



- (12) BẢN MÔ TẢ SÁNG CHẾ THUỘC BẰNG ĐỘC QUYỀN SÁNG CHẾ
(19) Cộng hòa xã hội chủ nghĩa Việt Nam (VN) (11) CỤC SỞ HỮU TRÍ TUỆ
(51)⁷ G10L 19/008; G10L 19/18; G10L 21/038; G10L 19/02 (13) B

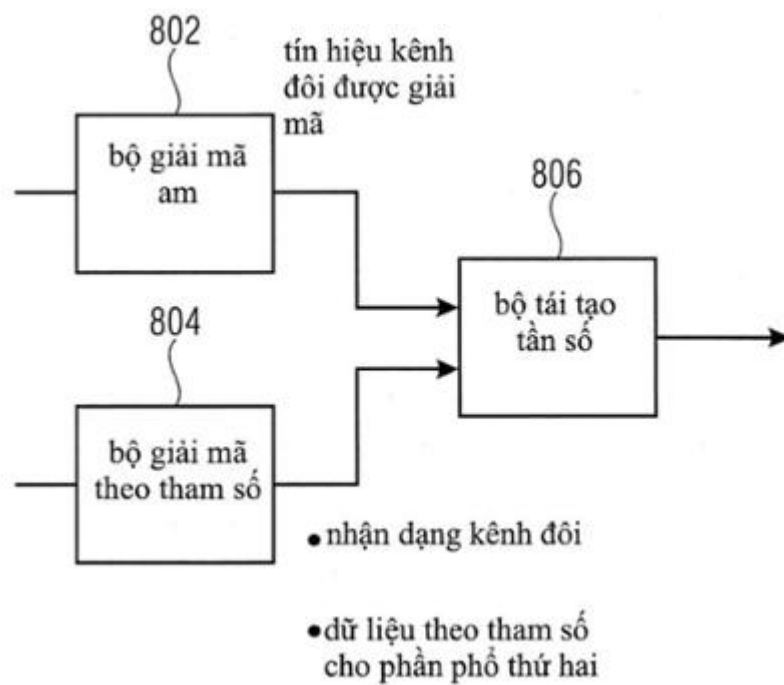


1-0026216

- (21) 1-2016-00385 (22) 15/07/2014
(86) PCT/EP2014/065106 15/07/2014 (87) WO 2015/010947 A1 29/01/2015
(30) 13177346 22/07/2013 EP; 13177348 22/07/2013 EP; 13177350 22/07/2013 EP; 13177353 22/07/2013 EP; 13189366 18/10/2013 EP
(45) 25/11/2020 392 (43) 25/05/2016 338A
(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);
THOSKHAHNA, Balaji Nagendran (IN); SCHMIDT, Konstantin (DE); BAYER,
Stefan (AT); 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 TẠO TÍN HIỆU KÊNH ĐÔI ĐƯỢC GIẢI MÃ, BỘ MÃ HÓA ÂM THANH VÀ PHƯƠNG PHÁP MÃ HÓA TÍN HIỆU ÂM THANH KÊNH ĐÔI

(57) Sáng chế đề cập đến thiết bị và phương pháp tạo tín hiệu kênh đôi được giải mã, bộ mã hóa âm thanh và phương pháp mã hóa tín hiệu âm thanh kênh đôi. Thiết bị tạo tín hiệu kênh đôi được giải mã, bao gồm: bộ xử lý âm thanh (802) để giải mã tín hiệu kênh đôi được mã hóa để thu được tập hợp thứ nhất của các phần phổ thứ nhất; bộ giải mã theo tham số (804) để cung cấp dữ liệu theo tham số cho tập hợp thứ hai của các phần phổ thứ hai và sự nhận dạng kênh đôi hoặc nhận dạng phép biểu diễn kênh đôi khác nhau thứ nhất hoặc thứ hai cho các phần phổ thứ hai; và bộ tái tạo tần số (806) để tái tạo phần phổ thứ hai phụ thuộc vào phần phổ thứ nhất của tập hợp thứ nhất của các phần phổ thứ nhất, dữ liệu theo tham số cho phần thứ hai và sự nhận dạng kênh đôi cho phần thứ hai.



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 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 tâm thính học. Hiện nay, các bộ mã hóa-giải mã âm thanh điển hình cần khoảng 60kbp/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 được nhắm vào việc làm giảm tốc độ bit bằng cách lợi dụng sự đồng dạng 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 được thiết lập để biểu diễn các thành phần tần số cao (high frequency - HF). Phổ HF được điền đầy với nội dung phổ từ các vùng tần số thấp (low frequency - LF) và dạng phổ, độ nghiêng và tính liên tục theo thời gian được điều chỉnh để duy trì âm sắc và sắc thái của tín hiệu ban đầu. Các phương pháp BWE như vậy cho phép các bộ mã hóa-giải mã âm thanh giữ được chất lượng tốt với tốc độ bit thậm chí thấp khoảng 24kbps/kênh.

Việc lưu giữ hoặc truyền các tín hiệu âm thanh thường chịu 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 chỉ có tốc độ bit rất thấp.

Bộ mã hóa-giải mã âm thanh hiện đại ngày nay có khả năng mã hóa được tín hiệu băng rộng bằng cách sử dụng phương pháp mở rộng băng thông BWE [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 ("ráp nối") 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 ráp nối phổ. Điển hình, vùng HF gồm nhiều ráp nối liên kế và từng ráp nối này được bắt nguồn từ các vùng thông dải (band-pass -

BP) của phổ LF thấp hơn tần số giao nhau đã có. Các hệ thống thuộc tình trạng kỹ thuật thực hiện hiệu quả việc ráp nối 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 mục tiêu.

Kỹ thuật khác đã tìm ra các bộ mã hóa-giải mã âm thanh ngày nay mà làm tăng hiệu quả nén và do đó cho phép băng thông âm thanh được mở rộng ở tốc độ bit thấp như sự thay thế tổng hợp được dẫn bởi tham số của các phần phù hợp của phổ âm thanh. Ví dụ, các phần tín hiệu như nhiễu âm của tín hiệu âm thanh ban đầu có thể được thay thế mà không có tổn thất đáng kể về chất lượng chủ quan bởi nhiễu âm nhân tạo được tạo ra trong bộ giải mã và được tỷ lệ bởi các tham số thông tin phụ. Một ví dụ công cụ thay thế nhiễu âm theo cảm giác (Perceptual Noise Substitution - PNS) được chứa trong mã hóa âm thanh tiên tiến (Advanced Audio Coding – AAC) MPEG-4 [5].

Đề xuất thêm nữa cũng cho phép băng thông âm thanh được mở rộng ở tốc độ bit thấp là kỹ thuật điền đầy nhiễu âm được chứa trong mã hóa âm thanh và tiếng nói không đồng 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 tỉ lệ bởi sau xử lý dẫn theo tham số.

Hệ thống theo tình trạng kỹ thuật khác được gọi là thay thế phổ chính xác (Accurate Spectral Replacement - ASR) [2-4]. Ngoài bộ mã hóa-giải mã dạng sóng, ASR sử dụng tầng tổng hợp tín hiệu dành riêng mà phục hồi các phần đường hình sin quan trọng theo cảm giác của tín hiệu tại phía bộ giải mã. Ngoài ra, hệ thống được mô tả trong [5] dựa trên việc mô hình hóa đường hình sin trong vùng HF của bộ mã hóa dạng sóng để cho phép băng thông âm thanh được mở rộng có chất lượng cảm giác tốt tại 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 tầng phân tích/tổng hợp khá phức tạp cho phép 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 tại vạch phổ

1300 được nhập vào hệ thống bộ lọc bao gồm thông thấp 1302 và thông cao 1304. Đầu ra tín hiệu bởi bộ lọc thông cao 1304 được nhập vào bộ trích xuất/bộ lập mã tham số 1306. Bộ trích xuất/bộ lập mã tham số 1306 được tạo cấu hình để tính toán và lập mã các tham số như tham số đường bao phổ, tham số bổ sung nhiễu âm, tham số sóng hài bị thiếu, hoặc tham số lọc nghịch đảo. Các tham số được trích xuất được nhập vào trong bộ dồn kênh dòng bit 1308. Tín hiệu đầu ra thông thấp được nhập vào trong bộ xử lý tiêu biểu gồm có chức năng của bộ lấy mẫu giảm 1310 và bộ lập mã lỗi 1312. Thông thấp 1302 hạn chế băng thông được mã hóa thành băng thông nhỏ hơn đáng kể so với việc xảy ra trong tín hiệu âm thanh đầu vào ban đầu trên vạch phổ 1300. Điều này cung cấp độ khuếch đại mã hóa đáng kể do thực tế rằng toàn bộ các chức năng xảy ra trong bộ mã hóa lỗi chỉ phải hoạt động trên tín hiệu với băng thông được giảm. Khi đó, ví dụ, băng thông của tín hiệu âm thanh trên vạch phổ 1300 là 20 kHz và khi bộ lọc thông thấp 1320 có băng thông là 4 kHz, 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 8 kHz, đó là sự giảm bớt đáng kể để tốc độ lấy mẫu được yêu cầu cho tín hiệu âm thanh 1300 tối thiểu là 40kHz.

Fig.13b minh họa biểu đồ dưới dạng sơ đồ bộ giải mã mở rộng băng thông tương ứng. Bộ giải mã bao gồm bộ dồn kênh dòng bit 1320. Bộ giải ghép dòng bit 1320 trích xuất tín hiệu đầu vào cho bộ giải mã lỗi 1322 và tín hiệu đầu vào cho bộ giải mã tham số 1324. Tín hiệu đầu ra bộ giải mã lỗi có, trong ví dụ ở trên, tốc độ lấy mẫu là 8kHz và, do đó, khi băng thông 4kHz, cho sự khôi phục băng thông hoàn thiện, tín hiệu đầu ra của bộ khôi phục tần số cao 1330 phải ở 20kHz đòi hỏi tốc độ lấy mẫu là ít nhất 40kHz. Để thực hiện tính khả thi này, bộ xử lý giải mã được yêu cầu 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 trên Fig.13a sử dụng biểu diễn tham số của băng tần cao. Bộ khôi phục tần số cao 1330 có chức năng ghép như tái tạo khoảng tần số phía trên sử dụng khoảng nguồn trong khoảng tần số thấp, việc điều chỉnh đường bao phổ, chức năng bổ sung nhiễu 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 trên 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 trên phía bộ giải mã và được đặt chính xác ở giữa băng khôi phục. Vì thế, tất cả các đường sóng hài khuyết được xác định trong băng khôi phục nhất định không được đặt tại giá trị tần số nơi 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 băng nhất định. Do vậy, khi đường sóng hài khuyết trong tín hiệu gốc được đặt rất gần viền băng khôi phục trong tín hiệu gốc, lỗi trong tần số được đưa vào bằng cách đặt đường sóng hài khuyết trong tín hiệu đã được khôi phục tại trung tâm của băng gần tới 50% băng khôi phục riêng, để tham số được tạo ra và được truyền.

Hơn nữa, mặc dù bộ mã hóa lỗi âm thanh điển hình hoạt động trong miền phổ, bộ giải mã lỗi vẫn tạo ra tín hiệu miền thời gian mà sau đó, một lần nữa, được biến đổi thành miền phổ bởi chức năng giàn lọc 1326. Điều này đưa ra 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 yêu cầu lượng tính toán phức tạp đáng kể và bởi nguồn điện, mà đặc biệt là vấn đề khi công nghệ mở rộng băng thông được ứng dụng trong các thiết bị di động như điện thoại di động, máy tính bảng hoặc máy tính xách tay, v.v.

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 không thể thiếu 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 trên tần số giao nhau đã cho được mã hóa dạng sóng. Vì thế, bộ mã hóa-giải mã âm thanh hiện thời hoặc là giảm chi tiết HF hoặc âm sắc khi BWE được thực hiện, vì việc căn chỉnh chính xác của sóng hài âm của tín hiệu không được xem xét trong hầu hết các hệ thống.

Thiếu sót khác của tình trạng hiện tại của hệ thống BWE trong tình trạng kỹ thuật là nhu cầu biến đổi tín hiệu âm thanh thành miền mới để thực hiện BWE (ví dụ chuyển từ miền MDCT sang miền QMF). Điều này dẫn tới sự rắc rối của việc đồng bộ hóa, sự phức tạp tính toán bổ sung và tăng yêu cầu bộ nhớ.

Trong trường hợp các cặp kênh đôi, về cơ bản tồn tại phép biểu diễn một vài kênh như phép biểu diễn kênh ghép hoặc phép biểu diễn kênh phân tách. Phép biểu diễn kênh ghép được biết đến rộng rãi là phép biểu diễn giữa/bên mà trong đó kênh

giữa là tổng của kênh trái và phải và trong đó kênh bên là sự chênh lệch giữa kênh trái và phải.

Phép biểu diễn khác là kênh trộn giảm và kênh dư và hệ số dự báo bổ sung mà cho phép tạo lại kênh trái và phải từ sự trộn giảm và phần dư. Phép biểu diễn phân tách là, trong trường hợp này, kênh phân tách trái và phải hoặc thường là, kênh thứ nhất và kênh thứ hai.

Ngoài ra, có trường hợp, mà khoảng nguồn cho cho các hoạt động làm đầy khoảng trống thể hiện mối tương quan mạnh mẽ, trong khi các khoảng mục tiêu không thể hệ mối tương quan mạnh mẽ này. Trong phương án này, khi khoảng nguồn được mã hóa sử dụng phép biểu diễn âm lập thể như sự biểu diễn giữa/bên để giảm tốc độ truyền dữ liệu đối với phần tần số lõi, sau đó hình ảnh kênh đôi sai được tạo cho phần khôi phục hoặc khoảng mục tiêu. Khi đó, mặt khác, khoảng nguồn không thể hiện bất kỳ tương quan hoặc chỉ tương quan nhỏ và khoảng mục tiêu có tương quan nhỏ hoặc không tương quan, sau đó một lần nữa hoạt động điền đầy khoảng trống đơn giản sẽ tạo ra các thành phần lại.

Tài liệu tham khảo

- [1] Dietz, L. Liljeryd, K. Kjörling and O. Kunz, "Spectral Band Replication, a novel approach in audio coding," in 112th AES Convention, Munich, May 2002.
- [2] Ferreira, D. Sinha, "Accurate Spectral Replacement", Audio Engineering Society Convention, Barcelona, Spain 2005.
- [3] D. Sinha, A. Ferreira¹ and E. Harinarayanan, "A Novel Integrated Audio Bandwidth Extension Toolkit (ABET)", Audio Engineering Society Convention, Paris, France 2006.
- [4] R. Annadana, E. Harinarayanan, A. Ferreira and D. Sinha, "New Results in Low Bit Rate Speech Coding and Bandwidth Extension", Audio Engineering Society Convention, San Francisco, USA 2006.
- [5] T. Żernicki, M. Bartkowiak, "Audio bandwidth extension by frequency scaling of sinusoidal partials", Audio Engineering Society Convention, San Francisco, USA 2008.
- [6] J. Herre, D. Schulz, Extending the MPEG-4 AAC Codec by Perceptual Noise Substitution, 104th AES Convention, Amsterdam, 1998, Preprint 4720.

- [7] M. Neuendorf, M. Multrus, N. Rettelbach, et al., MPEG Unified Speech and Audio Coding-The ISO/MPEG Standard for High-Efficiency Audio Coding of all Content Types, 132nd AES Convention, Budapest, Hungary, April, 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), August 1986.
- [9] Smith, J.O., Serra, X. "PARSHL: An analysis/synthesis program for non-harmonic sounds based on a sinusoidal representation", Proceedings of the International Computer Music Conference, 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", 101st AES Convention, Los Angeles 1996
- [13] J. Herre, "Temporal Noise Shaping, Quantization and Coding methods in Perceptual Audio Coding: A Tutorial introduction", 17th AES International Conference on High Quality Audio Coding, August 1999
- [14] J. Herre, "Temporal Noise Shaping, Quantization and Coding methods in Perceptual Audio Coding: A Tutorial introduction", 17th AES International Conference on High Quality Audio Coding, August 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", in Proceedings of 1st IEEE Benelux Workshop on MPCA, Leuven, November 2002

[18] F. Nagel, S. Disch, S. Wilde, A continuous modulated single sideband bandwidth extension, ICASSP International Conference on Acoustics, Speech and Signal Processing, Dallas, Texas (USA), April 2010

Bản chất kỹ thuật của sáng chế

Do đó, mục tiêu của sáng chế là đề xuất khái niệm bộ mã hóa/bộ giải mã được cải thiện cho các phép biểu diễn kênh đôi.

Mục tiêu này đạt được nhờ thiết bị giải mã điểm 1, bộ mã hóa âm thanh theo điểm 14, phương pháp giải mã theo điểm 19, phương pháp mã hóa theo điểm 20 hoặc chương trình máy tính theo điểm 21.

Sáng chế dựa trên việc tìm ra rằng trạng thái tương quan không chỉ quan trọng đối với khoảng nguồn mà còn quan trọng đối với cả khoảng mục tiêu. Hơn nữa, sáng chế này xác nhận trạng thái rằng trạng thái tương quan khác nhau có thể xảy ra trong khoảng nguồn hoặc khoảng mục tiêu. Ví dụ, khi tín hiệu tiếng nói với nhiều âm tần số cao được xem xét, trạng thái có thể là băng 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 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 được thực hiện mà bỏ qua trạng thái này, thì phần tần số cao cũng sẽ tương quan, và điều này có thể tạo ra giả âm phân chia không gian nghiêm ngặt trong tín hiệu được khôi phục. Để giải quyết vấn đề này, dữ liệu theo tham số cho băng khôi phục 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 băng khôi phục. Trên phía bộ mã hóa, vì thế việc nhận dạng kênh đôi được tính toán cho các phần phổ thứ hai, tức là, cho các phần mà bổ sung thông tin năng lượng cho các băng khôi phục được tính toán. Bộ tái tạo tần số trên 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, khoảng 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 băng khôi phục dưới việc xem xét lại.

Việc nhận dạng kênh đôi tốt hơn là được truyền dưới dạng dấu hiệu cho từng băng 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 được biểu thị bởi dấu hiệu tính toán tốt nhất cho các băng lỗi. Sau đó, trong việc thực hiện, tín hiệu lỗi được lưu giữ 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, phép biểu diễn ô nguồn được chọn để thích hợp với phép biểu diễn ô mục tiêu như được biểu thị bởi các dấu hiệu nhận dạng kênh đôi cho điền đầy khoảng trống thông minh hoặc các băng khôi phục, tức là, cho khoảng mục tiêu.

Cần nhấn mạnh rằng quy trình này không chỉ thực hiện với các tín hiệu âm lập thể, tức là, cho kênh trái và kênh phải mà còn hoạt động với 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 chẳng hạn như 7.1, 11.1, v.v.

Khía cạnh khác dựa trên việc phát hiện ra rằng các vấn đề liên quan đến việc phân tách phần mở rộng băng thông là một mặt và mặt khác là việc mã hóa lỗi có thể được giải quyết và khắc phục bằng việc thực hiện mở rộng băng thông trong miền phổ giống nhau mà trong đó bộ giải mã lỗi hoạt động. Do đó, bộ giải mã lỗi tốc độ đầy đủ được đề xuất là bộ mã hóa và giải mã khoảng tín hiệu âm thanh đầy đủ. 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, tất cả việc xử lý được thực hiện trong tốc độ lấy mẫu đầy đủ hoặc miền băng thông đầy đủ. 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 trong cách thức 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. Trên phía bộ giải mã, bộ giải mã lỗi, là bộ giải mã băng 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 cứ hiểu biết rằng có bất kỳ tái tạo tần số bổ sung nào. Tuy nhiên, phổ đã tạo ra có nhiều khoảng trống phổ. Những khoảng trống này sau đó được điền đầy với công nghệ điền đầy khoảng trống thông minh sáng tạo (Intelligent Gap Filling - 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 nhiều âm thay vì mô phỏng 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 nhiều âm mà không cần tái tạo tần số hoặc bằng cách tái tạo tần số sử dụng ô tần số tại khoảng tần số khác nhau.

