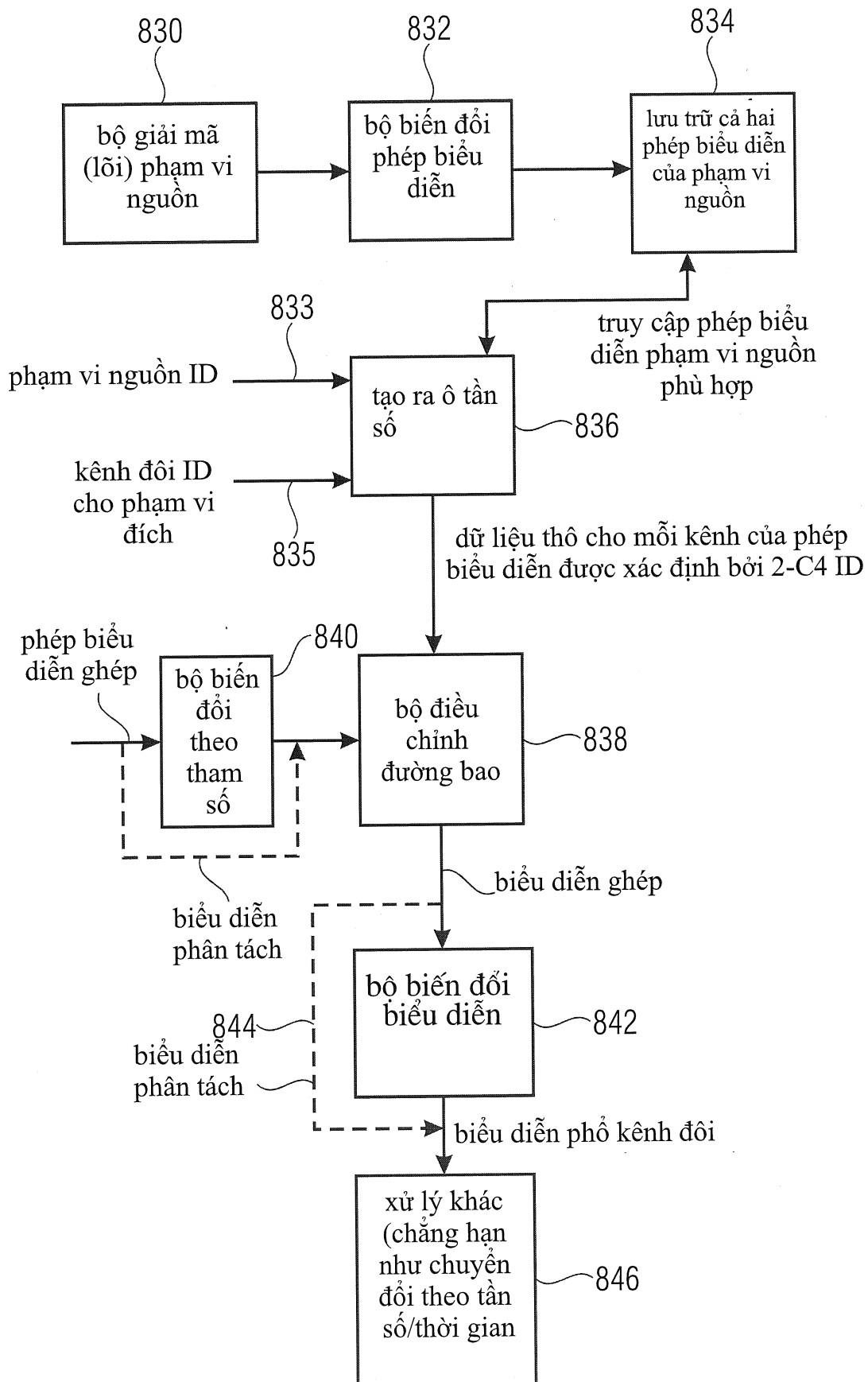


FIG 8D

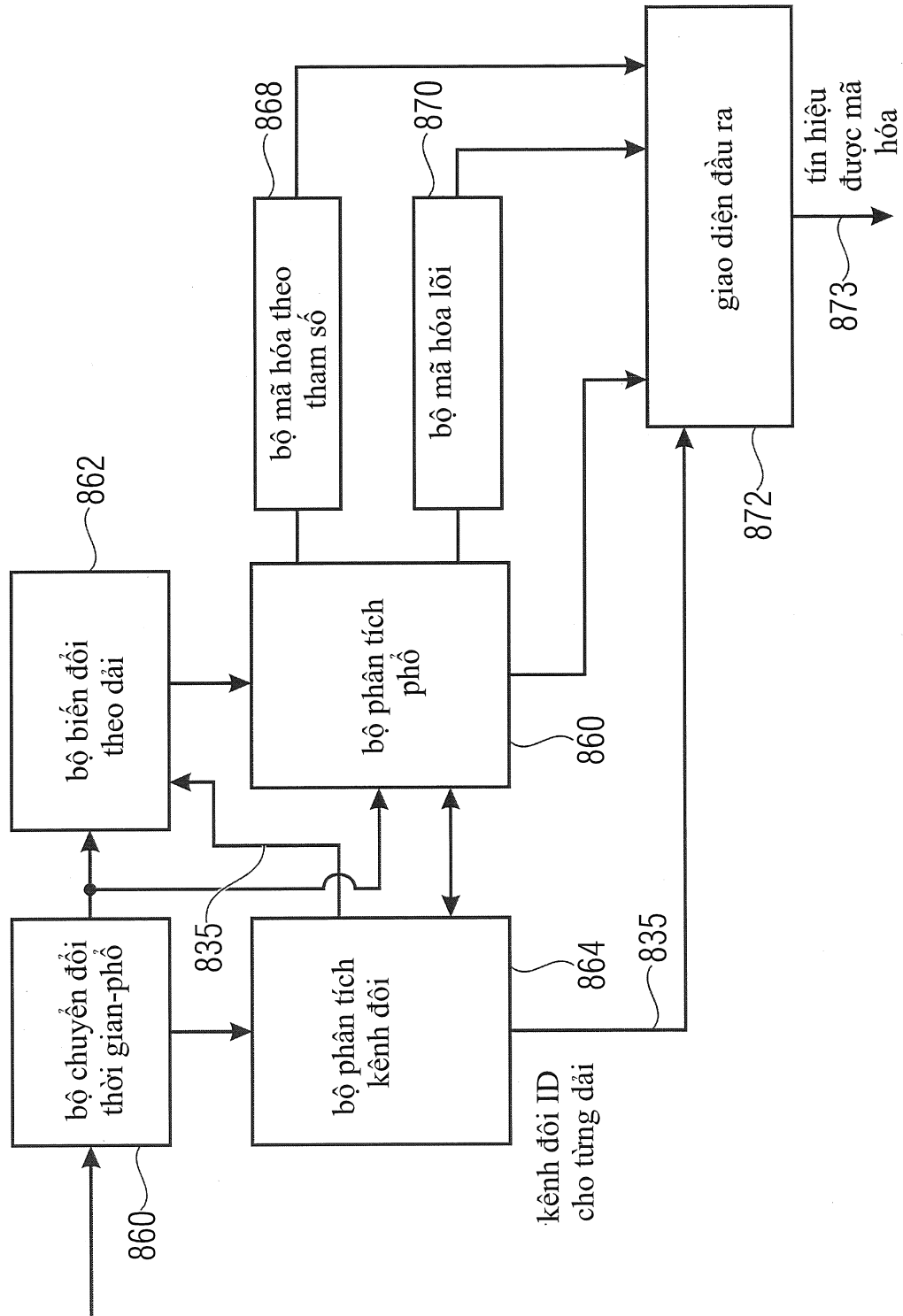


FIG 8E

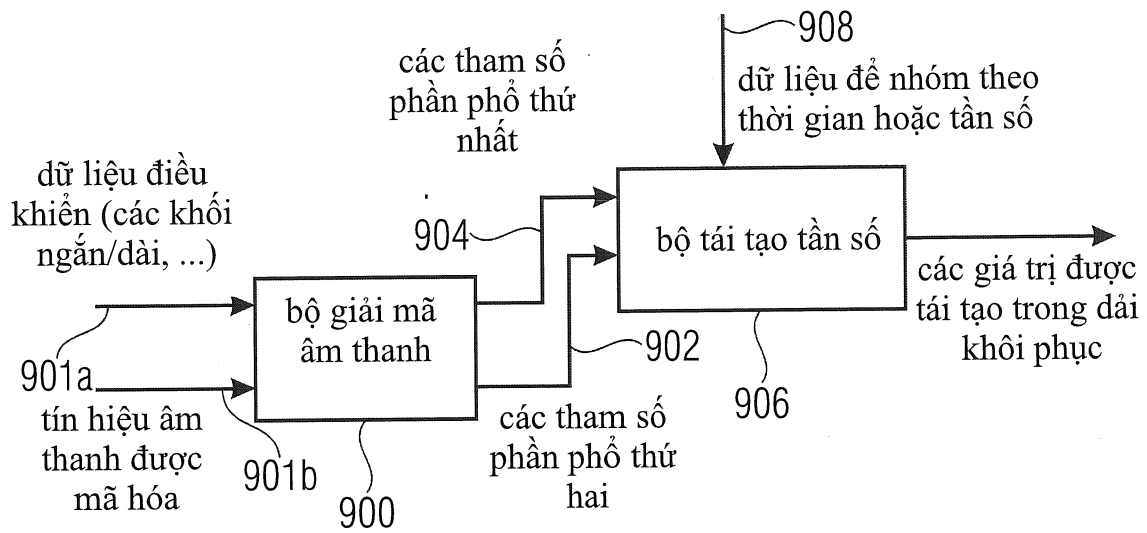


FIG 9A

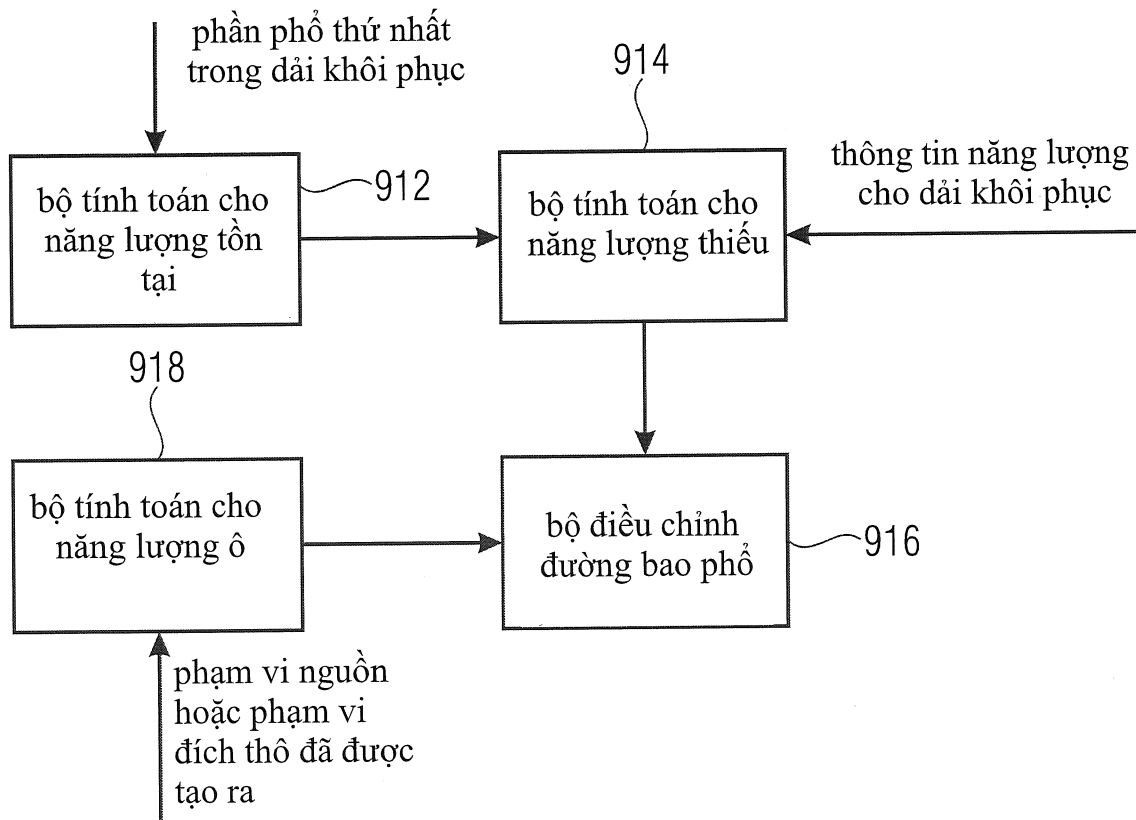
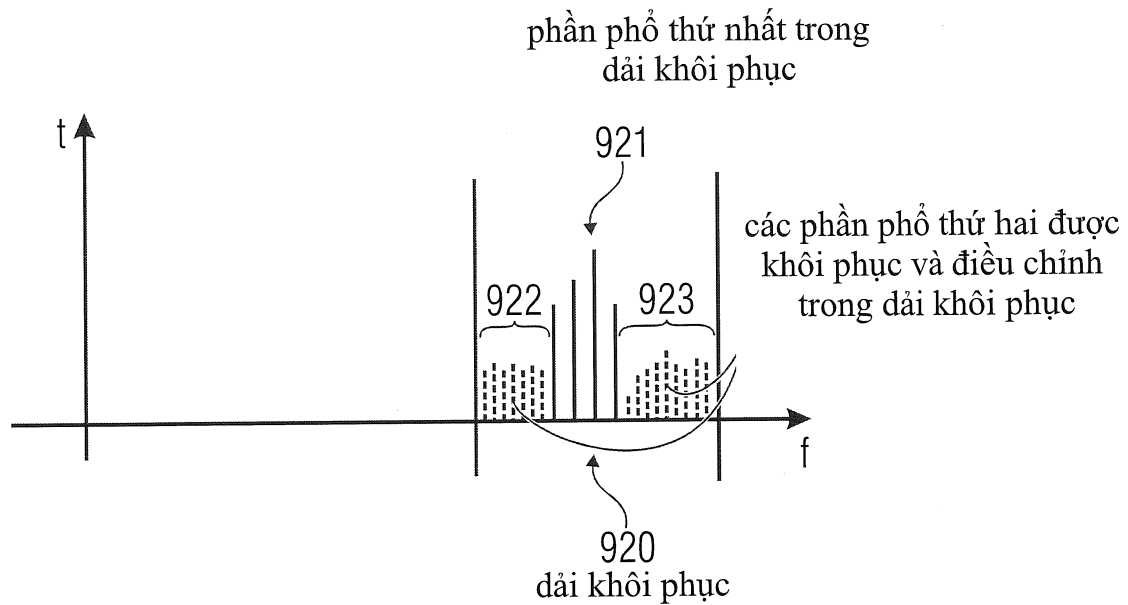