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

Khía cạnh khác dựa trên việc phát hiện ra rằng chất lượng âm thanh của tín hiệu được khôi phục có thể được cải thiện qua IGF vì toàn bộ phổ có thể tới được bộ mã hóa lỗi, để mà, ví dụ, các phần âm quan trọng theo cảm giác trong khoảng phổ cao có thể vẫn được mã hóa bởi bộ mã hóa lỗi hơn là thay thế tham số. Ngoài ra, hoạt động điền đầy khoảng trống sử dụng các ô tần số từ tập hợp thứ nhất của các phần phổ thứ nhất mà, ví dụ, tập hợp các thành phần âm thanh thường từ phạm vi tần số thấp hơn,

nhưng cũng từ phạm vi tần số cao hơn nếu có thể dùng được, được thực hiện. Đối với sự điều chỉnh đường bao phổ trên phía bộ giải mã, tuy nhiên, các phần phổ từ tập hợp thứ nhất của các phần phổ được định vị trong băng khôi phục không được xử lý sau thêm nữa bằng ví dụ điều chỉnh đường bao phổ. Chỉ các giá trị phổ còn lại trong băng khôi phục mà không bắt nguồn từ bộ giải mã lỗi là được điều chỉnh đường bao sử dụng thông tin đường bao. Tốt hơn là, thông tin đường bao là thông tin đường bao băng đầy đủ tính cho năng lượng của tập hợp thứ nhất của các phần phổ thứ nhất trong băng khôi phục và tập hợp thứ hai của các phần phổ thứ hai trong băng khôi phục tương tự, tại đó các giá trị phổ sau cùng trong tập hợp thứ hai của các phần phổ thứ hai được chỉ ra là không và do đó không được mã hóa bởi bộ mã hóa lỗi, nhưng được tạo mã theo tham số với thông tin năng lượng phân giải thấp.

Người ta đã phát hiện ra rằng các giá trị năng lượng tuyệt đối, hoặc được chuẩn hóa với băng thông của băng tương ứng hoặc không được chuẩn hóa, là hữu dụng và rất hiệu quả trong việc ứng dụng trên phía bộ giải mã. Các ứng dụng đặc biệt này khi các hệ số khếch đại phải được tính toán dựa trên năng lượng còn lại trong băng khôi phục, năng lượng khuyết trong băng khôi phục và thông tin ô tần số trong băng khôi phục.

Hơn nữa, tốt hơn là dòng bit được mã hóa không chỉ bao trùm thông tin năng lượng cho các băng khôi phục, mà ngoài ra còn có, các thừa số định tỉ lệ cho các băng định tỉ lệ mở rộng đến tần số cực đại. Điều này đảm bảo rằng với mỗi băng khôi phục, mà cho phần âm nhất định, tức là phần phổ thứ nhất là sẵn có, tập hợp thứ nhất này của phần phổ thứ nhất có thể thực sự được giải mã với biên độ đúng. Hơn nữa, ngoài thừa số định tỉ lệ cho mỗi băng khôi phục, năng lượng cho băng khôi phục này được tạo ra trong bộ mã hóa và được truyền đến bộ giải mã. Hơn nữa, tốt hơn là các băng khôi phục trùng với các băng thừa số định tỉ lệ hoặc trong trường hợp nhóm năng lượng, ít nhất các đường biên của băng khôi phục trùng với đường biên của các băng thừa số định tỉ lệ.

Một khía cạnh nữa dựa trên việc tìm ra những khiếm khuyết nhất định trong chất lượng âm thanh có thể được khắc phục bằng cách ứng dụng sơ đồ điền đầy ô tần số thích ứng. Để đạt được điều này, sự phân tích trên phía bộ mã hóa được thực hiện để tìm ra ứng viên vùng nguồn phù hợp nhất cho vùng mục tiêu nhất định. Việc nhận

dạng thông tin phù hợp cho vùng mục tiêu vùng nguồn nhất định cùng với nhau với việc tùy chọn một số thông tin bổ sung được tạo ra và truyền dưới dạng 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 băng 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 băng khôi phục. Sau đó, kết quả này của hoạt động điền đầy ô tần số, cụ thể, dữ liệu phổ thô cho băng khôi phục, được tạo hình sử dụng thông tin đường bao phổ để cuối cùng thu được băng 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 hạt thấp. Trong việc thực hiện này, vùng nguồn được chia nhỏ thành vùng nguồn chồng lấp điển hình và vùng mục tiêu hoặc các băng khôi phục có được bởi những vùng mục tiêu tần số không chồng lấp. Sau đó, những điểm tương đồng giữa từng vùng nguồn và từng vùng mục tiêu được xác định trên phía bộ mã hóa và cặp phù hợp nhất của vùng nguồn và vùng mục tiêu được nhận dạng bằng thông tin phù hợp và, trên 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 băng khôi phục.

Với mục tiêu thu được độ hạt 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 phù hợp thậm chí tốt hơn giữa vùng nguồn và vùng mục tiêu.

Thêm nữa, 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, thậm chí dấu hiệu có thể được truyền. Khi dấu hiệu được xác định là âm trên phía bộ mã hóa, sau đó dấu hiệu biểu hiện tương quan cũng được truyền trong thông tin phù hợp và, trên phía bộ giải mã, các giá trị phổ vùng nguồn được nhân lên với "-1" hoặc, trong biểu diễn phức hợp, được "xoay" 180 độ.

Việc thực hiện tiếp theo của sáng chế này áp dụng việc hoạt động làm trắng ô. Việc làm trắng phổ loại bỏ thông tin đường bao phổ thô và nhấn mạnh cấu trúc phổ tinh mà quan tâm trước nhất cho việc ước lượng ô giống nhau. 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 sử dụng quy trình được xác định trước, dấu hiệu làm trắng được truyền biểu thị tới bộ giải mã mà để chu trình làm trắng được xác định trước tương tự 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 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 điều biến bổ sung thông qua phép nhân bởi chuỗi theo thời gian thay thế của $-1/1$ để bù cho phép biểu diễn nghịch đảo tần số của mọi băng khác trong MDCT. Thêm nữa, tín hiệu của kết quả tương quan được áp dụng khi tạo ra ô tần số.

Ngoài ra, ưu tiên để sử dụng lược bớt ô và làm ổn định để chắc chắn rằng tránh được giả âm được tạo ra bởi các vùng nguồn thay đổi nhanh cho vùng khôi phục giống nhau hoặc vùng mục tiêu giống nhau. 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, sau đó ô nguồn này có thể được giảm từ bộ của các ô nguồn tiềm năng vì nó tương quan cao với các ô nguồn khác. Thêm nữa, như một loại 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 mục tiêu trong khung hiện tại.

Khía cạnh khác dựa trên việc phát hiện rằng chất lượng được cải thiện và tốc độ bit được giảm đặc biệt cho các tín hiệu bao gồm các phần tạm thời vì chúng xuất hiện rất thường xuyên trong các tín hiệu âm thanh mà thu được bằng cách tổ 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 trên phía bộ mã hóa được thực hiện bằng việc dự báo trên tần số khôi phục đường bao thời gian của tín hiệu âm thanh. Phụ thuộc vào việc thực hiện, tức là, khi bộ lọc tạo hình nhiễu âm tần số theo thời gian được xác định trong khoảng tần số không chỉ bao trùm khoảng tần số nguồn mà còn có khoảng tần số mục tiêu được khôi phục

trong bộ giải mã tái tạo tần số, đường bao theo thời gian không chỉ áp dụng tín hiệu âm thanh lỗi lên đến 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 được khôi phục. Do vậy, tiếng vang trước hoặc tiếng vang sau có thể sẽ xuất hiện mà không có việc tạo hình 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 lỗi 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 trên phía bộ giải mã trước khi áp dụng dự báo trên tần số. Tuy nhiên, tần số trên 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 trên một hoặc nhiều ô tần số cũng thiết lập tính liên tục của mối tương quan giữa khoảng nguồn và khoảng khôi phục hoặc trong hai khoảng khôi phục liền kề hoặc ô tần số liền kề.

Trong việc thực hiện, tốt hơn là sử dụng dạng lọc TNS/TTS phức hợp. Bằng cách đó, tránh được các giả âm sai số lấy mẫu (theo thời gian) của phép biểu diễn thực tế đã lấy mẫu tới hạn, như MDCT. Bộ lọc TNS phức hợp có thể được tính toán trên phía bộ mã hóa bằng việc áp dụng không chỉ biến đổi cosin rời rạc được thay đổi mà còn là biến đổi sin rời rạc được thay đổi ngoài ra để thu được biến đổi được thay đổi phức hợp. Tuy nhiên, chỉ các giá trị biến đổi cosin rời rạc được thay đổi, tức là, phần thực của biến đổi phức hợp được truyền. Tuy nhiên, về phía bộ giải mã, có khả năng ước lượng phần ảo của biến đổi sử dụng phổ MDCT của các khung đứng trước hoặc khung kế tiếp để mà, trên phía bộ giải mã, bộ lọc phức hợp có thể lại được áp dụng trong dự báo nghịch đảo trên tần số và, cụ thể, dự báo trên đường biên giữa khoảng nguồn và khoảng khôi phục và cũng trên đường biên giữa ô tần số liền kề tần số trong khoảng khôi phục.

Hệ thống mã hóa âm thanh theo sáng chế mã hóa hiệu quả tín hiệu âm thanh tùy ý tại khoảng rộng của tốc độ bit. Trong khi, 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. Vì thế, phần chính của tốc độ bit có sẵn được sử dụng để mã

hóa dạng sóng chỉ cấu trúc liên quan chủ yếu đến cảm giác của tín hiệu trong bộ mã hóa, và các khoảng trống phổ thu được được điền đầy trong bộ giải mã với lượng tín hiệu mà xấp xỉ phổ ban đầu. Gói bit rất hạn chế được dùng để kiểm soát tham số được dẫn được gọi là điền đầy khoảng trống thông minh (IGF) phổ bằng thông tin phụ dành riêng 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 được mô tả liên tiếp với việc tham chiếu tới 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 đã mã hóa phù hợp với bộ mã hóa trên 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 trên Fig.1b;

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

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

Fig.4b minh họa việc thực hiện các chức năng trên 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 trên 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 phép biểu diễn tần số miêu 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 minh họa các tổ hợp khác nhau của phép biểu diễn và các khoảng nguồn/đích;

Fig.8c minh họa lưu đồ minh họa chức năng của bộ giải mã kênh đôi với sự tái tạo tần số trên Fig.8a;

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

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

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

Fig.9c minh họa sơ đồ minh họa các chức năng trên Fig.9b;

Fig.9d minh họa việc thực hiện khác của bộ giải mã trên Fig.9a;

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

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

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

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

Fig.11a minh họa bộ giải mã khác có sự nhận dạng khoảng 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 trên Fig.11a;

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

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

Fig.12a và Fig.12b minh họa các phép toán tần số để minh họa khoảng nguồn và khoảng mục tiêu;

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 được nhập 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 được xuất ra bởi bộ biến đổi phổ theo thời gian. Phổ 101 được nhập 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 được nhập vào bộ tính toán theo tham số hoặc bộ mã hóa theo tham số 104 để tính toán thông tin đường bao phổ có độ phân giải phổ thứ hai. Hơn nữa, bộ mã hóa âm thanh miền phổ 106 được 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. Thêm nữa, 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 được nhập 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 trên Fig.3a sẽ được bao quanh bởi hai phần phổ thứ hai là 307a, 307b. Đây không phải là trường hợp trong HE AAC, trong đó khoảng tần số bộ mã hóa lõi bị giới hạn bằng.

Fig.7b minh họa bộ giải mã phù hợp với bộ mã hóa trên Fig.1a; phép biểu diễn được mã hóa thứ nhất 107 được nhập 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. Thêm nữa, phép biểu diễn được mã hóa thứ hai 109 được nhập 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 khoảng khôi phục hoặc băng 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 vạch phổ 117 được nhập vào bộ chuyển đổi phổ - 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 bộ mã hóa trên Fig.1a. Tín hiệu đầu vào âm thanh 99 được nhập vào giàn lọc phân tích 220 tương ứng với bộ chuyển đổi phổ theo thời gian 100 trên Fig.1a. Sau đó, hoạt động tạo hình nhiễu âm theo thời gian được biểu diễn trong khối TNS 222. Do đó, việc nhập vào bộ phân tích phổ 102 trên Fig.1a tương ứng với màn chắn âm khối 226 trên 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 được minh họa trên Fig.2b, khối 222 được áp dụng. Các tín hiệu kênh đôi hoặc các tín hiệu đa kênh,

sự lập mã kênh ghép 228 có thể được thực hiện bổ sung, để bộ mã hóa miền phổ 106 trên Fig.1a có thể gồm có khối lập mã kênh ghép 228. Ngoài ra, bộ mã hóa entropi 232 để thực hiện việc nén dữ liệu không tổn hao được đề xuất mà là phần của bộ mã hóa miền phổ 106 trên Fig.1a.

Bộ phân tích phổ/màn chắn âm 226 phân tách đầu ra của khối TNS 222 thành băng 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 trên Fig.1a. Khối 224 được chỉ định là mã hóa trích xuất theo tham số IGF tương ứng với bộ mã hóa theo tham số 104 trên Fig.1a và bộ trộn kênh dòng bit 230 tương ứng với bộ trộn kênh dòng bit 108 trên 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 tác dụng biến đổi cosin rời rạc được biến đổi như công cụ phân tích tần số.

Bộ phân tích phổ 226 áp dụng tốt hơn là 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 tâm thính học. Giai đoạn ước lượng màn chắn âm có thể được thực hiện trong nhiều cách thức khác nhau và tốt hơn được thực hiện tương tự trong chức năng của nó để giai đoạn ước lượng vết hình sin được sử dụng trong hình sin hoặc nhiễu âm mô hình hóa cho việc giải mã âm thanh/tiếng nói tài liệu tham khảo [8, 9] hoặc bộ mã hóa âm thanh dựa trên mô hình HILN được miêu tả trong tài liệu tham khảo [10]. Tốt hơn là, việc thực hiện được sử dụng mà dễ dàng thực hiện 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 khoảng nguồn và khoảng mục tiêu. Khoảng mục tiêu được biểu diễn bởi phổ từ khoảng nguồn. Phép đo sự giống nhau giữa khoảng nguồn và khoảng mục tiêu được thực hiện sử dụng phương pháp tương quan chéo. Vùng mục tiêu được chia thành $nTar$ các ô tần số không chồng lấp. Với mỗi ô trong vùng mục tiêu, các ô nguồn $nSrc$ được tạo từ tần số bắt đầu được cố định. Các ô nguồn này chồng lấp với hệ số 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 mục tiêu tại các độ trễ khác nhau để tìm ra ô nguồn phù hợp nhất với ô mục tiêu. Số ô phù hợp nhất được lưu trữ trong $tileNum[idx_tar]$, độ trễ tại nơi tương quan nhất với mục tiêu 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 âm cao, ô nguồn cần được làm nhân với -1 trước quy trình điền đầy ô tại bộ giải mã. Môđun IGF cũng lưu ý không ghi đè các thành phần âm trong phổ vì các thành phần âm đã bảo toàn sử dụng màn chắn âm. Tham số năng lượng theo băng được sử dụng để lưu trữ năng lượng của vùng mục tiêu cho phép chúng ta khôi phục phổ chính xác.

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

Bên cạnh đó, kỹ thuật làm ổn định chọn ô được đề cập nhằm loại bỏ giả âm miền tần số như nhiễu âm âm rung và nhiễu âm âm nhạc.

Trong trường hợp của cặp kênh âm lập thể việc xử lý âm lập thể ghép bổ sung được áp dụng. Điều này là cần thiết, vì với khoảng cách nhất định tín hiệu có thể là nguồn âm thanh xoay chuyển tương quan cao. Trong trường hợp vùng nguồn được chọn cho vùng cụ thể không tương quan tốt này, mặc dù năng lượng phù hợp cho vùng đích, hình ảnh không gian có thể tồi tệ do các vùng nguồn không tương quan. Bộ mã hóa phân tích từng băng năng lượng vùng đích, thực hiện đặc trưng tương quan chéo của các giá trị phổ và nếu ngưỡng nhất định bị vượt quá, thiết lập dấu hiệu ghép cho băng năng lượng này. Trong bộ giải mã các băng năng lượng kênh phải và trái được xử

lý riêng biệt nếu dấu hiệu âm lập thể ghép này không được thiết lập. Trong trường hợp dấu hiệu âm lập thể ghép được thiết lập, cả năng lượng và việc ráp nối đều được thực hiện trong miền âm lập thể ghép. Thông tin âm lập thể ghép cho các vùng IGF được ký hiệu tương tự thông tin âm lập thể ghép cho việc mã hóa lõi, 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 thanh lập thể ghép cho các băng nơi mà âm thanh lập thể ghép là hoạt động, vì thế nên không cần việc 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 ghép → 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 được tín hiệu hóa từ bên tới giữa:

$$midTile1[k] = midTile[k] - prediction\ Coeff \cdot sideTile[k]$$

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

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

Quy trình này đảm bảo rằng từ các ô được sử dụng để tái tạo những vùng đích tương ứng cao và vùng đích xoay chuyển, kênh trái và phải thu được vẫn biểu diễn nguồn âm thanh tương quan và xoay chuyển thậm chí nếu 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 ghép được truyền mà biểu thị L/R hoặc M/S như là ví dụ cho việc mã hóa âm lập thể ghép 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 ghép cho các băng lỗi. Thứ hai, tín hiệu lỗi được lưu trữ trong cả phép biểu diễn L/R và M/S. Với điền đầy ô IGF, phép biểu diễn ô nguồn được chọn để phù hợp với phép biểu diễn ô mục tiêu như được biểu thị bởi thông tin âm lập thể ghép cho các băng IGF.

Tạo hình nhiễu âm theo thời gian (Temporal Noise Shaping - TNS) là kỹ thuật tiêu chuẩn và là một phần của AAC [11 – 13]. TNS có thể được coi như 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 ẩn 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 tập hợp các hệ số dự đoán sử dụng "sự dự đoán phía trước" trong miền biến đổi, ví dụ MDCT. Các hệ số sau đó được sử dụng để làm phẳng đường bao theo thời gian của các tín hiệu. Vì lượng tử hóa tác động tới phổ đã lọc TNS, nhiễu âm lượng tử hóa cũng được làm phẳng theo thời gian. Bằng việc áp dụng việc lọc TNS nghịch đảo trên phía bộ giải mã, nhiễu âm lượng tử hóa được tạo hình theo đường bao theo thời gian của bộ lọc TNS và vì thế nhiễu âm lượng tử hóa trở thành được chắn bởi phần chuyển tiếp.

IGF được dựa trên biểu diễn MDCT. Để việc 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 tiếng vang trước và tiếng vang sau có thể nghe được xuất hiện trong băng phổ IGF do việc điền đầy ô. Fig.7c cho thấy hiệu quả tiếng vang trước 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 tiếng vang trước 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 yêu cầu được tính toán và áp dụng sử dụng phổ đầy đủ trên phía bộ mã hóa như thường lệ. Các tần số bắt đầu và tần số dừng TNS/TTS không bị ảnh hưởng bởi tần số bắt đầu IGF $f_{IGFstart}$ của công cụ IGF. Khi so sánh với TNS kế thừa, tần số dừng TTS được tăng đến tần số dừng của công cụ IGF, mà cao hơn $f_{IGFstart}$. Trên 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 tiếng vang trước đã thể hiện được làm giảm. Bên cạnh đó, nó vẫn tạo hình nhiều âm lượng tử hóa trong tín hiệu dưới $f_{IGFstart}$ như thường lệ với TNS.

Trong các bộ giải mã kế thừa, việc ráp nối phổ trên tín hiệu âm thanh làm sai độ tương quan phổ tại đường biên ráp nối và do đó làm suy giảm đường bao thời gian của tín hiệu âm thanh bằng việc đưa vào sự phân tán. Vì thế, lợi ích khác của việc thực hiện điền đầy ô IGF trên tín hiệu dư là, sau khi áp dụng bộ lọc tạo hình, các đường biên ô được tương quan đều, thu được sự tái tạo theo thời gian của tín hiệu âm thanh trung thực hơn.

Trong bộ mã hóa theo sáng chế, phổ có trải qua việc điền đầy TNS/TTS, xử lý màn chắn âm và ước lượng tham số IGF là rộng của tín hiệu bất kỳ ở trên tần số bắt đầu IGF ngoại trừ các thành phần âm. Phổ rải rác này bây giờ được mã hóa bằng bộ mã hóa lỗi sử dụng các nguyên lý của việc mã hóa số học và mã hóa dự báo. Các thành phần được mã hóa này cùng với các bit tín hiệu dạng 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 được nhập vào bộ giải ghép/bộ giải mã mà được nối, liên quan đến Fig.1b, tới các khối 112 và 114. Bộ giải ghép dòng bit phân tách tín hiệu âm thanh đầu vào thành phép biểu diễn được mã hóa thứ nhất 107 trên Fig.1b và phép biểu diễn được mã hóa thứ hai 109 trên Fig.1b. Phép biểu diễn

được mã hóa thứ nhất có tập hợp thứ nhất của các phần phổ thứ nhất được nhập vào khối giải mã kênh ghép 204 tương ứng với bộ giải mã miền phổ 112 trên Fig.1b. Phép biểu diễn được mã hóa thứ hai được nhập vào bộ giải mã theo tham số 114 không được minh họa trên Fig.2a sau đó đầu vào vào khối IGF 202 tương ứng với bộ tái tạo tần số 116 trên Fig.1b. Tập hợp thứ nhất của các phần phổ thứ nhất yêu cầu sự tái tạo tần số được nhập vào khối IGF 202 qua vạch phổ 203. Ngoài ra, sau sự giải mã kênh ghép 204 sự giải mã lỗi cụ thể được áp dụng trong khối màn chắn âm 206 để đầu ra của màn chắn âm 206 tương ứng với đầu ra của bộ giải mã miền phổ 112. Sau đó, sự tổng hợp bởi bộ tổ hợp 208 được thực hiện, tức là, sự dựng khung trong đó đầu ra của bộ tổ hợp 208 hiện thời có phổ khoảng đầy, nhưng vẫn có 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 qua vạch phổ 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 lỗi 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 mà là tần số khoảng đầy đủ được định rõ 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 băng thừa số định tỉ lệ SCB tại đó có bảy thừa số định tỉ lệ SCB1 tới SCB7 trong ví dụ minh họa trên Fig.3a. Các băng thừa số định tỉ lệ có thể là băng định tỉ lệ AAC mà được định rõ trong 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 từ phần bắt đầu phổ, tức là, tại các tần số thấp, nhưng bắt đầu hoạt động IGF tại tần số bắt đầu IGF được minh họa tại 309. Vì thế, băng tần lỗi 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 các 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 lỗi 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 đến 0 hoặc là được đặt đến 0 trước khi lượng tử hóa hoặc sau khi lượng tử hóa. Dù thế nào đi nữa, bộ mã hóa lỗi hoạt động trong khoảng đầy đủ, tức là, nếu khi phổ được minh họa, tức là, bộ giải mã lỗi không nhất thiết phải nhận biết 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 định rõ bằng việc mã hóa theo vạch phổ của các vạch phổ phổ như các vạch phổ MDCT, trong khi độ phân giải thứ hai hoặc độ phân giải thấp được định rõ bằng, chẳng hạn như, cách tính toán chỉ một giá trị phổ đơn trên băng thừa số định tỉ lệ, tại đó băng thừa số định tỉ lệ bao trùm các vạch phổ tần số khác nhau. Vì thế, độ phân giải thấp thứ hai là, về khía cạnh để nó là độ phân giải phổ, thấp hơn nhiều độ phân giải thứ nhất hoặc cao được định rõ bởi việc mã hóa theo vạch phổ được ứng dụng bởi bộ mã hóa lỗi như AAC hoặc USAC.

Đối với thừa số định tỉ lệ hoặc việc tính toán năng lượng, tình huống này được minh họa trên Fig.3b. Do thực tế 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 băng, bộ mã hóa lỗi tính toán thừa số định tỉ lệ cho mỗi băng 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 $f_{IGFstop}$ nhỏ hơn hoặc bằng một nửa tần số lấy mẫu, tức là $f_s/2$. Vì thế, các phần âm được mã hóa 302, 304, 305, 306, 307 trên Fig.3a và trong phương án cùng với các thừa số định tỉ 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 thừa số định tỉ lệ 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 băng lỗi, tức là, thấp hơn về tần số so với tần số bắt đầu IGF, tức là, trong băng thừa số định tỉ lệ SB1 tới SB3 có thể được áp dụng bổ sung. Trong việc điền đầy âm, tồn tại một số vạch phổ liền kề mà được lượng tử hóa về 0. Trên phía bộ giải mã, các giá trị phổ đã lượng tử hóa tới 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 miêu tả ở 308 trên Fig.3b. Năng lượng điền đầy nhiều

âm, có thể có được các giới hạn tuyệt đối hoặc các giới hạn liên quan cụ thể với thừa số định tỉ lệ như trong USAC tương ứng với năng lượng của bộ giá trị phổ được lượng tử hóa tới 0. Các vạch phổ điền đầy nhiều âm này có thể cũng được xem xét là bộ thứ ba của phần phổ thứ ba mà được tái tạo bằng việc tổng hợp điền đầy nhiều âm truyền thẳng 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 khoảng nguồn và thông tin năng lượng E_1, E_2, E_3, E_4 .

Tốt hơn là, các băng, cho thông tin năng lượng được tính toán trùng với băng thừa số định tỉ lệ. Trong phương án khác, việc tạo nhóm giá trị thông tin năng lượng được áp dụng sao cho, ví dụ, cho các băng thừa số định tỉ lệ 4 và 5, chỉ giá trị thông tin năng lượng đơn được truyền, nhưng thậm chí trong phương án này, các đường biên của các băng khôi phục được nhóm trùng với các đường biên của băng thừa số định tỉ lệ. Nếu việc phân tách băng khác nhau được áp dụng, sau đó 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 trên Fig.1a là bộ mã hóa dẫn theo tâm thính học như minh họa trên Fig.4a. Điền hình, như ví dụ được minh họa trong tiêu chuẩn MPEG2/4 AAC hoặc tiêu chuẩn MPEG1/2, tiêu chuẩn lớp 3, thành tín hiệu âm thanh được mã hóa sau khi được biến đổi thành khoảng phổ (401 trên Fig.4a) được chuyển đến bộ tính toán thừa số định tỉ lệ 400. Bộ tính toán thừa số định tỉ lệ được điều khiển bởi mô hình âm thính học nhận bổ sung tín hiệu âm thanh được lượng tử hóa hoặc nhận, như trong tiêu chuẩn MPEG1/2 Layer 3 hoặc MPEG AAC, phép biểu diễn phổ phức tạp của tín hiệu âm thanh. Mô hình tâm thính học tính toán, cho từng băng thừa số định tỉ lệ, thừa số định tỉ lệ biểu diễn ngưỡng tâm thính học. Ngoài ra, sau đó các thừa số định tỉ lệ, bởi việc phối hợp của vòng lặp bên trong và bên ngoài đã biết hoặc bởi bất kỳ quy trình mã hóa thích hợp được điều chỉnh sao cho các điều kiện tốc độ bit nhất định được đáp ứng. Sau đó, giá trị phổ được lượng tử hóa trên một mặt và các thừa số định tỉ lệ được tính toán trên mặt khác được nhập vào bộ xử lý được lượng tử hóa 404. Trong hoạt động bộ mã hóa âm thanh dễ dàng, các giá trị phổ được lượng tử hóa được gán trọng số bởi các thừa số định tỉ lệ, và các giá trị phổ được gán trọng số sau đó được nhập vào bộ lượng tử hóa cố định điền hình có chức năng nén

thành các khoảng biên độ phía trên. 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 entropi đặc trưng có sự mã hóa cụ thể và rất hiệu quả cho bộ các chỉ số lượng tử hóa 0 cho các giá trị tần số liên kề hoặc, cũng được gọi trong kỹ thuật, là "sự vận hành" của các giá trị 0.

Tuy nhiên, trong bộ mã hóa âm thanh trên Fig.1a, bộ xử lý lượng tử hóa đặc trưng nhận thông tin về các phần phổ thứ hai từ bộ phân tích phổ. Vì thế, 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 là 0 hoặc có phép biểu diễn được xác nhận bởi bộ mã hóa hóa hoặc bộ giải mã như 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 được nhập vào tập hợp khối 0 410. Sau đó, các phần phổ thứ hai đã sẵn sàng được thiết lập về 0 trước khi gán trọng số bởi các thừa số định tỉ lệ trong khối 412 được thực hiện. Trong việc thực hiện bổ sung, khối 410 không được đề xuất, nhưng thiết lập về kết hợp 0 được thực hiện trong khối 418 sau khối gán trọng số 412. Trong việc thực hiện khác, thiết lập về hoạt động 0 có thể cũng được thực hiện trong thiết lập về khối 0 422 sau sự lượng tử hóa trong khối bộ lượng tử hóa 420. Trong việc thực hiện này, các 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 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 được đưa vào bộ mã hóa entropi như 232 trên Fig.2b mà có thể là bộ mã hóa Huffman hoặc bộ mã hóa số học như, ví dụ, được định rõ trong tiêu chuẩn USAC.

Thiết lập về các khối 0 410, 418, 422, mà được đề xuất thay thế cho nhau hoặc song song được điều khiển bởi bộ phân tích phổ 424. Bộ phân tích phổ ưu tiên bao gồm bất kỳ thực hiện của các bộ phát hiện âm điệu được biết đến rộng rãi hoặc bao gồm bất kỳ loại bộ phát hiện khác nhau đang 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 nhiễu âm, bộ phát hiện tiếng nói hay bất

kỳ bộ phát hiện 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 trên Fig.1a như, ví dụ, được thực hiện 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 sự chuyển tiếp 504. Khi bộ phát hiện sự chuyển tiếp 504 phát hiện sự chuyển tiếp, sau đó chuyển sang từ các cửa sổ dài đến các cửa sổ ngắn được phát tín hiệu đến bộ tạo cửa sổ. Bộ tạo cửa sổ 502 sau đó tính toán, cho các khối chồng lấp, các khung cửa sổ, nơi từng khung cửa sổ đặc trưng có hai N giá trị như các giá trị 2048. 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 sự lấy thập phân sao cho sự lấy thập phân/biến đổi được kết hợp được thực hiện để thu được khung phổ với N giá trị như các giá trị phổ MDCT. Do đó, đối với hoạt động cửa sổ dài, khung tại đầu vào của khối 506 bao gồm hai N giá trị như các giá trị 2048 và khung phổ sau đó có các giá trị 1024. 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 từng khối phổ có 1/8 giá trị phổ khi so sánh với khối dài. Do đó, sự 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 tới Fig.5b minh họa việc thực hiện cụ thể của bộ tái tạo tần số 116 và bộ chuyển đổi phổ-thời gian 118 trên Fig.1b, hoặc của hoạt động kết hợp của các khối 208, 212 trên Fig.2a. Trên Fig.5b, bảng khôi phục cụ thể được xem xét như bảng thừa số định tỉ lệ 6 trên Fig.3a. Phần phổ thứ nhất trong bảng khôi phục này, tức là, phần phổ thứ nhất 306 trên Fig.3a được nhập vào khối bộ dựng/bộ điều chỉnh khung 510. Ngoài ra, phần phổ thứ hai được khôi phục cho bảng thừa số được định tỷ lệ cũng được nhập vào bộ dựng/bộ điều chỉnh 510. Ngoài ra, thông tin năng lượng như E_3 trên Fig.3b cho bảng thừa số định tỉ lệ cũng được nhập vào khối 510. Phần phổ thứ hai được khôi phục trong bảng khôi phục sẵn sàng được tạo bởi sự điền đầy ô tần số sử dụng khoảng nguồn và bảng khôi phục sau khi đáp lại khoảng mục tiêu. 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 như, ví dụ, được thu tại đầu ra của bộ tổ

hợp 208 trên Fig.2a. Sau đó, trong khối 512, sự biến đổi/ nội suy khối nghịch đảo được thực hiện để thu được các giá trị miền thời gian 248 cho ví dụ 124 các giá trị phổ tại khối đầu vào 512. Sau đó, sự hoạt động của sổ tổng hợp được thực hiện trong khối 514 mà lại được điều khiển bởi sự chỉ thị của 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, N giá trị miền thời gian 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 được minh họa tại 301 trên Fig.3a, hoạt động điền đầy nhiều 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 băng khôi phục đã dự tính trùng với băng thừa số định tỉ lệ 6 trên Fig.3a. Sau đó, các giá trị phổ điền đầy nhiều âm có thể cũng được nhập vào bộ điều chỉnh/bộ xây dựng khung 510 và việc điều chỉnh các giá trị phổ điền đầy nhiều âm có thể cũng được áp dụng trong khối này hoặc giá trị phổ điền đầy nhiều âm có thể được điều chỉnh sử dụng năng lượng điền đầy nhiều 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 băng cao trên tần số bắt đầu IGF mà còn cũng được áp dụng trong cả băng thấp. Hơn nữa, việc điền đầy ô tần số không có nhiều 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, tìm ra rằng việc mã hóa âm thanh có chất lượng cao và hiệu quả cao có thể thu được khi hoạt động điền đầy nhiều âm được giới hạn tới 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 minh họa trên Fig.3a.

Tốt hơn là, các ô mục tiêu (target tiles - TT) (có tần số lớn hơn tần số bắt đầu IGF) bị giới hạn tới các đường biên băng thừa số định tỉ lệ của bộ mã hóa tỉ lệ đầy đủ. Các ô nguồn (Source tiles - 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 đường biên băng thừa số định tỉ lệ.

Kích thước của ST nên tương ứng với kích thước của TT đã kết hợp. Điều này được miêu tả 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 (như $4 + 6$). Sau đó, tất cả ST có thể mà tương quan với TT[0], cũng có độ dài 10 ngăn. Ô mục tiêu thứ hai TT[1] liên kế với TT[0] có độ dài của 15 ngăn I (SCB có độ dài $7+8$). Sau đó, ST để mà có độ dài của 15 ngăn chứ không phải 10 ngăn như đối với TT[0].

Nếu trường hợp phát sinh là không thể tìm thấy TT cho ST với độ dài của ô mục tiêu (khi, ví dụ, độ dài của TT lớn hơn khoảng nguồn hiện có), sau đó sự tương quan không được tính toán và khoảng nguồn được sao chép một số lần vào TT (việc sao chép được thực hiện một lần sau khi việc khác để sao cho vạch phổ tần số cho tần số thấp nhất của việc sao chép thứ hai theo sau ngay lập tức - theo tần số - đường tần số cho tần số cao của việc sao chép thứ nhất), đến khi ô mục tiêu TT được điền đầy hoàn toàn.

Tiếp đó, tham chiếu đến Fig.5c minh họa phương án được ưu tiên khác của bộ tái tạo tần số 116 trên Fig.1b hoặc khối IGF 202 trên Fig.2a. Khối 522 là bộ tạo ô tần số nhận, không chỉ bằng mục tiêu ID mà còn nhận thêm bằng nguồn ID. Ví dụ, điều này được xác định trên bộ mã hóa bên mà định tỉ lệ bằng thừa số trên Fig.3a rất thích hợp cho bằng thừa số định tỉ lệ khôi phục. Do đó, bằng nguồn ID sẽ là 2 và bằng đích ID sẽ là 7. Dựa trên thông tin này, bộ tạo ô tần số 522 áp dụng bản sao hoặc sự hoạt động điền đầy ô sóng hài hoặc bất kỳ sự hoạt động điền đầy ô khác để tạo phần thứ hai thô của các thành phần phổ 523. Phần thứ hai thô của các thành phần phổ có độ phân giải tần số đồng nhất với độ phân giải tần số được gồm có trong tập hợp thứ nhất của các phần phổ thứ nhất.

Sau đó, phần phổ thứ nhất của bằng khôi phục như 307 trên Fig.3a được nhập vào bộ dựng khung 524 và phần thứ hai thô 523 cũng được nhập vào bộ dựng khung 524. Sau đó, khung được khôi phục được điều chỉnh bởi bộ điều chỉnh 526 sử dụng các thừa số khuếch đại cho bằng khôi phục được tính toán bởi bộ tính toán thừa số khuếch đại 528. Tuy nhiên, điều 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. Với mục đích này, bộ tính toán thừa số khuếch đại 528 phân tích bằng nguồn hoặc phần thứ hai thô 523 và phân tích bổ sung phần phổ thứ

nhất trong băng khôi phục để tìm thừa số khuếch đại chính xác 527 để năng lượng của đầu ra khung được điều chỉnh bởi bộ điều chỉnh 526 có năng lượng E_4 khi băng thừa số định tỉ lệ được dự tính trước.

Trong phạm vi này, rất quan trọng để ước tính độ chính xác khôi phục tần số cao của sáng chế so với HE-AAC. Điều này được giải thích với băng thừa số định tỉ lệ 7 trên Fig.3a. Điều này được giả định rằng bộ mã hóa đã có trước trong tình trạng kỹ thuật như miêu 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 băng khôi phục như băng thừa số định tỉ 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 huyền 307 sẽ được khôi phục bởi bộ giải mã trong tình trạng kỹ thuật trên Fig.13b sẽ ở giữa băng 7 tại tần số được biểu thị bởi tần số khôi phục 390. Do đó, sáng chế tránh được lỗi tần số 391 mà được đưa ra bởi bộ giải mã trong tình trạng kỹ thuật trên Fig.13d.

Trong việc thực hiện, bộ phân tích phổ cũng được thực hiện để tính toán những điểm tương đồng giữa phần phổ thứ nhất và phần phổ thứ hai và để xác định dựa trên điểm tương đồng đã tính toán, để phần phổ thứ hai trong khoảng khôi phục phần phổ thứ nhất phù hợp với phần phổ thứ hai như là có thể xảy ra. Sau đó, trong thực hiện khoảng nguồn/khoảng đích có thể thay đổi, bộ mã hóa theo tham số sẽ được đưa bổ sung vào trong phép biểu diễn được mã hóa thứ hai thông tin phù hợp biểu thị cho từng khoảng đích khoảng nguồn phù hợp. Trên phía bộ giải mã, thông tin này sau đó sẽ được sử dụng bằng bộ tạo ô tần số 522 trên Fig.5c minh họa sự tạo ra phần thứ hai tần số 523 dựa trên băng nguồn ID và băng mục tiêu ID.

Hơn nữa, như 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 ít nhất là một phần tư của tần số lấy mẫu hoặc thường cao hơn.

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

Thêm vào đó, 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 thúc 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 những phần phổ nữa 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 được 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 là 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 để 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 thừa số định tỉ lệ cho bằng thừa số định tỉ lệ tồn tại, mà được tạo ra và được truyền không kể có hay không tất cả giá trị phổ trong băng thừa số định tỉ lệ này được thiết lập tới không hoặc không như đã thảo luận phạm vi của các Fig.3a và Fig.3b.

Do đó, sáng chế có lợi đối với kỹ thuật tham số khác để làm tăng hiệu quả nén, ví dụ, việc thay thế nhiều âm và điền đầy nhiều âm (các kỹ thuật này dành riêng cho phép 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 phép biểu diễn tham số hiệu quả của lượng tín hiệu tùy ý bởi việc điền đầy khoảng trống phổ không hạn chế phân chia ưu tiên được cố định trong băng thấp (LF) và băng cao (HF).

Các phương án của hệ thống theo sáng chế cải thiện cách tiếp cận 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 gồm có:

- mã hóa lõi băng đầy đủ
- điền đầy khoảng trống thông minh (điền đầy ô hoặc điền đầy nhiều âm)
- các phần âm điệu thừa thớt trong lõi được chọn bởi màn chắn âm
- mã hóa cặp âm thanh lập thể ghép cho băng đầy đủ, chứa 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, như AAC là ví dụ, sử dụng MDCT như việc 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 được đề xuất có tên là điền đầy khoảng trống thông minh (IGF). Fig.2 thể hiện sơ đồ khối của hệ thống được đề xuất trên phía bộ mã hóa và Fig.2 thể hiện hệ thống trên phía bộ giải mã.

Fig.6a minh họa thiết bị để 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 tập hợp 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 được minh họa ở 603, các giá trị phổ trong phần phổ thứ nhất và phần phổ thứ hai là các giá trị dư dự báo phổ. Để biến đổi các giá trị dư 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 dư 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 dư dự báo phổ 603 được nhập vào bộ tạo ô tần số 612 để tạo các giá trị phổ thô cho băng khôi phục hoặc phần tần số thứ hai nhất định và dữ liệu thô này hiện tại có độ phân giải giống phép biểu diễn phổ thứ nhất độ phân giải cao được nhập vào bộ tạo hình phổ 614. Bộ tạo hình phổ hiện thời tạo hình phổ sử dụng thiết bị dự báo đường bao được truyền trong dòng bit và dữ liệu được tạo hình theo phổ sau đó được áp dụng vào bộ lọc dự báo phổ 616

cuối cùng tạo khung giá trị phổ đầy sử dụng thông tin bộ lọc 607 được truyền từ bộ mã hóa tới bộ giải mã qua dòng bit.