FIG 9B



- vd
- năng lượng còn lại 5 bộ phận
 - giá trị năng lượng cho dải khôi phục: 10 bộ phận (bao phủ các phần phổ thứ nhất và thứ hai trong dải khôi phục)
 - năng lượng của dữ liệu phạm vi nguồn hoặc dữ liệu phạm vi đích thô 8 bộ phận
 - năng lượng thiếu 5 bộ phận
 - hệ số khuếch đại: $g := \sqrt{\frac{mE_k}{pE_k}} = 0.79$

→ chỉ các giá trị phổ cho các phần phổ thứ hai được điều chỉnh

→ phần phổ thứ nhất không bị ảnh hưởng bởi việc điều chỉnh đường bao

FIG 9C

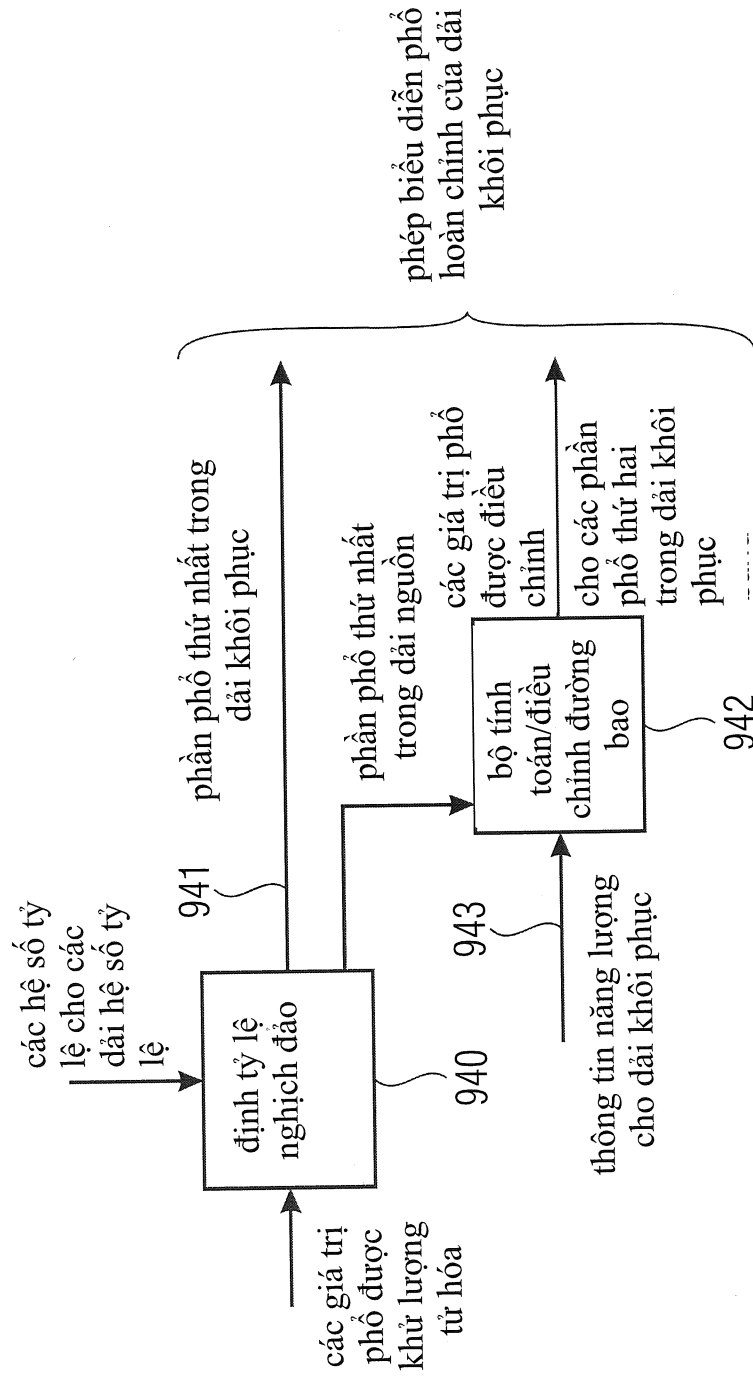


FIG 9D

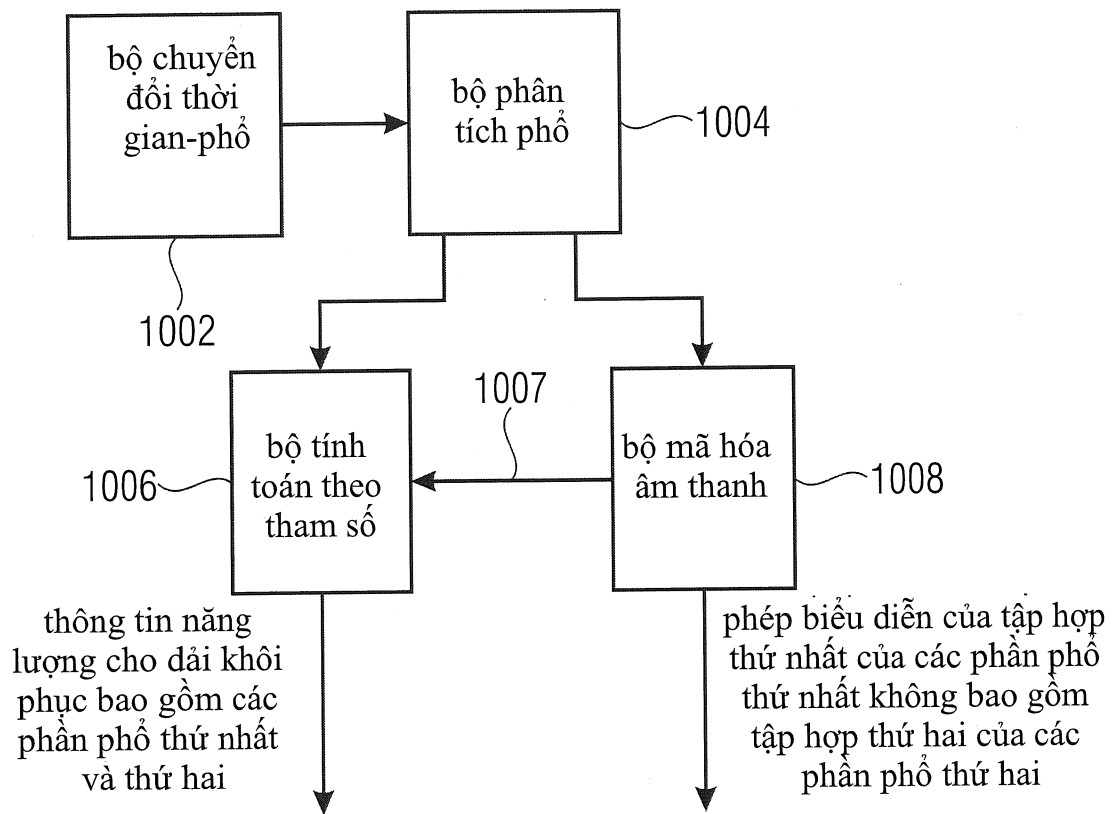


FIG 10A

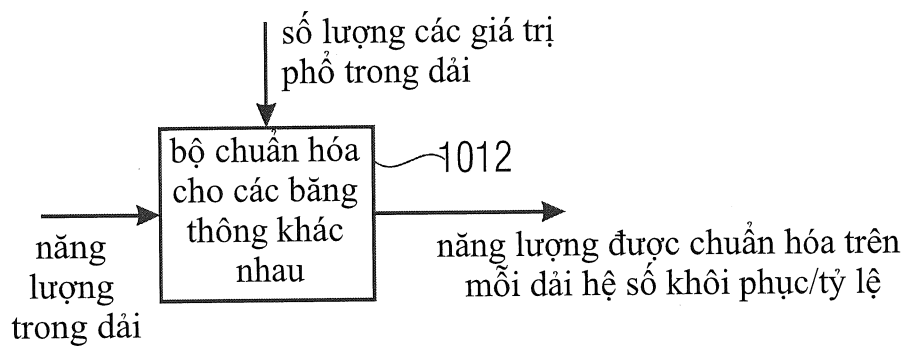


FIG 10B

thông tin điều khiển từ bộ mã hóa lỗi

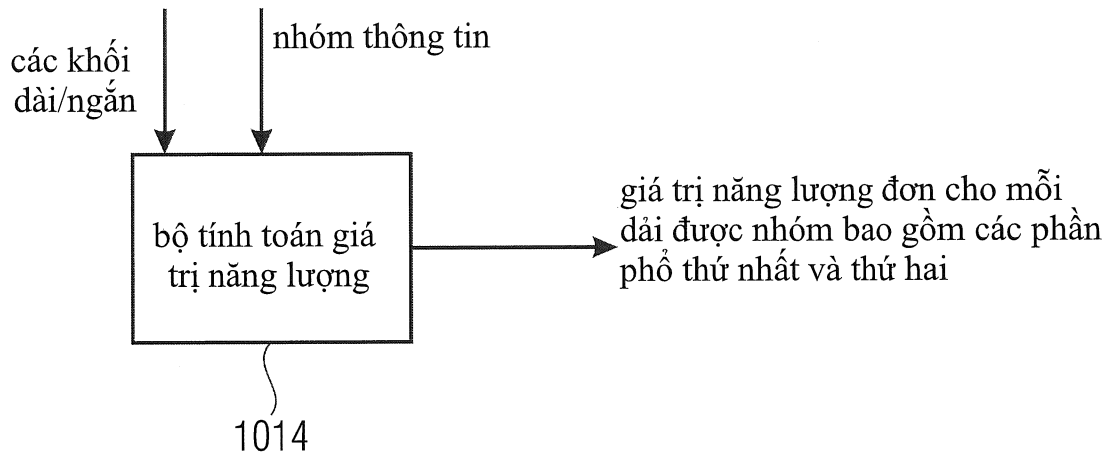


FIG 10C

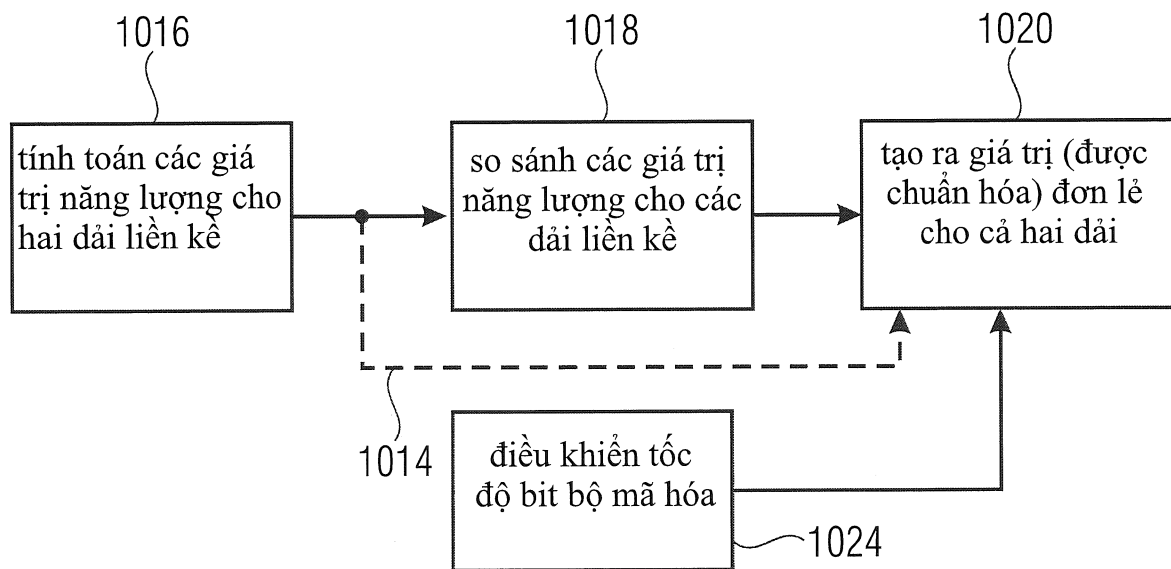


FIG 10D

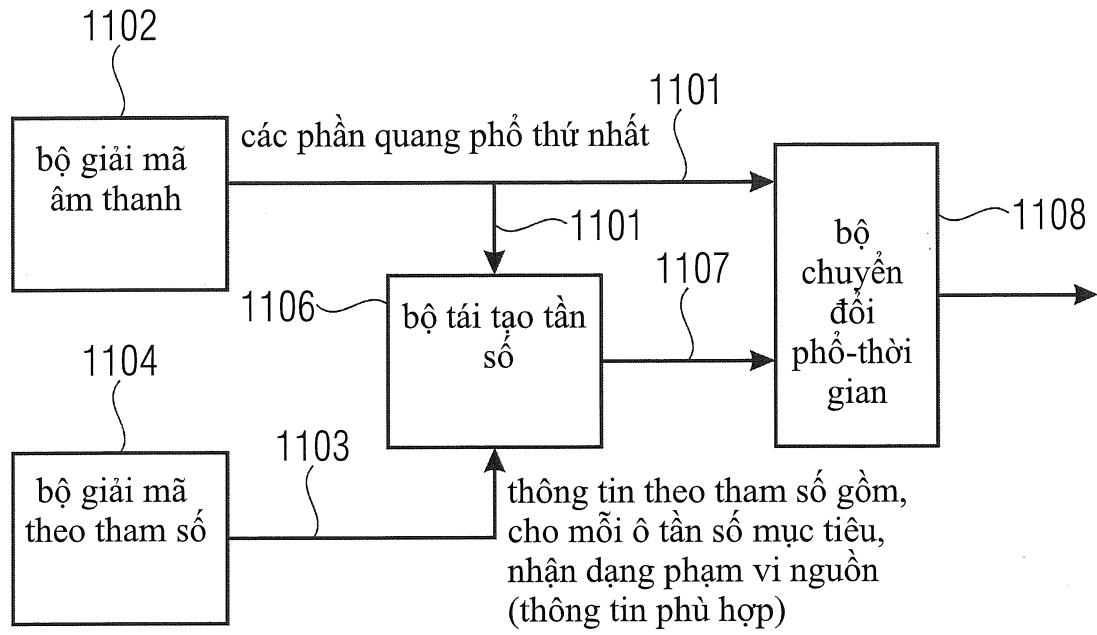


FIG 11A

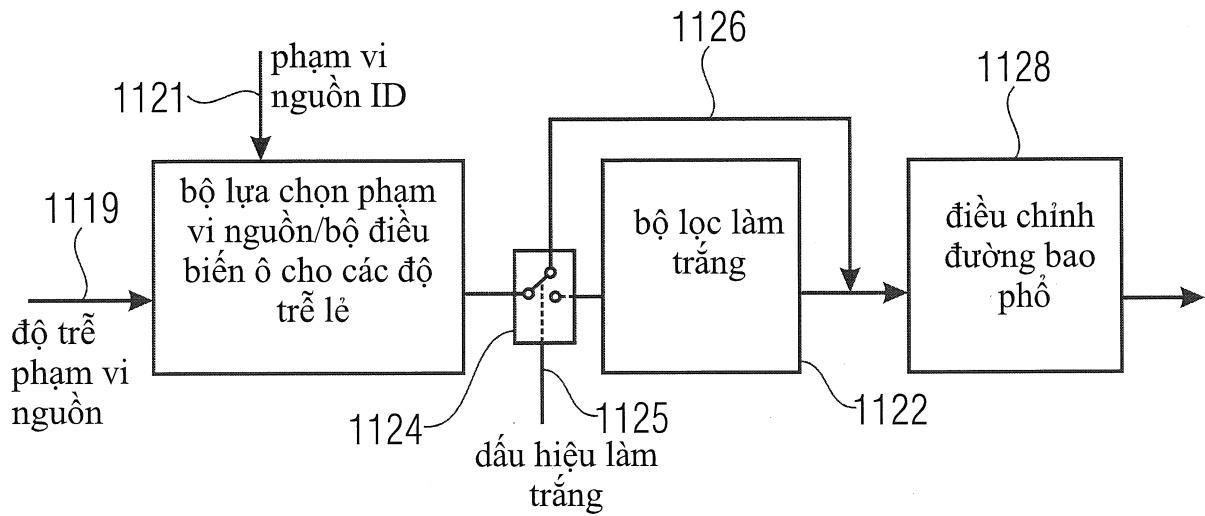


FIG 11B