Trên Fig.6b, giả định rằng, trên phía bộ mã hóa, việc tính toán thông tin bộ lọc được truyền thông qua dòng bit và sử dụng thông qua vạch phổ 607 được thực hiện tiếp theo tính toán thông tin đường bao. Vì thế nên, nói cách khác, bộ mã hóa phù hợp với bộ giải mã trên 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 những thực hiện nhất định, tại đó thông tin đường bao được tính toán trước khi thực hiện lọc TNS hoặc TTS trên phía bộ mã hóa. Sau đó, bộ lọc dự đoán phổ 622 được áp dụng trước khi thực hiện tạo hình phổ trong khối 624. Do đó, nói cách khác, các giá trị phổ (đầy) được tạo trước khi 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 ban đầu được nhập vào khối MDCT phức tạp 702. Sau đó, việc tính toán bộ lọc TTS việc sự 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. Sau đó, tập hợp thứ nhất của phần phổ thứ nhất 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 chỉ thị tại $X(k)$ trên Fig.7a và tất cả dữ liệu này được đưa tới bộ trộn dòng bit 710.

Trên phía bộ giải mã, dữ liệu được mã hóa được nhập vào bộ tách kênh 720 để một mặt phân tách thông tin phụ IGF, mặt khác phân tách thông tin phụ TTS và phép 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 băng khôi phục. Sau đó, khung băng đầy được điền ô và được thu hoàn toàn, hoạt động TTS nghịch đảo 728 được thực hiện và, trên bộ giải mã bên, hoạt động MDCT phức tạp nghịch đảo cuối cùng được thực hiện trong khối 730. Do đó, việc sử dụng thông tin bộ lọc TNS phức tạp cho phép, khi được ứng dụng không chỉ vào trong

bằng lỗi hoặc trong các băng phân tách như được ứng dụng trên các đường viền lỗi/ô hoặc các đường viền ô/ô tự động tạo ra việc xử lý đường viền ô, mà, với mục đích đưa lại sự tương phổ giữa các ô. Tương quan phổ trên đường biên ô không thu được bằng cách chỉ 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 so sánh của tín hiệu gốc (bảng trái) và tín hiệu được mở rộng không có TTS. Có thể thấy rằng có các thành phần lạ mạnh được minh họa bởi các phần được mở rộng trong khoảng tần số phía trên được minh họa tại 750. Đây là, tuy nhiên, không xảy ra trên Fig.7e khi phần phổ giống tại 750 được so sánh với các thành phần lạ liên quan 750 trên 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 hạn chế 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 nghiêm ngặt. 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 băng rộng bằng việc sử dụng các phương pháp mở rộng băng thông (BWE) như sao chép băng 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 ("ráp nối") 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 ráp nối phổ. Điển hình, vùng HF gồm nhiều việc ráp nối liên kế và từng việc ráp nối này được bắt nguồn từ các vùng thông dải (BP) của phổ LF thấp hơn tần số giao nhau đã có. Hệ thống thuộc tình trạng kỹ thuật thực hiện hiệu quả với việc ráp nối

trong biểu diễn giàn lọc bởi việc sao chép bộ hệ số băng con liền kề từ vùng nguồn tới vùng mục tiêu.

Nếu hệ thống BWE được thực hiện trong giàn 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. Diễn hình, độ kết hạt theo thời gian bị giới hạn bởi kích thước bước nhảy được sử dụng giữa các cửa sổ biến đổi liền kề. Điều này có thể dẫn tới tiếng vang trước hoặc tiếng vang sau không mong muốn trong khoảng phổ BWE.

Từ việc mã hóa âm thanh theo cảm giác, người ta biết rằng tạo hình đường bao thời gian của tín hiệu âm thanh có thể được lưu trữ lại bằng việc sử dụng kỹ thuật lọc phổ như tạo hình đường bao theo thời gian (Temporal Envelope Shaping - TNS) [14]. Tuy nhiên, bộ lọc TNS được biết tới trong tình trạng kỹ thuật là bộ lọc giá trị thật trên phổ giá trị thật. Như bộ lọc giá trị thật trên phổ giá trị thật có thể bị suy yếu rất nhiều bởi giả âm sai số, đặc biệt nếu việc biến đổi thực tế bên dưới là biến đổi cosin rời rạc được biến đổi (Modified Discrete Cosine Transform - 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 thay đổi phức hợp (Complex Modified Discrete Cosine Transform - CMDCT). Bằng cách đó, giả âm sai số được ngăn chặn.

Tạo hình ô theo thời gian gồm có:

- 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 nhiễu âm theo thời gian (Temporal Noise Shaping - TNS) bằng dự báo tuyến tính theo hướng tần số, để sử dụng cách thức thay đổi trong phạm vi mở rộng băng thông.

Thêm nữa, thuật toán mở rộng băng thông theo sáng chế dựa trên điền đầy khoảng trống thông minh (Intelligent Gap Filling - IGF), nhưng sử dụng sự lấy mẫu quá, biến đổi giá trị phức hợp (CMDCT), ngược với cấu hình tiêu chuẩn IGF dựa vào biểu diễn MDCT lấy mẫu tới hạn 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 phổ giá trị phức hợp.

Mặc dù cách tiếp cận mới được miêu tả trong phạm vi IGF, quy trình theo sáng chế có thể sử dụng trong tổ hợp với bất kỳ phương pháp BWE 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 như tạo hình nhiễu âm theo thời gian, mà đúng hơn như kỹ thuật tạo hình ô theo thời gian (temporal tile shaping - TTS). Việc đổi tên là hợp lý, bởi thực tế là thành phần tín hiệu điền đầy ô được tạo hình theo thời gian bằng TTS 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 giá trị phức hợp
- có được thông tin phụ cho BWE và loại bỏ thông tin phổ mà được sao chép lại bởi bộ giải mã
- áp dụng lượng tử hóa sử dụng môđun tâm thính học (psycho acoustic module - PAM)
- lưu giữ/truyền dữ liệu, chỉ các hệ số MDCT giá trị thực được truyền

Fig.7b thể hiện bộ giải mã tương ứng. Nó chủ yếu ngược với các bước 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 khối bộ giải mã trễ) và tổ hợp các hệ số MDCT và MDST thành các hệ số CMDCT giá trị phức hợp
- thực hiện điền đầy ô với 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, ngoài ra, để tổng hợp TTS và xử lý sau 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 này chứa các phần chuyển tiếp, các tiếng vang trước và tiếng vang sau có thể nghe được xuất hiện trong băng phổ IGF do việc điền đầy ô. Fig.7c thể hiện hiệu quả tiếng vang trước điển hình trước sự bắt đầu phần chuyển tiếp do IGF. Trên bảng trái trên 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 được thể hiện. Trong ví dụ, tần số bắt đầu IGF $f_{IGFstart}$ hoặc f_{split} giữa băng lõi và băng đã điền đầy ô được chọn là $f_s/4$. Trong bảng phải của Fig.7c, tiếng vang trước và tiếng vang sau rõ rệt bao quanh có thể thấy được xung quanh 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 chắn theo thời gian của nhận thức con người. Do đó, các hệ số dự báo TTS được yêu cầu được tính toán và áp dụng 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 TTS bắt đầu [...] hoặc TTS dừng [...] sẽ không có trong ô IGF, mà phải được căn chỉnh với các tần số $f_{IGF...}$ tương ứng. 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 $f_{IGFstart}$ thu được trong 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, lưu ý bổ sung đưa ra là năng lượng trong bộ giải mã IGF được điều chỉnh chính xác. Điều này đặc biệt

đúng nếu, trong xử lý TTS và IGF, 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 mục tiêu (như bộ lọc tạo hình không là bản sao chính xác của bộ lọc phẳng đã nêu) của một ô IGF. Trong trường hợp này, tỉ lệ lợi ích dự báo của hai bộ lọc TTS được áp dụng không bằng một nửa và do đó phải điều chỉnh năng lượng theo tỉ lệ này.

Trong biến thể thực hiện thay thế, đề xuất 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 quá trình lọc TTS và do đó là bước xử lý cuối cùng trước khi biến đổi tổng hợp. Do đó, bất chấp 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.

Trên phía bộ giải mã, các hệ số bộ lọc TTS được áp dụng trên phổ đầy đủ một lần nữa, 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 tiếng vang trước 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 ráp nối phổ trên tín hiệu âm thanh (ví dụ, SBR) làm sai lệch sự tương quan phổ tại đường biên ráp nối và theo đó làm suy yếu đường bao thời gian của tín hiệu âm thanh bởi sự phân tán được đưa vào. Vì thế, 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 đường biên ô được tương quan đều, thu được sự tái tạo theo thời gian trung thành hơn của tín hiệu âm thanh.

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

Thêm nữa, như được 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 gồm có bộ chuyển đổi thời gian-phổ như 702 để chuyển đổi tín hiệu âm thanh thành phép biểu diễn phổ. Phép biểu diễn phổ có thể là phép 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 định rõ 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 được minh họa ở 714 trên Fig.7a. Hơn nữa, bộ mã hóa âm thanh như bộ mã hóa âm thanh dẫn tâm thính học 704 được cung cấp. 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 gồm có 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 toán 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 từ Fig.8a tới Fig.8e. Hơn nữa, tham chiếu được thực hiện để mô tả các chi tiết tương ứng trong phạm vi của Fig.2a, Fig.2b (sự lập mã kênh ghép 228 và sự 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 gồm có 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 theo 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 các phần phổ thứ nhất và dữ liệu theo 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 khoảng nguồn và khoảng đích. Khoảng 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, khoảng 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. Thêm nữa, khoảng 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 được minh họa trong cột thứ ba trên Fig.8b. Cuối cùng, cả khoảng 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 trong đó hai kênh của tín hiệu kênh đôi được biểu diễn riêng rẽ. Sau đó, phép biểu diễn kênh đôi thứ nhất là phép biểu diễn ghép trong đó hai kênh của phép biểu diễn kênh đôi được biểu diễn ghép, tức là, trong đó việc xử lý khác hoặc biến đổi phép 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 ghép. 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 đồ thị dòng cho hoạt động được thực hiện bởi bộ giải mã âm thanh trên Fig.8a. Trong bước 812, bộ giải mã âm thanh 802 thực hiện việc giải mã khoảng nguồn. Khoảng nguồn có thể gồm có, đối với Fig.3a, các băng thừa số định tỉ lệ SCB1 tới SCB3. Thêm nữa, có thể là sự nhận dạng kênh đôi cho từng băng thừa số định tỉ lệ và băng thừa số định tỉ lệ 1, ví dụ, là trong phép biểu diễn thứ nhất (như L/R) và trong băng thừa số định tỉ 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 băng khác nhau. Tiếp đó, trong bước 814, bộ tái tạo tần số 806 được tạo cấu hình để lựa chọn khoảng nguồn cho việc tái tạo tần số. Trong bước 816,

bộ tái tạo tần số 806 sau đó kiểm tra phép biểu diễn của khoảng nguồn và trong khối 818, bộ tái tạo tần số 806 so sánh phép biểu diễn kênh đôi của khoảng nguồn với phép biểu diễn kênh đôi của khoảng mục tiêu. 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, sau đó dòng tín hiệu 824 được giữ và khối 822 tính toán phép biểu diễn kênh đôi khác từ khoảng nguồn và sử dụng phép biểu diễn kênh đôi khác được tính toán này cho sự tái tạo khoảng mục tiêu. Do đó, bộ giải mã trên 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 khoảng nguồn trong phép biểu diễn kênh đôi thứ nhất. Tất nhiên, việc bổ sung theo sáng chế cho phép tái tạo khoảng mục tiêu sử dụng khoảng nguồn có sự nhận dạng kênh đôi giống nhau. Và, thêm vào đó, sáng chế cho phép tái tạo khoảng mục tiêu 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 giữ hoặc truyền tới loa phát tương ứng cho tín hiệu kênh đôi.

Điều này được nhấn mạnh rằng hai kênh của phép biểu diễn kênh đôi có thể là kênh đôi âm lập thể 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 tới 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. Khoảng nguồn hoặc bộ giải mã lõi 830 có thể đáp lại 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 trên Fig.8a. Cụ thể, khối 832 là bộ biến đổi biểu diễn để biến đổi các phép biểu diễn khoảng nguồn trong từng băng riêng biệt sao cho, tại đầu ra của khối 832, bộ hoàn thiện của khoảng 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 khoảng nguồn hoàn chỉnh này có thể được lưu giữ trong bộ lưu 834 cho cả hai phép biểu diễn của khoảng nguồn.

Sau đó, khối 836 áp dụng việc tạo ra ô tần số sử dụng, như đầu vào, khoảng nguồn ID và sử dụng bổ sung như đầu vào kênh đôi ID cho khoảng mục tiêu. Dựa trên ID kênh đôi cho khoảng mục tiêu, bộ tạo ô tần số truy nhập bộ lưu trữ 834 và nhận bộ tạo kênh đôi của khoảng nguồn phù hợp với ID kênh đôi cho đầu vào khoảng đích vào trong bộ tạo ô tần số tại 835. Do đó, khi ID kênh đôi cho khoảng đích chỉ ra sự xử lý âm lập thể ghép, sau đó bộ tạo ô tần số 836 truy nhập bộ lưu trữ 834 để thu được phép biểu diễn âm lập thể của khoảng nguồn được chỉ ra bởi ID khoảng nguồn 833.

Bộ tạo ra ô tần số 836 thực hiện hoạt động này cho từng khoảng mục tiêu 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, các tham số điều chỉnh đường bao được yêu cầu 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 miêu tả. Khi đó, sự nhận dạng kênh đôi trong khoảng mục tiêu 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 mục tiêu này, sau đó bộ biến đổi theo tham số 840 biến đổi tham số đường bao thành phép biểu diễn kênh đôi đã yêu cầu. Khi, ví dụ, sự nhận dạng kênh đôi cho một băng biểu thị mã hóa âm lập thể ghép và khi các tham số cho khoảng mục tiêu này được truyền như tham số đường bao L/R, sau đó bộ biến đổi theo tham số tính toán các tham số đường bao âm lập thể ghép từ tham số đường bao L/R như đã miêu tả sao cho phép biểu diễn theo tham số chính xác được sử dụng để điều chỉnh đường bao phổ của khoảng mục tiêu.

Trong phương án ưu tiên khác các tham số đường bao được truyền như tham số âm lập thể ghép khi âm lập thể ghép được sử dụng trong băng mục tiêu.

Khi giả định rằng đầu vào vào bộ điều chỉnh đường bao 838 là bộ của các khoảng mục tiêu có các phép biểu diễn kênh đôi khác nhau, sau đó đầu ra của bộ điều chỉnh đường bao 838 là bộ của khoảng mục tiêu cũng trong các phép biểu diễn kênh đôi khác nhau. Khi, khoảng mục tiêu có phép biểu diễn được ghép như M/S, sau đó khoảng mục tiêu 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 mục tiêu 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 tiếp theo này có thể, ví dụ, là sự chuyển đổi tần số/thời gian hoặc là các xử lý được yêu cầu khác.

Tốt hơn là, các phần phổ thứ hai tương ứng với các băng tần, và việc nhận dạng kênh đôi được cung cấp như dãy tín hiệu tương ứng với bảng trên Fig.8b, tại đó một tín hiệu cho mỗi băng 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 trong sự 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 khoảng khôi phục bắt đầu với tần số bắt đầu IGF 309 trên Fig.3a có các nhận dạng kênh đôi cho các băng 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 nữa, việc nhận dạng băng nguồn và nhận dạng băng mục tiêu 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 khoảng nguồn tới khoảng mục tiêu. Khoảng nguồn có thể được sử dụng để tái tạo lại với tần số, khoảng mục tiêu 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 hai hoạt động điền đầy ô tần số tương tự với việc xử lý nhiều việc ráp nối 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 gồm có 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. Hơn nữa, 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, tức là, 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.

Thêm nữa, 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 rõ việc nhận dạng kênh đôi nhận dạng hoặc phép biểu diễn kênh đôi thứ nhất hoặc phép biểu diễn kênh đôi thứ hai.

Phụ thuộc vào kết quả của bộ phân tích kênh đôi, băng trong phép biểu diễn phổ thứ hai được tham số hóa sử dụng phép biểu diễn kênh đôi thứ nhất hoặc phép biểu diễn kênh đôi thứ hai, và được thực hiện bằng bộ mã hóa tham số 868. Phạm vi tần số lỗi, tức là băng tần dưới tần số bắt đầu 309 trên Fig.3a được mã hóa bởi bộ mã hóa lỗi 870. Kết quả của các khối 868 và 870 được nhập vào giao diện đầu ra 872. Như được chỉ ra, bộ phân tích kênh đôi đề xuất sự nhận dạng kênh đôi cho mỗi băng trên tần số bắt đầu IGF hoặc cho toàn bộ băng tần, và sự nhận dạng kênh đôi cũng được chuyển đến giao diện đầu ra 872, để dữ liệu này cũng được bao gồm trong tín hiệu được mã hóa 873 được đưa ra bởi giao diện đầu ra 872.

Ngoài ra, tốt hơn là bộ mã hóa âm thanh gồm có bộ biến đổi theo băng 862. Dựa trên quyết định của bộ phân tích kênh đôi 862, tín hiệu đầu ra của bộ chuyển đổi phổ theo thời gian 862 được biến đổi thành phép biểu diễn được chỉ ra bởi bộ phân tích kênh đôi, và cụ thể, bởi ID kênh đôi 835. Do đó, đầu ra của bộ biến đổi theo băng 862 là tập hợp các băng tần trong đó mỗi băng tần có thể là phép biểu diễn kênh đôi thứ nhất hoặc phép biểu diễn kênh đôi thứ hai khác. Khi sáng chế được áp dụng trong băng đầy đủ, tức là, khi khoảng nguồn và khoảng khôi phục đều được xử lý bởi bộ biến đổi theo băng, bộ phân tích phổ 860 có thể phân tích phép biểu diễn này. Ngoài ra, tuy nhiên, bộ phân tích phổ 860 có thể cũng phân tích đầu ra tín hiệu bởi bộ chuyển đổi phổ theo thời gian được biểu thị bởi vạch phổ điều khiển 861. Do đó, bộ phân tích phổ 860 có thể ứng dụng sự phân tích âm điệu được ưu tiên trên đầu ra của bộ biến đổi theo băng 862 hoặc đầu ra của bộ chuyển đổi phổ theo thời gian 860 trước khi được xử lý bởi bộ biến đổi theo băng 862. Ngoài ra, bộ phân tích phổ có thể áp dụng sự nhận dạng của phạm vi nguồn phù hợp nhất cho phạm vi đích nhất định trên kết quả của bộ chuyển đổi theo băng 862 hoặc trên kết quả của bộ chuyển đổi phổ theo thời gian 860.

Sau đó, tham chiếu tới các hình vẽ từ Fig.9a đến Fig.9d để minh họa việc tính toán được ưu tiên của các giá trị thông tin năng lượng được 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.

Bên trong eSBR [15] tín hiệu âm thanh được lấy mẫu giảm ít nhất bởi các hệ số của hai và phần tần số cao của phổ được đưa về không một cách hoàn chỉnh ngoài [1, 17]. Phần được xóa này được thay thế bởi kỹ thuật theo tham số, eSBR, trên 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 lập mã USAC [15] cung cấp sự khả thi để điền đầy các lỗ trống phổ (các vạch phổ phổ 0) với âm ngẫu nhiên như có các khuyết điểm sau: âm ngẫu nhiên có thể không bảo vệ được cấu trúc mịn theo thời gian của tín hiệu truyền tiếp và nó không thể bảo vệ được cấu trúc sóng hài của tín hiệu âm.

Khu vực nơi mà eSBR hoạt động trên phía bộ giải mã được xóa hoàn toàn bởi bộ mã hóa [1]. Do đó eSBR dễ xóa các vạch phổ âm trong vùng tần số cao hoặc cấu trúc sóng hài biến dạng của tín hiệu gốc. Như 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 ráp nối, 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 dạng đường bao phổ. Tình trạng kỹ thuật này không xử lý phổ bị xóa một phần và do độ phân giải thời gian cao hoặc có thể cần lượng bit tương đối lớn để truyền giá trị năng lượng thích hợp hoặc áp dụng việc lượng tử hóa thô cho các giá trị năng lượng.

Phương pháp IGF không cần biến đổi bổ sung vì nó sử dụng biến đổi MDCT kế thừa mà được tính toán như miêu 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 băng thừa số định tỉ 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 $f_{IGFstart}$ và $f_{IGFstop}$ là các tham số do người sử dụng cung cấp.

Các giá trị đã tính toán từ bước c) và d) được mã hóa không tiêu hao và truyền như thông tin phụ với dòng bit đến bộ giải mã.

Bộ giải mã nhận giá trị được 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 giá trị năng lượng đã truyền
- e) Điều chỉnh theo băng thừa số định tỉ lệ đường bao phổ
- f) Áp dụng TNS/TTS nếu được tín hiệu hóa

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

Trong [16, 4.6.2] sự phân bố của $\hat{x}$ trong các băng thừa số định tỉ lệ được miêu tả. Các băng thừa số định tỉ lệ là bộ của bộ hệ số và được biểu thị và kí hiệu trong bản mô tả với scb .

Giới hạn của từng scb_k với $k = 0, 1, 2, \dots, max_sfb$ được định rõ bởi dãy - swb_offset (16, 4.6.2), trong đó $swb_offset[k]$ và $swb_offset[k+1]-1$ định rõ chỉ số đầu tiên và cuối cùng cho vạch phổ hệ số phổ thấp nhất và cao nhất có trong scb_k . Biểu thị băng thừa số định tỉ 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 định rõ tần số bắt đầu IGF và tần số dừng IGF. Hai giá trị này ánh xạ với chỉ số băng thừa số định tỉ lệ phù hợp nhất $igfStartSfb$ và $igfStopSfb$. Cả hai được tạo tín hiệu trong dòng truyền dữ liệu tới bộ giải mã.

[16] miêu 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 thừa số định tỉ lệ được truyền tới bộ giải mã. Để khối ngắn tám cửa sổ ngắn với tám tập hợp khác nhau của hệ số phổ được tính toán. Để lưu tốc độ bit, các thừa số định tỉ 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 băng thừa số định tỉ 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:

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

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

Ta giả định rằng bộ mã hóa này quyết định nhóm các bộ thừa số định tỉ lệ num_window_group .

Chỉ rõ w phân vùng nhóm này của tập hợp $\{0,1,2,\dots,7\}$ mà là các chỉ số của tám cửa sổ ngắn. w_l chỉ rõ l -th là tập con của w , trong đó l chỉ rõ chỉ số của nhóm cửa sổ, $0 \leq l < \text{num_window_group}$.

Để tính toán khối ngắn người sử dụng định rõ tần số bắt đầu/dừng IGF được ánh xạ với băng thừa số định tỉ lệ thích hợp. Tuy nhiên, để đơn giản biểu thị cho các khối ngắn $k = \text{igfStartSfb}, 1 + \text{igfStartSfb}, 2 + \text{igfStartSfb}, \dots, \text{igfEndSfb}$.

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

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

Với việc lượng tử hóa:

$$\hat{E}_{k,l} = n\text{INT}(4\log_2(E_{k,l}))$$

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

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

Để $\hat{x}_r \in \mathbb{R}^N$ được biến đổi MDCT, phép biểu diễn phổ 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 giá trị thực của phần tương tự của tín hiệu âm thanh. Phép biểu diễn phổ MDST $\hat{x}_i$ có thể được tính toán chính xác hoặc được ước lượng từ $\hat{x}_r$. $\hat{c} := (\hat{x}_r, \hat{x}_i) \in \mathbb{C}^N$ chỉ rõ phép biểu diễn phổ phức hợp của tín hiệu âm thanh được tạo cửa sổ, có $\hat{x}_r$ là phần thực của nó và $\hat{x}_i$ là phần ảo của nó. Bộ mã hóa áp dụng tùy ý TNS trên $\hat{x}_r$ và $\hat{x}_i$.

Bây giờ năng lượng của tín hiệu gốc trong khoảng IGF có thể được đo với:

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

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

$$E_{tk} = \frac{1}{|\text{scb}_k|} \sum_{i \in \text{tr}_k} \hat{c}_i^2, E_{rk} = \frac{1}{|\text{scb}_k|} \sum_{i \in \text{tr}_k} \hat{x}_{ri}^2$$

Trong đó tr_k là bộ biểu thị khoảng ô nguồn được kết hợp, trong sự phụ thuộc của scb_k . Trong hai cách thức trên đây, thay vì bộ chỉ số scb_k , bộ $\overline{scb_k}$ (được xác định sau bộ chỉ số này) có thể được sử dụng để tạo tr_k để đạt được các giá trị chính xác hơn E_t và E_r .

Tính toán:

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

Nếu $E_{tk} > 0$, khác $f_k = 0$.

Với:

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

bây giờ bản ổn định hơn của E_k được tính toán, vì sự tính toán của E_k với giá trị MDCT duy nhất được làm giảm bởi sự thật là giá trị MDCT không tuân theo định lý của Parseval, do đó chúng không phản ánh thông tin năng lượng hoàn chỉnh của các giá trị phổ. $\hat{E}_k$ được tính toán như ở trên.

Như lưu ý gần đây, với các khối ngắn ta thừa nhận rằng bộ mã hóa quyết định để nhóm *num_window_group* các bộ thừa số định tỉ lệ. Theo nêu trên, w_l chỉ rõ l -th là tập con của w , trong đó l chỉ rõ chỉ số của nhóm cửa sổ, $0 \leq l < num_window_group$.

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 định rõ 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 $\hat{x}_i \in \mathbb{R}^N$ là tín hiệu âm thanh cửa sổ MDCT đã biến đổi của đoạn $2N$, tính toán:

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

Tính toán tương tự:

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

và tiếp tục với thừa số $f_{k,l}$:

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

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

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

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

Quy trình không chỉ sử dụng năng lượng băng khôi phục hoặc nhận được từ băng khôi phục phức hợp hoặc từ giá trị MDCT, mà còn sử dụng thông tin năng lượng từ khoảng nguồn cung cấp việc 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 băng khôi phục sử dụng thông tin về năng lượng của băng khôi phục và sử dụng thêm thông tin về năng lượng của khoảng nguồn để sử dụng cho việc khôi phục băng 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 băng khôi phục của phổ phức hợp của tín hiệu gốc, để tính toán thông tin năng lượng tiếp theo (E_{rk}) trên khoảng nguồn của phần giá trị thực của phổ phức hợp trong tín hiệu gốc được sử dụng cho việc khôi phục băng khôi phục, và trong đó bộ tính toán theo tham số được tạo cấu hình để tính toán thông tin năng lượng cho băng khôi phục sử dụng thông tin năng lượng (E_{ok}) và thông tin năng lượng tiếp theo (E_{rk}).

Thêm nữa, bộ tính toán theo tham số 1006 được tạo cấu hình để định rõ thông tin năng lượng thứ nhất (E_{ok}) trên băng thừa số định tỉ lệ đã khôi phục của phổ phức hợp của tín hiệu gốc, để xác định thông tin năng lượng thứ hai (E_{tk}) trên khoảng nguồn của phổ phức hợp của tín hiệu gốc được sử dụng để khôi phục băng thừa số định tỉ lệ được khôi phục, để xác định thông tin năng lượng thứ ba (E_{rk}) trên khoảng nguồn của phần giá trị thực của phổ phức hợp của tín hiệu gốc được sử dụng để khôi phục băng thừa số định tỉ lệ được khôi phục, để xác định thông tin trọng số dựa trên mối liên hệ giữa ít nhất hai trong số 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 cho băng khôi phục.

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

A)

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

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

B)

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

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

C)

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

$$E_k = \sqrt{f_k * E_{ok}}$$

D)

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

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

Tất cả các ví dụ này xác nhận sự thật là mặc dù chỉ các giá trị MDCT thực được xử lý trên phía bộ giải mã, nhưng tính toán thực tế là - do sự chồng chéo 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 tiếp theo 922, 923 của băng khôi phục 920 cho các giá trị tần số khác với phần phổ thứ nhất 921 có các tần số trong băng khôi phục 920 dựa vào các giá trị MDCT thực. Vì thế, thông tin năng lượng đã truyền tới bộ giải mã sẽ nhỏ hơn thông tin năng lượng E_{ok} trên băng 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) sẽ nhỏ hơn 1.

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

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

cho tất cả $k = igfStartSfb, 1 + igfStartSfb, 2 + igfStartSfb, \dots, igfEndSfb$.

Bộ giải mã khử lượng tử hóa giá trị MDCT đã truyền tới $x \in \mathbb{R}^N$ và tính toán năng lượng dùng được 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 $\overline{scb_k} = \{i | i \in scb_k \wedge x_i = 0\}$. Bộ này gồm có tất cả các chỉ số của băng thừa số định tỉ lệ scb_k mà hệ số được lượng tử hóa tới giá trị 0 bởi bộ mã hóa.

IGF có phương pháp băng con (không được miêu tả ở đây) được sử dụng để điền đầy khoảng trống phổ thu được từ sự lượng tử hóa thô của giá trị phổ MDCT tại phía bộ mã hóa bằng việc sử dụng giá trị khác 0 của MDCT đã truyền. x sẽ bao gồm thêm các giá trị mà thay thế tất cả các giá trị đến 0 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 khuyết trong băng khôi phục được tính toán 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 thừa số khuếch đại là:

$$x_i := g' x_i$$

cho tất cả $i \in \overline{scb_k}$ và k trong khoảng như đã định rõ ở trên.

Khôi phục hình dạng đường bao phổ của x thành tạo hình đường bao phổ ban đầu $\hat{x}$.

Với trình tự của sổ ngắn, tất cả các tính toán như đã phác họa ở trên về nguyên lý là tương tự, nhưng việc nhóm các băng thừa số định tỉ lệ được xem xét. Ký hiệu là $E_{k,l}$ giá trị năng lượng đã lượng tử hóa, được nhóm lại thu được từ dòng bit. Tính toán:

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

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

Chỉ số miêu tả j chỉ số của sổ của trình tự 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:

$$x_{j,i} := g'x_{j,i}$$

cho tất cả $i \in \overline{scb_{k,l}}$.

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

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

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

Một lần nữa, sau việc 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 gồm có 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 theo tham số biểu thị năng lượng phổ cho tập hợp 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 theo 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 theo tham số để thu được dữ liệu theo tham số được giải mã 902 cho tập hợp thứ hai của phần phổ thứ hai biểu thị năng lượng riêng cho băng khôi phục riêng, trong đó phần phổ thứ hai được đặt trong băng khôi phục. Thêm nữa, bộ tái tạo tần số 906 được cung cấp để khôi phục giá trị phổ của băng khôi phục gồm có 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 băng khôi phục, trong đó băng khôi phục bao gồm phần phổ thứ nhất và phần phổ thứ hai. Bộ tái tạo tần số 906 gồm có bộ tính toán 912 để xác định thông tin năng lượng còn lại bao gồm năng lượng đã tích lũy của phần phổ thứ nhất có tần số trong băng khôi phục. Thêm nữa, bộ tái tạo tần số 906 bao gồm bộ tính toán 918 để xác định thông tin năng lượng ô của các phần phổ tiếp theo của băng khôi phục và các giá trị tần số khác với phần phổ thứ nhất, trong đó, các giá trị tần số này có các tần số trong băng khôi phục, trong đó phần phổ tiếp theo được tạo ra bởi việc tái tạo tần số sử dụng phần phổ thứ nhất khác với phần phổ thứ nhất trong băng khôi phục.

Bộ tạo tần số 906 còn bao gồm bộ tính toán 419 để năng lượng khuyết trong băng khôi phục, và bộ tính toán 914 hoạt động sử dụng năng lượng tách biệt cho băng khôi phục và năng lượng còn lại được tạo bởi khối 912. Ngoài ra, bộ tạo tần số 906 bao gồm bộ điều chỉnh đường bao phổ 916 để điều chỉnh phần phổ tiếp theo trong băng khôi phục dựa trên thông tin năng lượng khuyết và thông tin năng lượng ô được tạo bởi khối 918.

Tham chiếu tới Fig.9c minh họa băng khôi phục nhất định 920. Băng khôi phục bao gồm phần phổ thứ nhất trong băng khôi phục như phần phổ thứ nhất 306 trên Fig.3a được minh họa dưới dạng biểu đồ tại 921. Ngoài ra, phần còn lại của các giá trị phổ trong băng khôi phục được tạo sử dụng vùng nguồn, ví dụ, từ băng thừa số định tỉ lệ 1, 2, 3 dưới tần số bắt đầu điền đầy khoảng trống thông minh 309 trên Fig.3a. Bộ tạo tần số 906 được tạo cấu hình để tạo các giá trị phổ thô cho các phần phổ thứ hai 922 và 923. Sau đó, thừa số khuếch đại g được tính toán như được minh họa trên Fig.9c để cuối cùng điều chỉnh các giá trị phổ thô trong các băng tần 922, 923 để thu được các phần phổ thứ hai được điều chỉnh và được khôi phục trong băng khôi phục 920 mà hiện tại có cùng độ phân giải phổ, tức là khoảng cách vạch phổ tương tự như phần phổ thứ nhất 921. Quan trọng để hiểu rằng phần phổ thứ nhất trong băng khôi phục được minh họa tại 921 trên Fig.9c được giải mã bởi bộ giải mã âm thanh 900 và không bị ảnh hưởng bởi khối được thực hiện điều chỉnh đường bao 916 trên Fig.9b. Thay vào đó, phần phổ thứ nhất trong băng khôi phục đã biểu thị 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 vạch phổ 904.

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

Thêm nữa, giá trị năng lượng E3 cho băng khôi phục tương ứng với băng thừa số định tỉ lệ 6 trên Fig.3b hoặc Fig.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 băng khôi phục 920 như đã tính toán trên 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 băng khôi phục. Sau đó, nó được giả sử rằng năng lượng của dữ liệu khoảng nguồn cho khối 922, 923 hoặc cho dữ liệu khoảng mục tiêu thô cho khối 922, 923 là bằng tám đơn vị năng lượng. Do đó, năng lượng khuyết của năm đơn vị được tính toán.

Dựa trên năng lượng khuyết bị tách ra bởi năng lượng ô tEk, hệ số khuếch đại của 0,79 được tính toán. Sau đó, vạch phổ 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 vạch phổ 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, băng khôi phục hoàn chỉnh được tính toán gồm có phần phổ thứ nhất trong băng khôi phục, và gồm vạch phổ phổ trong phần phổ thứ hai 922, 923 trong băng khôi phục 920.

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

Thêm nữa, ưu tiên các đường biên băng khôi phục trùng với các đường biên băng thừa số định tỉ lệ. Do đó, băng khôi phục có, trong một phương án, kích thước của sự tương đồng băng thừa số định tỉ 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 băng khôi phục cung cấp năng lượng của hai hoặc nhiều số nguyên cao hơn của băng thừa số định tỉ lệ. Do đó, khi được giả sử rằng việc tích lũy năng lượng được thực hiện cho băng thừa số

định tỉ lệ 4, băng thừa số định tỉ lệ 5 và băng thừa số định tỉ lệ 6, sau đó đường biên tần số thấp hơn của băng khôi phục 920 là bằng với đường biên thấp hơn của băng thừa số định tỉ lệ 4 và đường biên tần số cao hơn của băng khôi phục 920 trùng với đường biên của băng cao hơn của thừa số định tỉ lệ 6.

Tiếp theo, Fig.9d được thảo luận để thể hiện chức năng hơn khác của bộ giải mã trên Fig.9a. Bộ giải mã âm thanh 900 nhận các giá trị phổ được khử lượng tử ứng với các phần phổ thứ nhất của tập hợp thứ nhất của các phần phổ và, ngoài ra, các thừa số định tỉ lệ cho các băng thừa số định tỉ lệ như được minh họa trên Fig.3b được cung cấp tới khối định tỉ lệ nghịch đảo 940. Khối định tỉ lệ nghịch đảo 940 cung cấp tất cả các tập hợp thứ nhất của các phần phổ thứ nhất dưới tần số bắt đầu IGF 309 trên Fig.3a và, ngoài ra, các phần phổ thứ nhất trên tần số bắt đầu IGF, tức là, các phần phổ thứ nhất 304, 305, 306, 307 trên Fig.3a mà được đặt trong băng khôi phục như được minh họa tại 941 trên Fig.9d. Thêm nữa, phần phổ thứ nhất trong băng nguồn được sử dụng để điền đầy ô tần số trong băng 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 băng khôi phục được cung cấp như thông tin phụ tham số tới tín hiệu âm thanh được mã hóa như minh họa ở 943 trên Fig.9d. Sau đó, bộ điều chỉnh/tính toán đường bao 942 cung cấp các chức năng trên Fig.9b và Fig.9c và cuối cùng xuất ra các giá trị phổ được điều chỉnh cho phần phổ thứ hai trong băng khôi phục. Các giá trị đã điều chỉnh này 922, 923 cho phần phổ thứ hai trong băng khôi phục và phần phổ thứ nhất 921 trong băng khôi phục biểu thị vạch phổ 941 trên Fig.9d biểu diễn ghép phép biểu diễn phổ hoàn chỉnh của băng khôi phục.

Sau đó, tham chiếu tới Fig.10a và 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, nạp bộ phân tích phổ 1004, và bộ phân tích phổ 1004 được nối với bộ tính toán tham số 1006 trên một mặt và bộ mã hóa âm thanh 1008 mặt khác. Bộ mã hóa âm thanh 1008 cung cấp phép biểu diễn được mã hóa của tập hợp thứ nhất của các phần phổ thứ nhất và không gồm tập hợp thứ hai của các phần phổ thứ hai. Mặt khác, bộ tính toán theo tham số 1006 cung cấp thông tin năng lượng cho băng 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

phép biểu diễn được mã hóa thứ nhất của tập hợp thứ nhất của các phần phổ thứ nhất có độ phân giải phổ thứ nhất, trong đó bộ mã hóa âm thanh 1008 cung cấp các thừa số định tỉ lệ cho tất cả các băng của phép biểu diễn phổ được tạo bởi khối 1002. Ngoài ra, như được minh họa trên Fig.3b, bộ mã hóa cung cấp thông tin năng lượng ít nhất cho các băng khôi phục đặt, liên quan đến tần số, trên tần số bắt đầu IGF 309 như được minh họa trên Fig.3a. Do đó, để băng khôi phục tốt hơn là trùng với băng thừa số định tỉ lệ hoặc với các nhóm thừa số định tỉ lệ, hai giá trị được đưa ra,... thừa số định tỉ 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 băng thừa số định tỉ 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 gồm có bộ chuẩn hóa 1012 để chuẩn hóa năng lượng cho băng thông khác nhau đối với băng thông của băng khôi phục cụ thể. Để đạt được điều này, bộ chuẩn hóa 1012 nhận, như đầu vào, năng lượng trong băng và số lượng các giá trị phổ trong băng và bộ chuẩn hóa 1012 sau đó xuất ra năng lượng đã chuẩn hóa mỗi băng khôi phục/thừa số tỉ lệ.

Thêm nữa, bộ tính toán theo tham số 1006 trên Fig.10a gồm có bộ tính toán giá trị năng lượng nhận thông tin điều khiển lỗi hoặc bộ mã hóa âm thanh 1008 như minh họa bởi vạch phổ 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. Vì thế nên, 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 cũng đề cập đến việc nhóm phổ, tức là, việc nhóm hai băng thừa số định tỉ lệ thành băng khôi phục đơn. Vì thế, bộ tính toán giá trị năng lượng 1014 cung cấp giá trị năng lượng đơn cho từng băng đã được nhóm bao gồm phần phổ thứ nhất và thứ hai khi chỉ phần phổ được nhóm.

Fig.10d minh họa phương án tiếp theo 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 băng liên kề. Sau đó, trong khối 1018, giá trị năng lượng cho băng liên 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 giá trị (được chuẩn hóa) đơn cho cả các băng được tạo ra như đã xác

định trong khối 1020. Như minh họa bởi vạch phổ 1019, khối 1018 có thể được bỏ qua. Ngoài ra, sự tạo giá trị đơn lẻ cho hai hoặc nhiều hơn hai băng được thực hiện bởi khối 1020 có thể được điều khiển bởi bộ mã hóa điều khiển tốc độ truyền dữ liệu 1024. Do đó, khi tốc độ truyền dữ liệu được giảm, điều khiển tốc độ truyền dữ liệu được mã hóa 1024 điều khiển khối 1020 để tạo giá trị được tiêu chuẩn hóa đơn lẻ cho hai hoặc nhiều hơn hai băng mặc dù sự so sánh trong khối 1018 sẽ không cho phép nhóm các giá trị thông tin năng lượng.

Trong trường hợp bộ mã hóa âm thanh thực hiện việc nhóm hai hoặc nhiều hơn hai 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 hơn hai khối ngắn, sau đó, đối với hai hoặc nhiều hơn hai khối này, chỉ bộ đơn của các thừa số định tỉ lệ được tính toán và truyền. Trên phía bộ giải mã, bộ giải mã âm thanh sau đó sử dụng tập hợp tương tự của thừa số định tỉ lệ cho cả hai cửa sổ được nhóm.