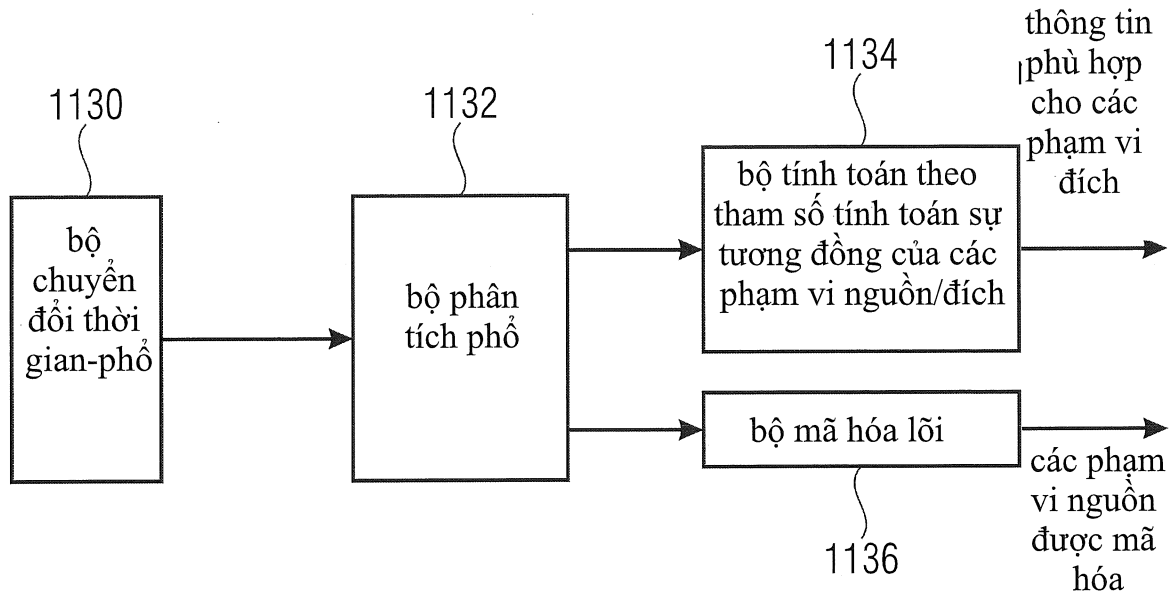


FIG 11C

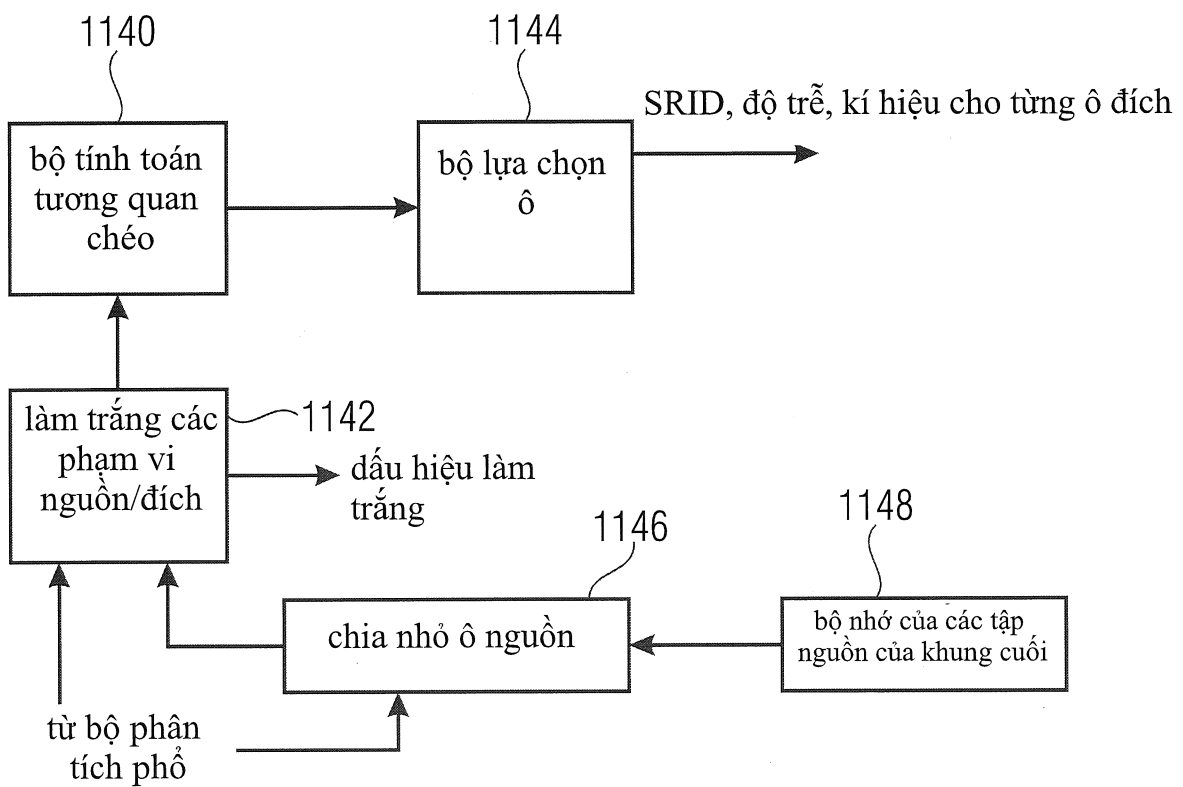


FIG 11D

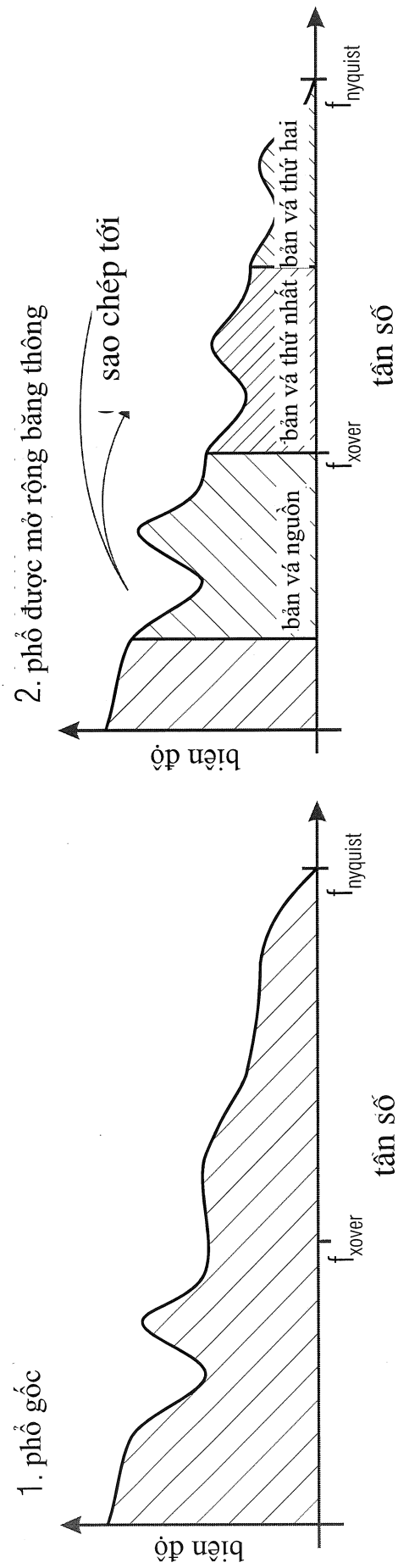


FIG 12A

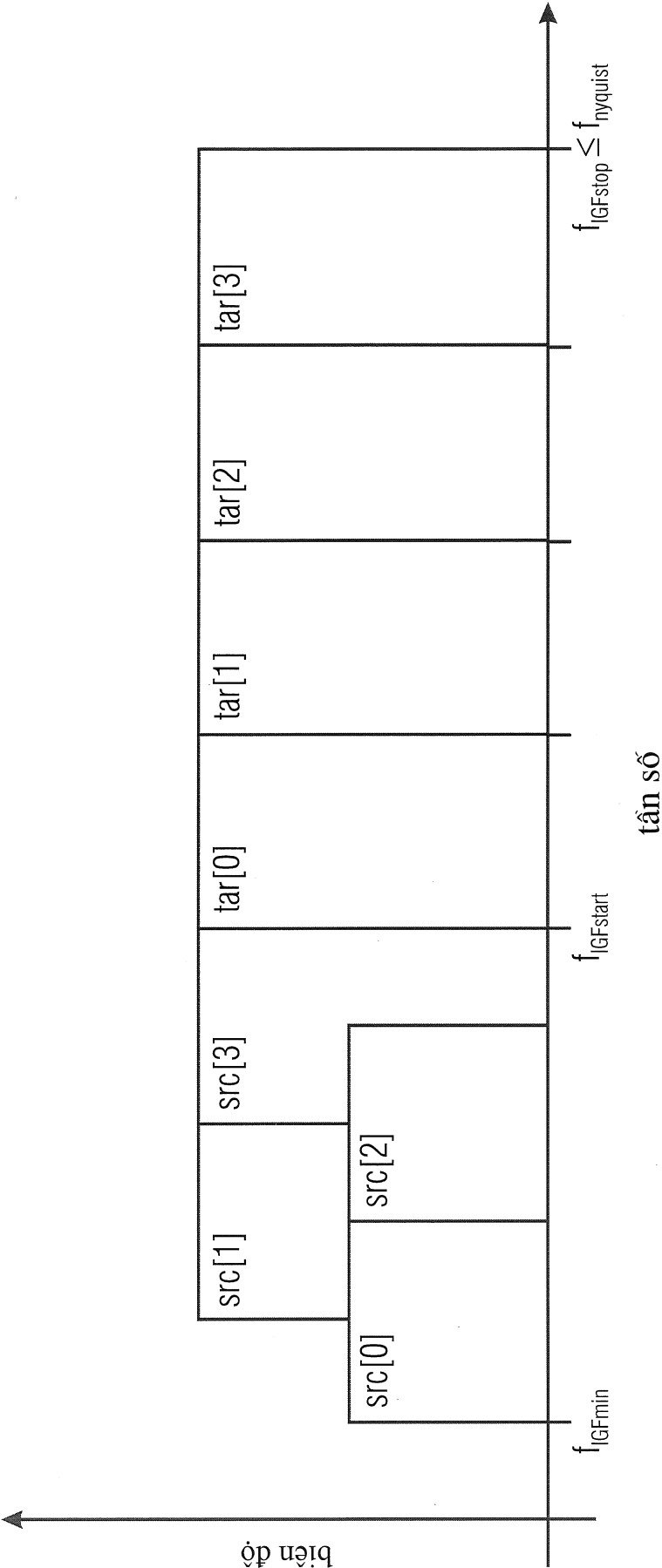


FIG 12B