Đối với việc tính toán thông tin năng lượng, giá trị phổ trong băng khôi phục được tích lũy trên hai hoặc nhiều hơn hai cửa sổ ngắn. Nói cách khác, có nghĩa là giá trị phổ trong băng 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 băng khôi phục này chứa khối ngắn. Sau đó, trên phía bộ giải mã, việc điều chỉnh đường bao được thảo luận đối với các hình vẽ từ Fig.9a tới Fig.9d không được thực hiện riêng rẽ cho từng khối ngắn nhưng được thực hiện cùng với nhau cho bộ các cửa sổ ngắn đã nhóm.

Việc chuẩn hóa tương ứng sau đó lại được sử dụng để thậm chí bất kỳ việc tạo nhóm trong tần số hoặc tạo nhóm trong thời gian được thực hiện, việc chuẩn hóa dễ dàng cho phép điều đó, để tính toán thông tin giá trị năng lượng trên phía bộ giải mã, một mặt chỉ giá trị thông tin năng lượng và số lượng các vạch phổ phổ trong băng khôi phục hoặc trong bộ các băng khôi phục đã được nhóm đã được biết đến.

Trong sơ đồ BWE thuộc tình trạng kỹ thuật, việc khôi phục vùng phổ HF ở trên đã có được gọi là tần số giao nhau thường dựa trên việc ráp nối phổ. Điển hình, vùng HF gồm nhiều việc ráp nối liền kề và từng việc ráp nối này được bắt nguồn từ các vùng thông dải (BP) của phổ LF thấp hơn tần số giao nhau đã có. Trong phép biểu diễn giàn lọc của tín hiệu như hệ thống sao chép bộ hệ số băng con liền kề ra khỏi phổ LF vào trong vùng mục tiêu. Đường biên của bộ đã lựa chọn là hệ thống điển hình phụ

thuộc và không phụ thuộc tín hiệu. Về lượng tín hiệu, việc lựa chọn việc ráp nối ổn định này dẫn tới âm sắc và màu sắc khó chịu của tín hiệu đã khôi phục.

Các cách tiếp cận khác chuyển tín hiệu LF tới HF thông qua sự biến điệu băng phụ đơn (Single Side Band - SSB) thích ứng tín hiệu. Như cách tiếp cận của sự phức tạp tính toán cao được so sánh với [1] khi chúng hoạt động tại tốc độ lấy mẫu cao trên mẫu miền thời gian. Cũng như vậy, việc ráp nối có thể không ổn định, đặc biệt cho tín hiệu không có âm điệu (ví dụ đoạn vô thanh), và do đó việc ráp nối thích ứng tín hiệu thuộc tình trạng kỹ thuật có thể đưa tới sự suy yếu thành tín hiệu.

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

Lợi thế của cấu hình IGF dựa trên MDCT là sự tích hợp trọn vẹn thành MDCT dựa trên lõi âm thanh, ví dụ việc mã hóa âm thanh tiên tiến MPEG (Advanced Audio Coding - AAC). Sự phân bổ việc biến đổi tương tự cho mã hóa âm thanh dạng sóng và để BWE làm giảm tất cả việc tính toán phức tạp cho việc mã hóa âm thanh một cách đáng kể.

Tuy nhiên, sáng chế cung cấp độ phân giải cho các vấn đề có tính ổn định vốn có được xây dựng trong sơ đồ ráp nối thích ứng trong tình trạng kỹ thuật.

Hệ thống được đề xuất dựa trên cơ sở quan sát rằng một số tín hiệu, lựa chọn ráp nối không được điều khiển có thể dẫn đến các thay đổi âm sắc và các sắc thái tín hiệu. Nếu tín hiệu là âm trong vùng nguồn phổ (spectral source region - SSR) nhưng như nhiều âm trong vùng đích phổ (spectral target region - STR), ráp nối như nhiều âm STR bởi âm SSR có thể dẫn đến âm sắc không tự nhiên. Âm sắc của tín hiệu có thể cũng thay đổi vì cấu trúc âm của tín hiệu có thể bị lệch đồng chỉnh hoặc thậm chí bị phá hủy bởi quy trình ráp nối.

Hệ thống IGF đã đưa ra thực hiện việc lựa chọn ô thông minh sử dụng tính tương quan chéo như việc đo độ giống nhau giữa SSR cụ thể và STR cụ thể. Tính tương quan chéo của hai tín hiệu cung cấp phép đo tương tự của các tín hiệu này và

việc trễ của tính tương quan tối đa và ký hiệu của nó. Vì thế, 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 có khả năng gần với cấu trúc phổ ban đầu.

Sự đóng góp cơ bản của hệ thống được đề xuất là sự lựa chọn của phép đo tương tự thích hợp, và các kỹ thuật để ổn định sự xử lý lựa chọn ô. Kỹ thuật được đề xuất cung cấp sự cân bằng tối ưu giữa sự thích ứng tín hiệu tức thời và, tại cùng một thời điểm, sự ổn định theo thời gian. Việc cung cấp tính ổn định theo thời gian là đặc biệt quan trọng cho các tín hiệu mà có sự giống nhau không đáng kể của SSR và STR và do đó biểu lộ giá trị có tương đồng chéo thấp hoặc nếu phép đo tương tự được sử dụng không xác định. 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 xác định đặc tính tập trung riêng của năng lượng của vùng phổ bất kỳ như được thể hiện trên Fig.12a (bên trái). Mặc dù đây là phương pháp có sẵn để điều chỉnh đường bao phổ và tính âm điệu của phổ đã khôi phục trong vùng mục tiêu, với một số tín hiệu các phương pháp này không có khả năng bảo toàn âm sắc tốt như được thể hiện trên Fig.12a (bên phải). Trong ví dụ thể hiện trên Fig.12a, độ lớn của phổ trong vùng mục tiêu của tín hiệu gốc như trên được gọi là tần số giao nhau f_{xover} (Fig.12a, bên trái) làm giảm gần như tuyến tính. Ngược lại, trong phổ đã khôi phục (Fig.12a, bên phải) bộ riêng biệt của độ nghiêng và các đỉnh được hiển thị được hiểu như giả âm sắc thái âm sắc.

Bước quan trọng của cách tiếp cận mới là để định rõ bộ các ô mà trong đó sự tương tự tiếp theo dựa trên việc lựa chọn có thể xảy ra. Thứ nhất, đường biên ô của cả khoảng nguồn và khoảng mục tiêu được định rõ trong sự phù hợp với từng đường. Do đó, vùng mục tiêu giữa tần số bắt đầu IGF của bộ mã hóa lõi $f_{IGFstart}$ và tần số có giá trị cao nhất $f_{IGFstop}$ bị tách thành số nguyên tùy ý $nTar$ của các ô, từng ô này có kích thước được định rõ riêng biệt. Sau đó, từng ô mục tiêu $tar[idx_tar]$, bộ của ô nguồn có kích thước bằng nhau $src[idx_src]$ được tạo ra. Bởi vậy, mức độ cơ bản của tính tự do của hệ thống IGF được xác định. Tổng số ô nguồn $nSrc$ được xác định bởi băng thông của vùng nguồn,

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

trong đó f_{IGFmin} là tần số có giá trị thấp nhất cho việc lựa chọn ô sao cho số nguyên $nSrc$ của các ô nguồn phù hợp với bw_{src} . Số tối thiểu của các ô nguồn là 0.

Để tăng thêm mức độ tự do cho việc lựa chọn và điều chỉnh, các ô nguồn có thể được xác định để chồng lấp lên nhau bằng hệ số chồng lấp 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 có giá trị.

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

Với ô mục tiêu, sự tương quan chéo được tính toán với các ô nguồn khác nhau tại độ trễ lên đến các ngăn $xcorr_maxLag$. Đối với ô mục tiêu đã cho idx_tar và ô nguồn idx_src , $xcorr_val[idx_tar][idx_src]$ 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 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 mục tiêu. 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.

Trong một số cảnh có thể xảy ra đó là kích thước của ô mục tiêu cụ thể lớn hơn kích thước của các ô nguồn có giá trị. Trong trường hợp này, ô nguồn có giá trị được lặp lại mỗi lần mà cần thiết để điền đầy ô mục tiêu đặc biệt một cách hoàn chỉnh. Việc này vẫn có thể thể hiện sự tương quan chéo giữa ô mục tiêu lớn và ô nguồn nhỏ hơn để có vị trí tốt nhất của ô nguồn trong ô mục tiêu 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 với 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, được quan tâm hàng đầu để đánh giá ô giống nhau. Việc làm trắng cũng giúp cho 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 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, thứ nhất việc 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 lợi ích 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 tách ra bởi 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 (discrete cosine transform - DCT), chỉ giữ lại các hệ số DCT thấp hơn (thiết lập tầng trên cùng tới 0) và sau đó tính toán DCT nghịch đảo.
- tính toán đường bao phổ của tập hợp các hệ số dự báo tuyến tính (Linear Prediction Coefficients - 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: Thứ nhất 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 bởi 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 mục tiêu, đối với tất cả ô mục tiêu $nTar$ 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 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 băng khác trong MDCT.

Fig.12c thể hiện ví dụ về sự tương quan giữa ô nguồn và ô mục tiêu. 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 miêu tả ở trên tính cho độ trễ lẻ.

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

- ô số $[nTar]$: chỉ số ô nguồn được lựa chọn trên ô mục tiêu
- ô dấu hiệu $[nTar]$: ký hiệu của ô mục tiêu
- ô điều biến $[nTar]$: độ trễ của sự tương quan trên ô mục tiêu

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 mục tiêu 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 $tileNum[nTar]$ 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 nhiễu âm có tính nhạc khó chịu giống giả âm.

Để ước tính dạng giả âm, tập hợp của các ô nguồn sẽ được lược bớt vì thế giữ được các thành phần của bộ nguồn không giống nhau một cách tối đa. Điều này đạt được trên tập hợp của các ô nguồn:

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

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

$$T[i] = S_x[i][1] + S_x[i][2]..+ S_x[i][n]$$

Ở đây T đại diện cho 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 hạ từ tập hợp của 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 như ô biểu diễn cho tập con này. Bằng cách này, chúng ta 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.

Đưa 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 chia nhỏ với s_3 để được tương quan tối đa với các ô khác, s_3 được giữ lại vì nó là ô nguồn hữu ích trong khung ưu tiên, và do đó việc giữ lại ô này trong bộ 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 mục tiêu. Điều này có thể xảy ra nếu tính tương quan chéo giữa nguồn i và mục tiêu j, được biểu diễn là $T_x[i][j]$ rất thấp cho tất cả 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 tất cả nTar của khung k này.

Hai kỹ thuật ở trên làm giảm rất nhiều giả âm mà xuất hiện từ sự thay đổi nhanh chóng việc thiết lập số lượng ô khung chéo. Ưu điểm được bổ sung khác của việc lược bớt ô và làm ổn định ô này là không thêm nhu cầu thông tin để gửi tới bộ giải mã hoặc là sự thay đổi của cấu trúc bộ giải mã là cần thiết. Việc lược bớt ô được đề xuất là cách nhẹ nhàng cho việc làm giảm nhiều âm có tính nhạc tiềm năng giống các giả âm hoặc nhiều âm quá thừa 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.

Thêm nữa, 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. Thêm nữa, bộ tạo ra 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 từng ô tần số mục tiêu hoặc băng khôi phục mục tiêu thông tin khoảng nguồn. Bộ tạo ra tần số 1106 sau đó sử dụng sự tái tạo tần số bằng việc sử dụng giá trị phổ từ khoảng nguồn được nhận dạng bởi thông tin phù hợp để tạo ra dữ liệu phổ cho khoảng mục tiêu. 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 đã 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 khoảng nguồn cho độ trễ lẻ, 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 học 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 sao, bộ tái tạo tần số 1106 có thể bao gồm bộ chuyển 1124 phản ứng lại dấu hiệu làm trắng được nhận 1123. Khi dấu hiệu làm trắng được thiết lập, đầu ra của ô khoảng nguồn cho độ trễ lẻ được nhập vào bộ lọc làm trắng 1120. Sau đó, tuy

nhiên, dấu hiệu làm trắng không được thiết lập cho băng khôi phục nhất định, sau đó vạch phổ rẽ nhánh 1126 được kích hoạt sao cho đầu ra khối 1120 được cung cấp cho khối điều chỉnh đường bao phổ 1128 mà không có bất kỳ sự làm trắng nào.

Có thể nhiều hơn một mức làm trắng (1123) đã ký hiệu trong đường truyền dữ liệu và các mức này có thể được tạo tín hiệu 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 và 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 được miêu tả trước đó).

Bộ tái tạo tần số phía bộ giải mã có thể được điều khiển bởi khoảng nguồn ID 1121 khi chỉ áp dụng sơ đồ lựa chọn ô phổ thô. Tuy nhiên, khi, sơ đồ lựa chọn ô phổ đã điều chỉnh chính xác được sử dụng, sau đó, thêm vào đó, độ trễ khoảng nguồn 1119 được cung cấp. Thêm nữa, được đề xuất rằng việc tính toán sự tương quan cung cấp kết quả âm, sau đó, thêm vào đó, ký hiệu của sự tương quan có thể cũng được sử dụng cho khối 1120 sao cho các vạch phổ 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 đã thảo luận trên Fig.11a và Fig.11b đảm bảo rằng chất lượng âm thanh tốt nhất thu được do thực tế là khoảng nguồn trùng hợp nhất cho khoảng dự định hoặc khoảng mục tiêu nhất định được tính toán trên phía bộ mã hóa và được sử dụng trên 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 lập mã lỗi 1136. Bộ lập mã lỗi 1136 xuất ra khoảng

nguồn được mã hóa và bộ tính toán tham số 1134 xuất ra thông tin phù hợp cho các khoảng mục tiêu.

Khoảng nguồn được mã hóa được truyền tới bộ giải mã cùng với thông tin trùng hợp cho khoảng mục tiêu sao cho bộ giải mã được minh họa trên Fig.11a trong vị trí để thực hiện việc tái tạo tần số.

Bộ tính toán tham số 1134 được tạo cấu hình để tính toán điểm tương tự giữa phần phổ thứ nhất và phần phổ thứ hai và để xác định, dựa trên điểm tương tự đã tính toán, để phần phổ thứ hai phù hợp phần phổ thứ nhất phù hợp với phần phổ thứ hai. Ưu tiên hơn, các kết quả phù hợp với khoảng nguồn khác nhau và khoảng mục tiêu khác nhau như được minh họa trên Fig.12a và 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 khoảng mục tiêu đã xác định trước trong bộ thứ hai của phần phổ thứ hai hoặc khoảng nguồn đã xác định trước trong tập hợp thứ nhất của phần phổ thứ nhất như minh họa, ví dụ, trên Fig.12b. Tốt hơn là, khoảng mục tiêu đã xác định trước là không chồng lấp hoặc khoảng nguồn được xác định trước là chồng lấp. Khi khoảng nguồn 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 trên Fig.3a, và tốt hơn là, vùng mục tiêu xác định trước bao trùm vùng phổ thấp hơn trùng khớp với tần số thấp hơn với tần số bắt đầu điền đầy khoảng trống vì thế khoảng mục tiêu được đặt ở trên tần số bắt đầu điền đầy khoảng trống và khoảng nguồn được đặt dưới tần số bắt đầu điền đầy khoảng trống.

Như đã thảo luận, độ tạo hạt mịn thu được bằng cách so sánh khoảng mục tiêu và khoảng nguồn mà không có bất kỳ độ trễ nào tới khoảng nguồn và khoảng nguồn tương tự, nhưng với độ trễ nhất định. Độ trễ này được áp dụng trong bộ tính toán tương quan chéo 1140 trên Fig.11d và việc lựa chọn cặp trùng hợp được thực hiện cuối cùng bởi bộ lựa chọn 1144.

Ngoài ra, ưu tiên hơn để thực hiện sự làm trắng các khoảng nguồn và/hoặc mục tiêu được minh họa tại khối 1142. Khối 1142 sau đó cung cấp dấu hiệu làm trắng tới các dòng bit mà được sử dụng để điều khiển bộ chuyển mã hóa bên 1123 trên Fig.11b. Thêm nữa, 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 khoảng nguồn ID cho khoảng mục tiêu, độ trễ, dấu hiệu và khối 1142 cung cấp bổ sung dấu hiệu làm trắng.

Thêm nữa, 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 khoảng nguồn tiềm năng mà trong đó việc ráp nối 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, sau đó một trong số hai ô nguồn bị loại bỏ từ tập hợp của nguồn tiềm năng và ô nguồn đã loại bỏ không được sử dụng nhiều hơn cho việc xử lý thêm nữa 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 khoảng nguồn khác nhau và khoảng mục tiêu khác nhau được thể hiện trong khối 1140.

Việc thực hiện khác nhau được miêu tả với các hình minh họa khác nhau. Các hình vẽ từ Fig.1a đến Fig.5c đề cập tới sơ đồ bộ mã hóa/giải mã đầy đủ bằng thông đầy đủ hoặc tốc độ đầy đủ. Các hình vẽ từ Fig.6a đến Fig.7e đề cập tới sơ đồ sự phối hợp bộ mã hóa/giải mã với việc xử lý TNS hoặc TTS. Các hình vẽ từ Fig.8a đến Fig.8e đề cập tới 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 thông tin năng lượng cụ thể và ứng dụng, và các hình vẽ từ Fig.11a đến 12c đề cập cách cụ thể của sự lựa chọn ô.

Tất cả các khía cạnh khác nhau này có thể là tính sáng tạo khi sử dụng độc lập của từng khía cạnh, nhưng, thêm nữa có thể luôn được áp dụng với nhau như minh họa cơ bản trên Fig.2a và Fig.2b. Tuy nhiên, việc xử lý kênh đôi cụ thể có thể được áp dụng cho giản đồ bộ mã hóa/bộ giải mã được minh họa trên Fig.13, tương tự đúng cho việc xử lý TNS/TTS, việc tính toán và áp dụng thông tin năng lượng đường bao trong băng khôi phục hoặc sự nhận dạng khoảng nguồn thích ứng và áp dụng tính tương quan trên 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 khoảng nguồn thích ứng hoặc với dạng khác của việc tính toán năng lượng cho phép 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 phương pháp có thể được thực hiện bằng (hoặc sử dụng) thiết bị phần cứng, như ví dụ, bộ vi xử lý, máy tính có thể lập trình hoặc mạch điện tử. Trong một số phương án, một số hoặc nhiều bước xử lý phương pháp quan trọng có thể có thực hiện bằng các thiết bị như vậy.

Phụ thuộc vào các yêu cầu thực hiện nhất định, các phương án của sáng chế có thể được thực hiện trong phần cứng hoặc trong phần mềm. Phương án có thể được thực hiện sử dụng vật ghi lưu trữ chuyên 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 điện tử mà có khả năng liên kết với hệ thống máy tính có thể lập trình được, sao cho 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 gồm có 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, phương án của phương pháp sáng chế là, do đó, chương trình máy tính có mã chương trình để thực hiện một trong các phương pháp được miêu 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) gồm có, 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, môi trường lưu trữ số hoặc môi trường được ghi thường là hữu hình và/hoặc không chuyển dịch.

Phương án khác của các phương pháp theo sáng chế là, do đó, 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, dòng dữ liệu hoặc chuỗi các tín hiệu có thể, ví dụ, được tạo cấu hình để được truyền thông qua sự kết nối truyền thông dữ liệu, ví dụ, thông qua liên mạng.

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

Phương án khác gồm có 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, điện tử hoặc quang học) chương trình máy tính thực hiện một trong các phương pháp được mô tả ở đây đến bộ nhận. Ví dụ, bộ nhận có thể là máy tính, thiết bị di động, thiết bị nhớ hoặc thiết bị tương tự. Thiết bị hoặc hệ thống có thể, ví dụ, bao gồm máy chủ tệp tin để chuyển chương trình máy tính đến bộ nhận.

Theo một số phương án, thiết bị logic lập trình (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. Trong một số phương án, mảng cổng lập trình được dạng trường có thể kết hợp với bộ vi xử lý để thực hiện một trong số các phương pháp được mô tả ở đây. Thông thường, các phương pháp tốt hơn là được thực hiện bởi thiết bị phần cứng bất kỳ.

Các phương án được mô tả ở trên chỉ mang tính minh họa cho các nguyên lý của sáng chế. Điều này được hiểu rằng việc điều chỉnh và biến đổi của sự sắp xếp và các chi tiết được mô tả ở đây sẽ rõ ràng với những người có trình độ trong lĩnh vực kỹ thuật. Do đó, mục đích chỉ bị giới hạn bởi phạm vi của các điểm yêu cầu bảo hộ sắp

đưa ra dưới đây và không bị giới hạn bởi các chi tiết cụ thể được biểu diễn bằng cách mô tả và giải thích của các phương án ở đây.

YÊU CẦU BẢO HỘ

1. Thiết bị tạo tín hiệu kênh đôi được giải mã, bao gồm:

bộ xử lý âm thanh (802) để giải mã tín hiệu kênh đôi được mã hóa để thu được tập hợp thứ nhất của các phần phổ thứ nhất;

bộ giải mã theo tham số (804) để cung cấp dữ liệu theo tham số cho tập hợp thứ hai của các phần phổ thứ hai và sự nhận dạng kênh đôi nhận dạng phép biểu diễn kênh đôi khác nhau thứ nhất hoặc thứ hai cho các phần phổ thứ hai; và

bộ tái tạo tần số (806) để tạo ra phần phổ thứ hai phụ thuộc vào phần phổ thứ nhất của tập hợp thứ nhất của các phần phổ thứ nhất, dữ liệu theo tham số cho phần thứ hai và sự nhận dạng kênh đôi cho phần thứ hai,

trong đó bộ xử lý âm thanh được tạo cấu hình để giải mã tập hợp thứ nhất của các phần phổ thứ nhất phù hợp với sự nhận dạng kênh đôi tiếp theo cho tập hợp thứ nhất của các phần phổ thứ nhất và để biến đổi (832) tập hợp thứ nhất của các phần phổ thứ nhất sao cho phép biểu diễn kênh đôi của tập hợp thứ nhất của các phần phổ thứ nhất và phép biểu diễn kênh đôi thứ hai của tập hợp thứ nhất của các phần phổ thứ nhất được thu, và

trong đó bộ tái tạo tần số (806) được tạo cấu hình để sử dụng phép biểu diễn kênh đôi thứ nhất hoặc phép biểu diễn kênh đôi thứ hai của tập hợp thứ nhất của các phần phổ thứ nhất như được biểu thị trong sự nhận dạng kênh đôi cho phần thứ hai.

2. Thiết bị theo điểm 1, trong đó nhận dạng kênh đôi nhận dạng hoặc xử lý phân tách của hai kênh của tín hiệu kênh đôi hoặc xử lý ghép của hai kênh của tín hiệu kênh đôi, và

trong đó bộ tái tạo tần số (806) được tạo cấu hình để tái tạo phần phổ thứ hai cho kênh thứ nhất trong số hai kênh và phần phổ thứ hai cho kênh thứ hai trong số hai kênh sử dụng phần thứ nhất của kênh thứ nhất và phần thứ nhất của kênh thứ hai, trong đó phần thứ nhất của kênh thứ nhất và phần thứ nhất của kênh thứ hai nằm trong phép biểu diễn kênh đôi được nhận dạng bởi sự nhận dạng kênh đôi cho phần phổ thứ hai.

3. Thiết bị theo điểm 1 hoặc điểm 2,

trong đó sự nhận dạng kênh đôi nhận dạng hoặc xử lý phân tách của hai kênh của tín hiệu kênh đôi hoặc xử lý ghép của hai kênh của tín hiệu kênh đôi, và

trong đó bộ tái tạo tần số (806) được tạo cấu hình để tái tạo phép biểu diễn ghép của hai kênh trong phân phổ thứ hai như được nhận dạng bởi sự nhận dạng kênh đôi, và

trong đó bộ tái tạo tần số (806) còn bao gồm bộ biến đổi phép biểu diễn (842) để biến đổi phép biểu diễn ghép của phân phổ thứ hai thành phép biểu diễn phân tách cho phân phổ thứ hai.

4. Thiết bị theo điểm 3, trong đó bộ biến đổi phép biểu diễn sử dụng các tham số biểu diễn ghép bổ sung cho biến đổi phép biểu diễn.

5. Thiết bị theo điểm 3, trong đó phép biểu diễn ghép là phép biểu diễn giữa/bên, và trong đó bộ biến đổi phép biểu diễn được tạo cấu hình để vận hành dựa trên phương trình

$$leftTile = 0,5 * (midTile + sideTile)$$

$$rightTile = 0,5 * (midTile - sideTile)$$

trong đó leftTile và rightTile là phép biểu diễn phân tách cho phân phổ thứ hai, và trong đó midTile và sideTile là phép biểu diễn ghép cho phân phổ thứ hai.

6. Thiết bị theo điểm 3, trong đó phép biểu diễn ghép là phép biểu diễn bao gồm hệ số dự báo bổ sung, và trong đó bộ biến đổi phép biểu diễn được tạo cấu hình để vận hành dựa trên phương trình sau:

$$sideTile[k] = sideTile[k] - predictionCoeff * midTile[k]$$

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

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

khi hướng dự báo từ bên tới giữa, hoặc:

$$midTile1[k] = midTile[k] - predictionCoeff * sideTile[k]$$

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

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

khi hướng dự báo được biểu thị từ bên tới giữa,

trong đó leftTile và rightTile là phép biểu diễn phân tách cho phân phổ thứ hai, và trong đó midTile và sideTile là phép biểu diễn ghép cho phân phổ thứ hai, và trong đó predictionCoefficient là hệ số dự báo bổ sung.

7. Thiết bị theo một điểm trong số các điểm đã nêu,

trong đó dữ liệu theo tham số cho tập hợp thứ hai của các phân phổ thứ hai được đưa ra riêng cho mỗi kênh của phép biểu diễn kênh đôi, và

trong đó bộ tái tạo tần số (806) được tạo cấu hình để biến đổi (840) dữ liệu theo tham số cho phần phổ thứ hai thành phép biểu diễn ghép cho phần phổ thứ hai và để áp dụng dữ liệu theo tham số vào phép biểu diễn ghép của phần phổ thứ nhất, khi sự nhận dạng kênh đôi nhận dạng phép biểu diễn ghép cho phần phổ thứ hai.

8. Thiết bị theo một trong số các điểm đã nêu, trong đó các phần phổ thứ hai tương ứng với các băng tần, và trong đó sự nhận dạng kênh đôi là dãy các dấu hiệu, một dấu hiệu cho mỗi băng tần, và trong đó bộ giải mã theo tham số (804) được tạo cấu hình để kiểm tra, liệu dấu hiệu được thiết lập hay không và để điều khiển sự tái tạo tần số phù hợp với dấu hiệu để sử dụng hoặc phép biểu diễn kênh đôi thứ nhất hoặc phép biểu diễn kênh đôi thứ hai của phần phổ thứ nhất của tín hiệu kênh đôi.

9. Thiết bị theo một điểm bất kỳ trong số các điểm đã nêu, trong đó bộ giải mã theo tham số (804) được tạo cấu hình để cung cấp sự nhận dạng kênh đôi nữa cho tập hợp thứ nhất của các phần phổ thứ nhất biểu thị hoặc phép biểu diễn kênh đôi khác nhau thứ nhất hoặc thứ hai cho phần phổ thứ nhất, và

trong đó thiết bị được tạo cấu hình để giải mã phép biểu diễn kênh đôi thứ hai như được biểu thị bởi sự nhận dạng kênh đôi cho phần phổ thứ nhất, và

trong đó bộ tái tạo tần số (806) được tạo cấu hình để biến đổi (832) phép biểu diễn kênh đôi thứ hai thành phép biểu diễn kênh đôi thứ nhất sau giải mã lỗi (830).

10. Thiết bị theo một trong số các điểm đã nêu, còn bao gồm bộ tổ hợp (846) để tổ hợp tập hợp thứ nhất của các phần phổ thứ nhất được tạo bởi bộ xử lý âm thanh (802) và phần phổ thứ hai được khôi phục được tạo bởi bộ tái tạo tần số (806) để thu được tín hiệu kênh đôi được giải mã.

11. Thiết bị theo một điểm trong số các điểm đã nêu,

trong đó bộ giải mã theo tham số được tạo cấu hình để cung cấp bổ sung, cho phần phổ thứ hai, sự nhận dạng băng nguồn (1211) biểu thị phần phổ thứ nhất cụ thể để được sử dụng để tái tạo phần phổ thứ hai, và

trong đó bộ tái tạo tần số (806) được tạo cấu hình để tái tạo phần phổ thứ hai sử dụng phần phổ thứ nhất được nhận dạng bởi sự nhận dạng băng nguồn (833).

12. Thiết bị theo một điểm trong số các điểm đã nêu,

trong đó bộ tái tạo tần số (806) bao gồm bộ biến đổi phép biểu diễn (832) để cung cấp phép biểu diễn kênh đôi thứ nhất và thứ hai của tập hợp thứ nhất của các phần phổ thứ nhất được tạo bởi bộ xử lý âm thanh (830),

trong đó bộ tái tạo tần số (806) còn bao gồm bộ tạo ô tần số (836) để tạo dữ liệu thô cho mỗi kênh của phép biểu diễn kênh được nhận dạng bởi sự nhận dạng kênh đôi và sử dụng sự nhận dạng khoảng nguồn (833) biểu thị các phần phổ thứ nhất được sử dụng để tạo dữ liệu thô,

trong đó bộ tái tạo tần số (806) còn bao gồm bộ biến đổi tham số (840) để biến đổi các tham số được cung cấp trong phép biểu diễn kênh đôi thứ nhất thành phép biểu diễn kênh đôi thứ hai cho các tham số, trong đó dữ liệu thô cho mỗi kênh được cung cấp trong phép biểu diễn kênh đôi thứ hai bởi bộ tạo ô tần số (836).

trong đó bộ tái tạo tần số (806) còn bao gồm bộ điều chỉnh đường bao (838) để điều chỉnh đường bao của mỗi kênh của phép biểu diễn kênh đôi, phép biểu diễn kênh đôi là phép biểu diễn kênh đôi thứ hai.

trong đó bộ tái tạo (806) còn bao gồm bộ biến đổi phép biểu diễn (842) để biến đổi phép biểu diễn kênh đôi của các giá trị phổ trong phần phổ thứ hai thành phép biểu diễn kênh đôi thứ nhất,

trong đó thiết bị còn bao gồm bộ chuyển đổi tần số-thời gian (846) để chuyển đổi phép biểu diễn được tạo bởi bộ biến đổi phép biểu diễn (842) từ miền phổ thành miền thời gian.

13. Bộ mã hóa âm thanh để mã hóa tín hiệu âm thanh kênh đôi, bao gồm:

bộ chuyển đổi thời gian-phổ (860) để chuyển đổi tín hiệu âm thanh kênh đôi thành phép biểu diễn phổ;

bộ phân tích phổ (860) để cung cấp sự biểu thị của tập hợp thứ nhất của các phần phổ thứ nhất được mã hóa với độ phân giải phổ thứ nhất và của tập hợp thứ hai của các phần phổ thứ hai được mã hóa bởi độ phân giải phổ thứ hai, độ phân giải phổ thứ hai nhỏ hơn độ phân giải phổ thứ nhất,

bộ phân tích kênh đôi (864) để phân tích các băng của tín hiệu âm thanh kênh đôi trong khoảng khôi phục để xác định nhận dạng kênh đôi cho băng khôi phục nhận dạng phép biểu diễn kênh đôi thứ nhất hoặc phép biểu diễn kênh đôi thứ hai, trong đó

bằng khôi phục bao gồm phần phổ thứ hai của tập hợp thứ hai của các phần phổ thứ hai;

bộ giải mã lời (870) để mã hóa tập hợp thứ nhất của các phần phổ thứ nhất để cung cấp phép biểu diễn được mã hóa thứ nhất; và

bộ tính toán tham số (868) để tính toán dữ liệu theo tham số trên tập hợp thứ hai của các phần phổ thứ hai,

trong đó tín hiệu âm thanh kênh đôi được mã hóa bao gồm phép biểu diễn được mã hóa thứ nhất, phép biểu diễn theo tham số được mã hóa và sự nhận dạng kênh đôi cho tập hợp thứ hai của các phần phổ thứ hai.

14. Bộ mã hóa âm thanh theo điểm 13, còn bao gồm bộ biến đổi (862) để biến đổi các phần phổ thứ nhất thành phép biểu diễn kênh đôi được biểu thị bởi sự nhận dạng kênh đôi và trong đó bộ phân tích phổ (860) được tạo cấu hình để phân tích phép biểu diễn kênh đôi được xuất ra bởi bộ biến đổi theo băng (862).

15. Bộ mã hóa âm thanh theo một điểm trong số các điểm 13 hoặc 14,

trong đó bộ phân tích kênh đôi (864) được tạo cấu hình để thực hiện tính toán tương quan giữa phần phổ thứ hai của kênh thứ nhất của phép biểu diễn kênh đôi và phần phổ thứ hai của kênh đôi của phép biểu diễn kênh đôi để xác định hoặc phép biểu diễn kênh đôi phân tách hoặc phép biểu diễn kênh đôi ghép.

16. Bộ mã hóa âm thanh theo một điểm trong số các điểm từ 13 đến 15,

trong đó bộ phân tích phổ (860) được tạo cấu hình để so sánh các kết quả phù hợp cho các phần phổ khác nhau của ít nhất một kênh của phép biểu diễn kênh đôi tới phần phổ thứ nhất khác của ít nhất một kênh của phép biểu diễn kênh đôi để xác định cặp phù hợp của phần phổ thứ nhất của ít nhất một kênh và phần phổ thứ hai của ít nhất một kênh và để cung cấp thông tin phù hợp (833) cho cặp phù hợp nhất, và

trong đó bộ mã hóa âm thanh được tạo cấu hình để xuất ra, ngoài tín hiệu âm thanh được mã hóa, sự biểu thị phù hợp (833) cho phần phổ thứ hai.

17. Bộ mã hóa âm thanh theo một trong số các điểm từ 13 đến 16 bao gồm bộ biến đổi theo băng (862) có đầu vào được nối với đầu ra của bộ chuyển đổi thời gian-phổ (860),

trong đó bộ phân tích phổ (860) được tạo cấu hình để nhận, như đầu vào, đầu ra của bộ biến đổi theo băng (862);

trong đó bộ phân tích kênh đôi (864) được tạo cấu hình để phân tích đầu ra của bộ chuyển đổi thời gian-phổ (860) và để cung cấp kết quả phân tích để điều khiển bộ biến đổi theo băng (862),

trong đó bộ mã hóa âm thanh được tạo cấu hình để mã hóa đầu ra của bộ biến đổi theo băng (862) được điều khiển bởi bộ phân tích phổ (860), sao cho chỉ tập hợp thứ nhất của các phần phổ thứ nhất được mã hóa bởi bộ mã hóa lõi (870), và

trong đó bộ mã hóa tham số (868) được tạo cấu hình để mã hóa theo tham số tập hợp thứ hai của các phần phổ thứ hai như được biểu thị bởi bộ phân tích phổ (860) trong đầu ra của bộ biến đổi theo băng (862).

18. Phương pháp tạo tín hiệu kênh đôi được giải mã, phương pháp này bao gồm:

giải mã (802) tín hiệu kênh đôi được mã hóa để thu được tập hợp thứ nhất của các phần phổ thứ nhất;

cung cấp (804) dữ liệu theo tham số cho tập hợp thứ hai của các phần phổ thứ hai và sự nhận dạng kênh đôi nhận dạng phép biểu diễn kênh đôi khác nhau thứ nhất hoặc thứ hai cho các phần phổ thứ hai; và

tái tạo (806) phần phổ thứ hai phụ thuộc vào phần phổ thứ nhất của tập hợp thứ nhất của các phần phổ thứ nhất, dữ liệu theo tham số cho phần thứ hai và nhận dạng kênh đôi cho phần thứ hai,

trong đó giải mã (802) bao gồm việc giải mã tập hợp thứ nhất của các phần phổ thứ nhất phù hợp với sự nhận diện kênh đôi tiếp theo cho tập hợp thứ nhất của các phần phổ thứ nhất và biến đổi (832) tập hợp thứ nhất của các phần phổ thứ nhất sao cho phép biểu diễn kênh đôi thứ nhất của tập hợp thứ nhất của các phần phổ thứ nhất và phép biểu diễn kênh đôi thứ hai của tập hợp thứ nhất của các phần phổ thứ nhất được thu, và

trong đó tái tạo (806) bao gồm sử dụng phép biểu diễn kênh đôi thứ nhất hoặc phép biểu diễn kênh đôi thứ hai của tập hợp thứ nhất của các phần phổ thứ nhất như được biểu thị trong sự nhận dạng kênh đôi cho phần thứ hai.

19. Phương pháp mã hóa tín hiệu âm thanh kênh đôi, phương pháp này bao gồm:

chuyển đổi (860) tín hiệu âm thanh kênh đôi thành phép biểu diễn phổ;

cung cấp (860) sự biểu thị của tập hợp thứ nhất của các phần phổ thứ nhất được mã hóa với độ phân giải phổ thứ nhất và của tập hợp thứ hai của các phần phổ thứ hai

được mã hóa bởi độ phân giải phổ thứ hai, độ phân giải phổ thứ hai nhỏ hơn độ phân giải phổ thứ nhất;

phân tích (864) các băng của tín hiệu âm thanh kênh đôi trong khoảng khôi phục để xác định sự nhận dạng kênh đôi cho băng khôi phục hoặc nhận dạng phép biểu diễn kênh đôi thứ nhất hoặc phép biểu diễn kênh đôi thứ hai, trong đó băng khôi phục bao gồm phần phổ thứ hai của tập hợp thứ hai của các phần phổ thứ hai;

mã hóa (870) tập hợp thứ nhất của các phần phổ thứ nhất để cung cấp phép biểu diễn được mã hóa thứ nhất; và

tính toán (868) dữ liệu theo tham số trên tập hợp thứ hai của các phần phổ thứ hai,

trong đó tín hiệu âm thanh kênh đôi được mã hóa bao gồm phép biểu diễn được mã hóa thứ nhất, phép biểu diễn theo tham số được mã hóa và sự nhận dạng kênh đôi cho tập hợp thứ hai của các phần phổ thứ hai.

20. Vật ghi có thể đọc được bằng máy tính bao gồm chương trình máy tính để thực hiện, khi chạy trên máy tính hoặc bộ xử lý, phương pháp theo điểm 18.

21. Vật ghi có thể đọc được bằng máy tính bao gồm chương trình máy tính để thực hiện, khi chạy trên máy tính hoặc bộ xử lý, phương pháp theo điểm 19.