ví dụ độ tương quan của hai tín hiệu

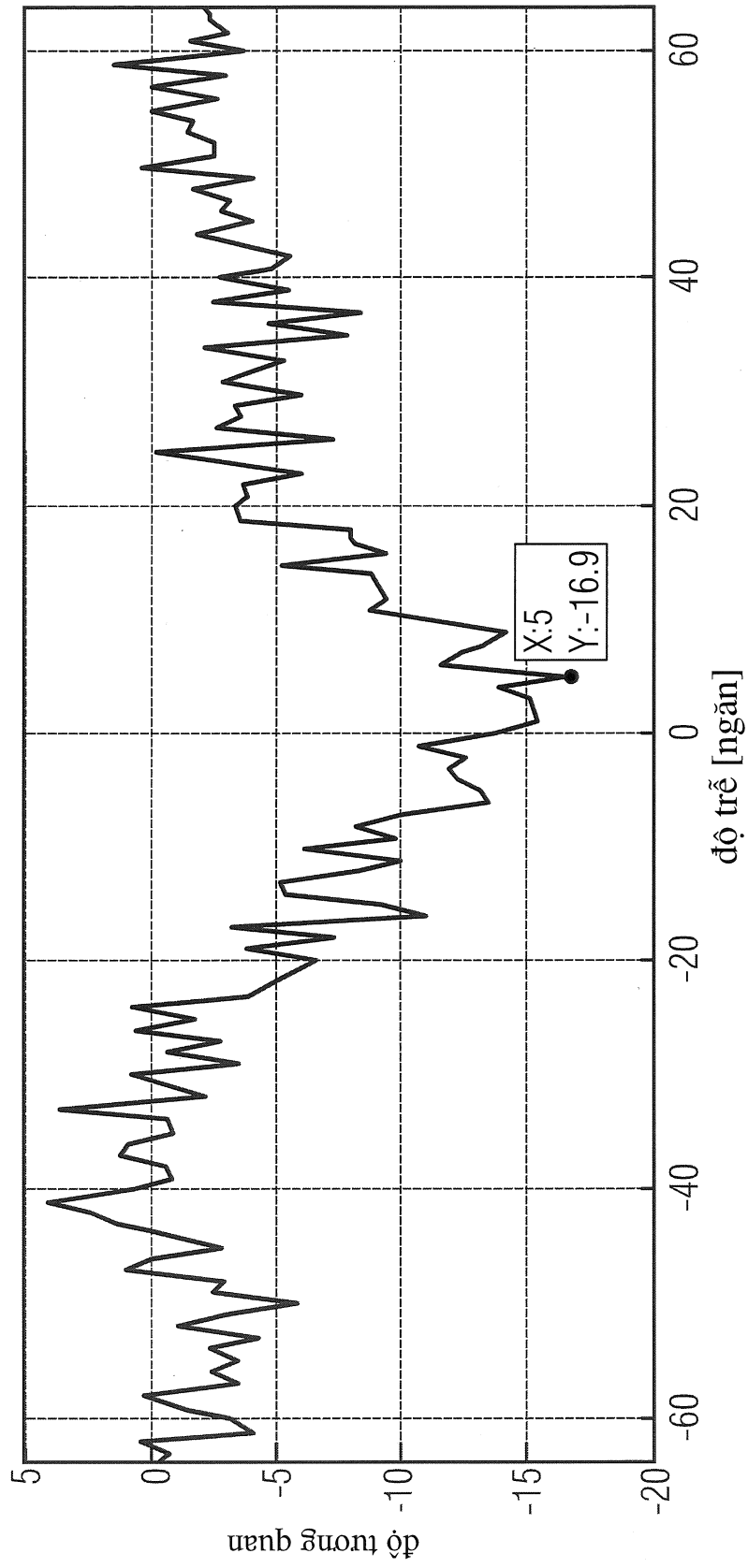


FIG 12C

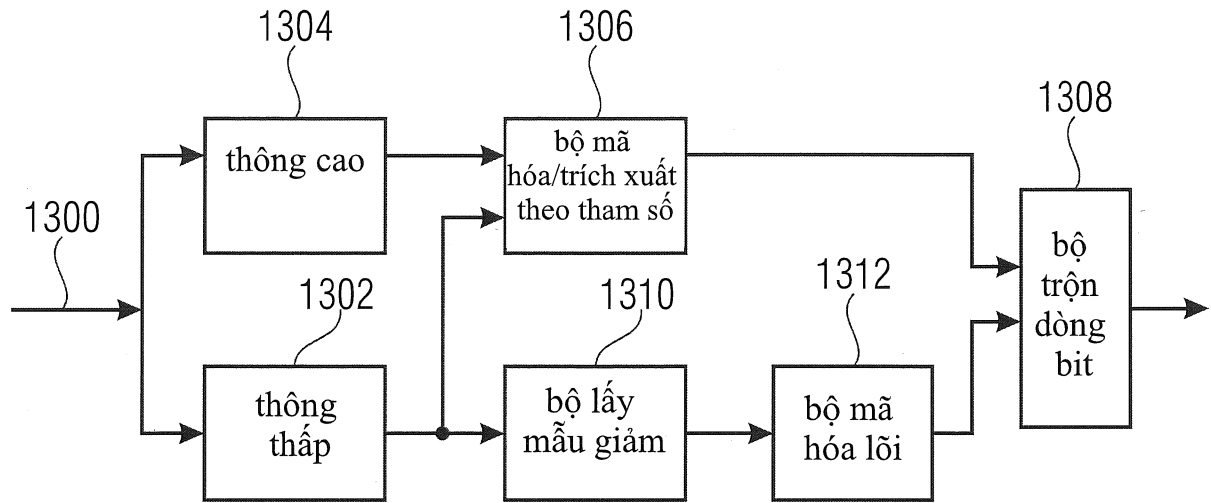


FIG 13A
(tình trạng kỹ thuật)

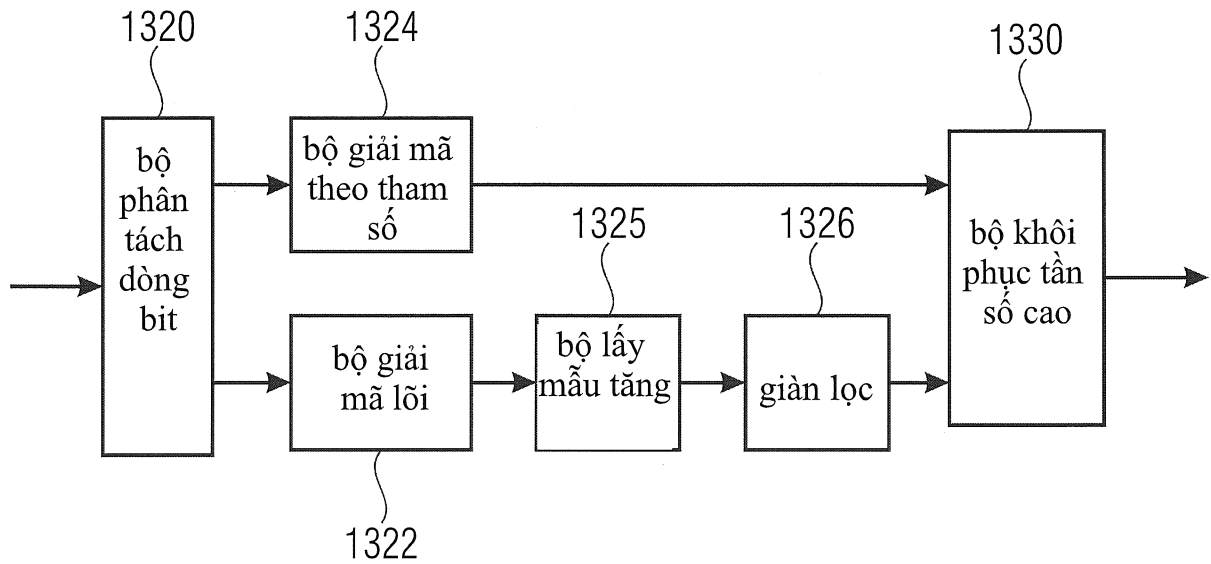


FIG 13B
(tình trạng kỹ thuật)