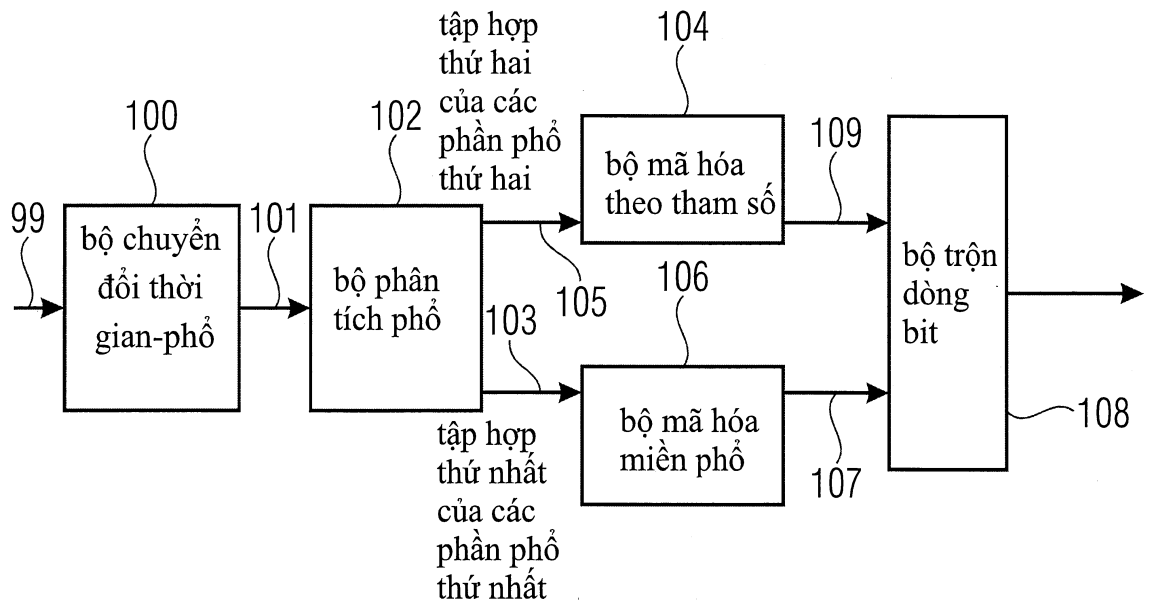


FIG 1A

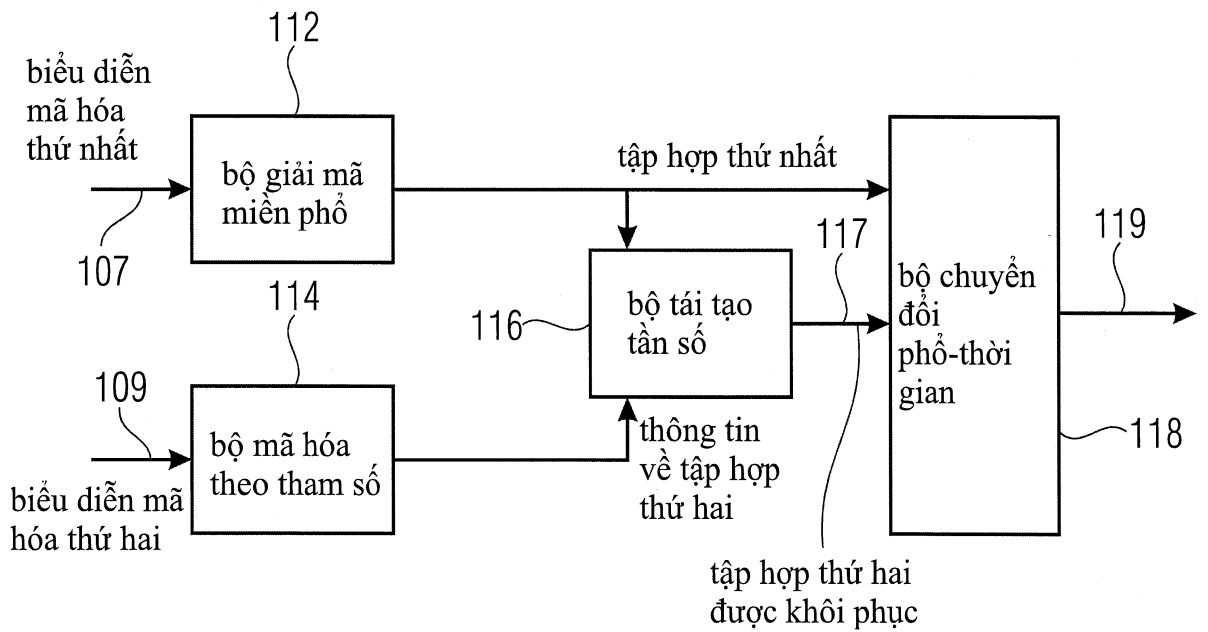


FIG 1B

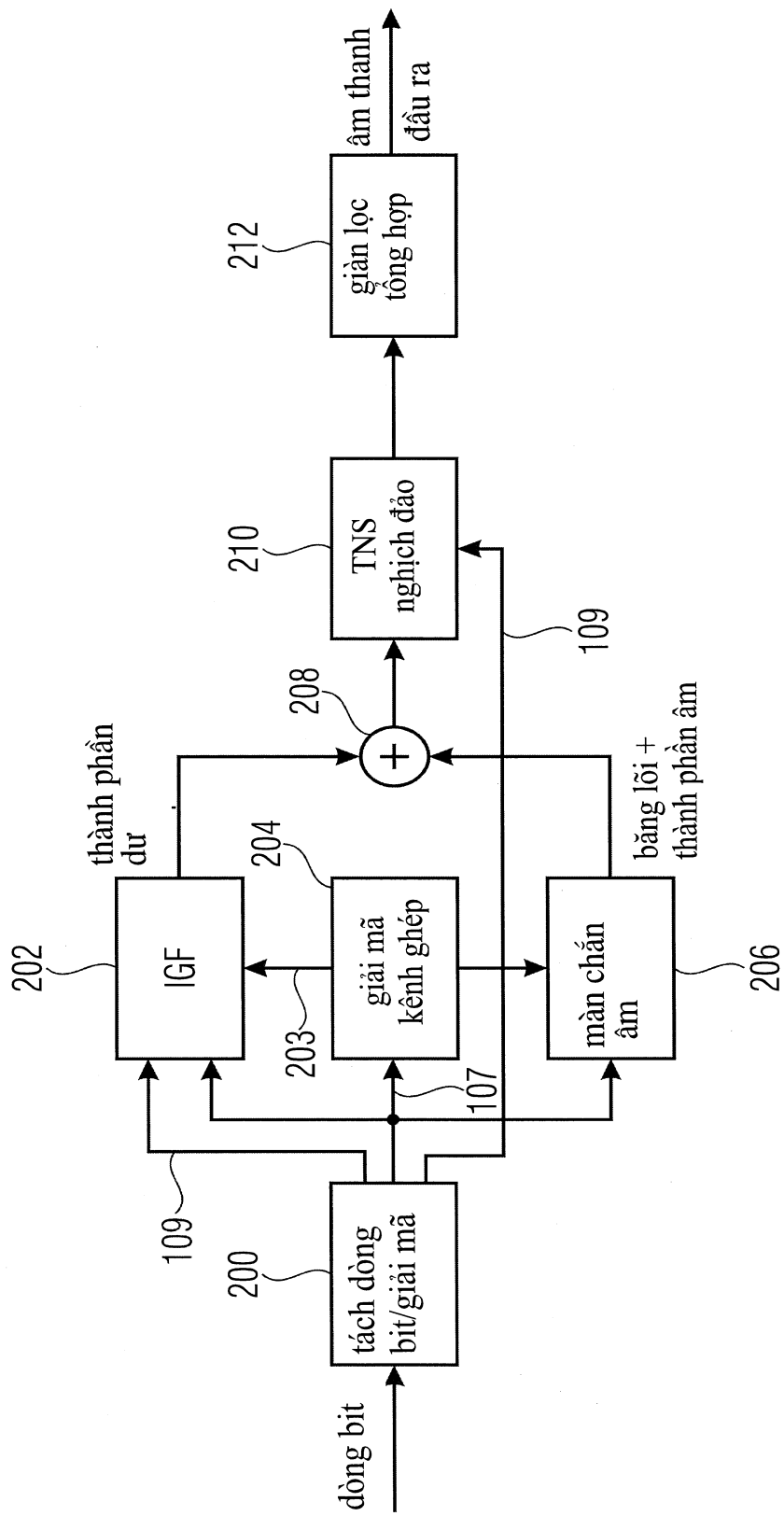


FIG 2A

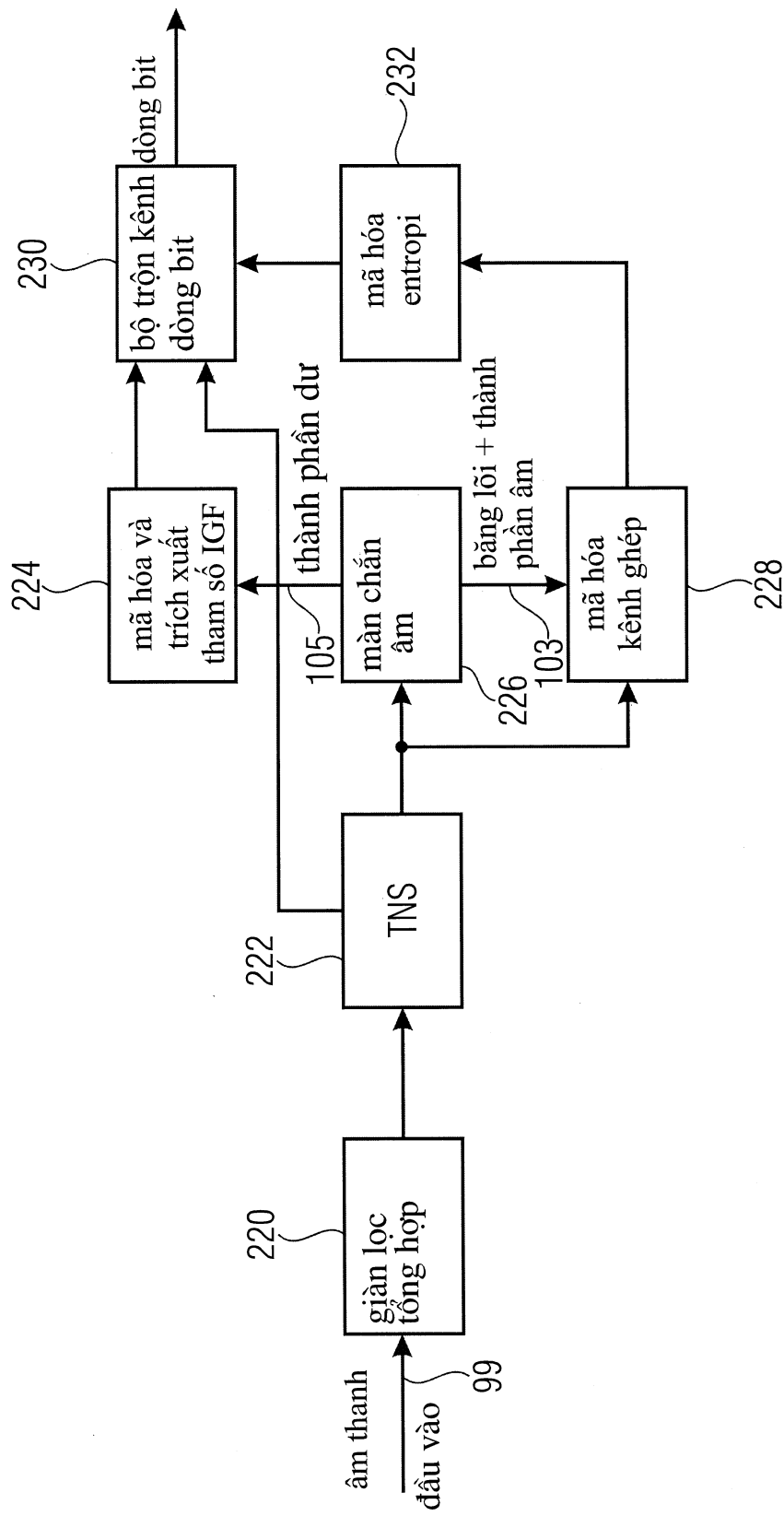


FIG 2B

- 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 mỗi SCB);

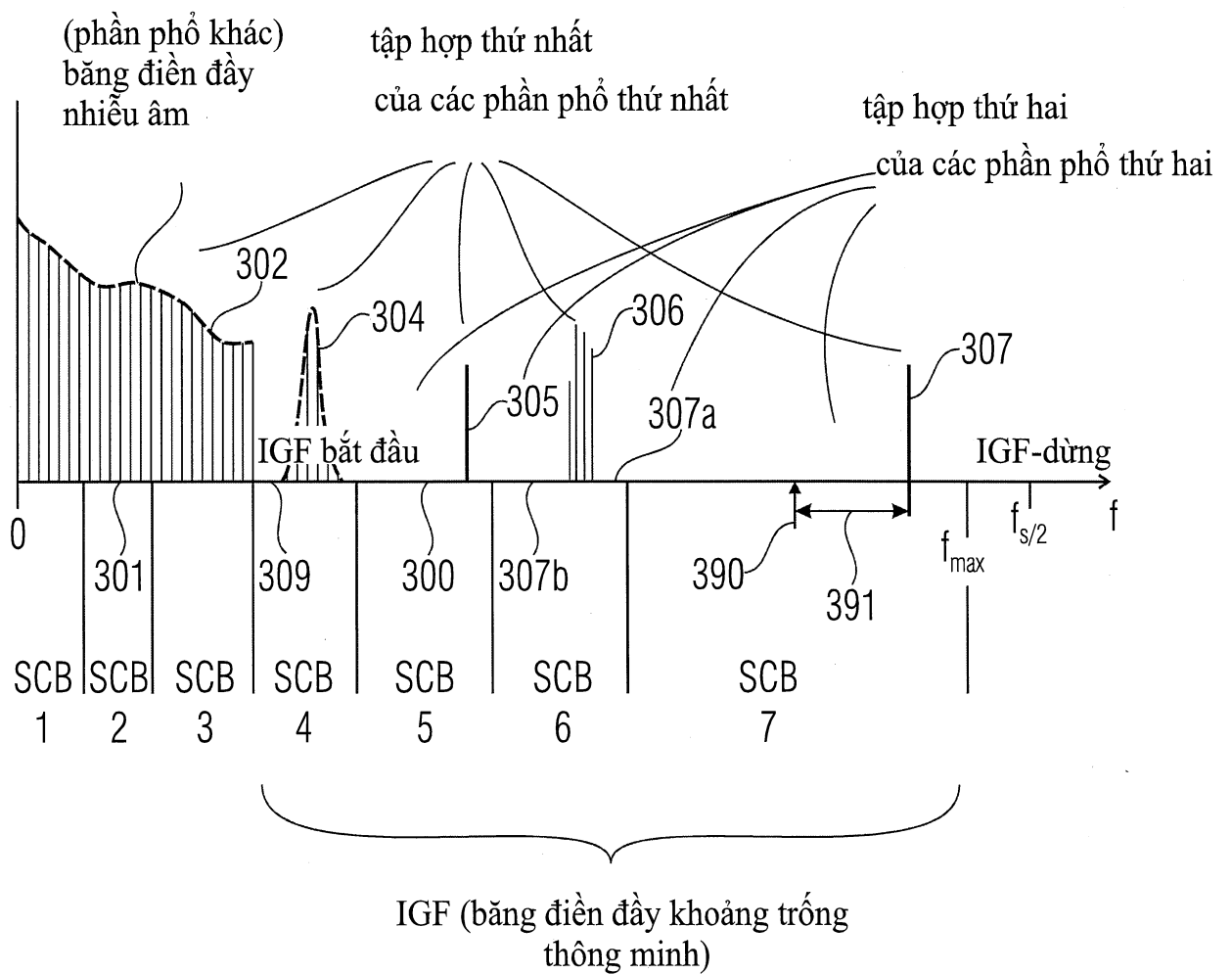


FIG 3A

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

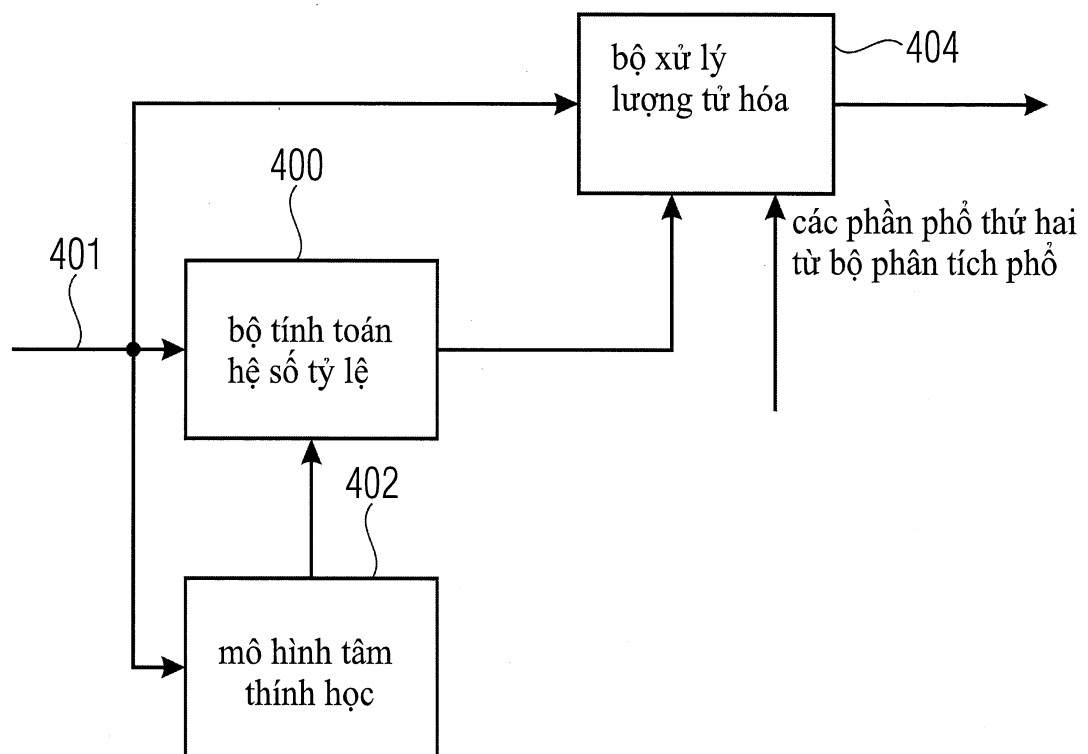


FIG 4A

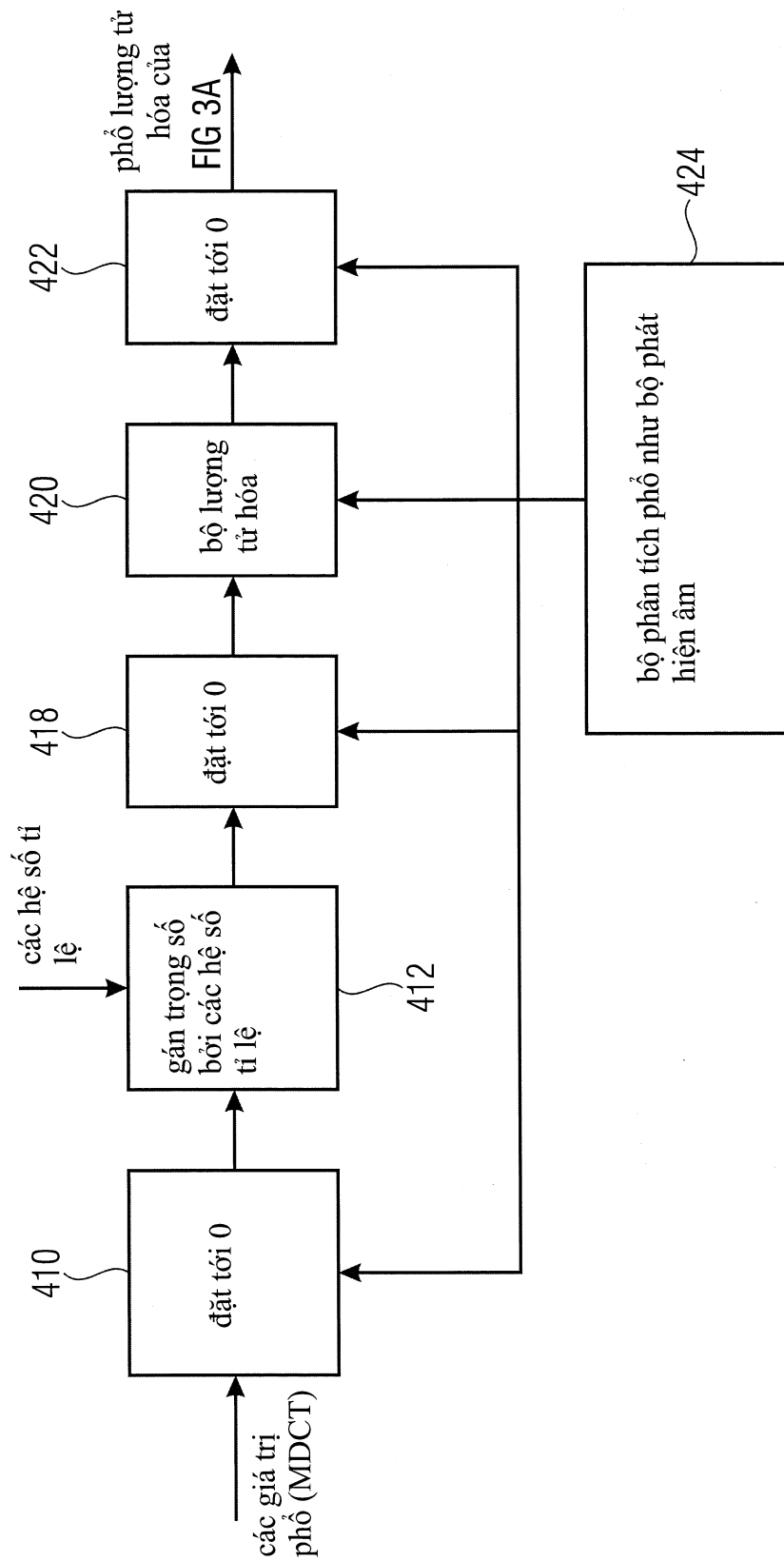


FIG 4B

(Bộ xử lý lượng tử hóa)

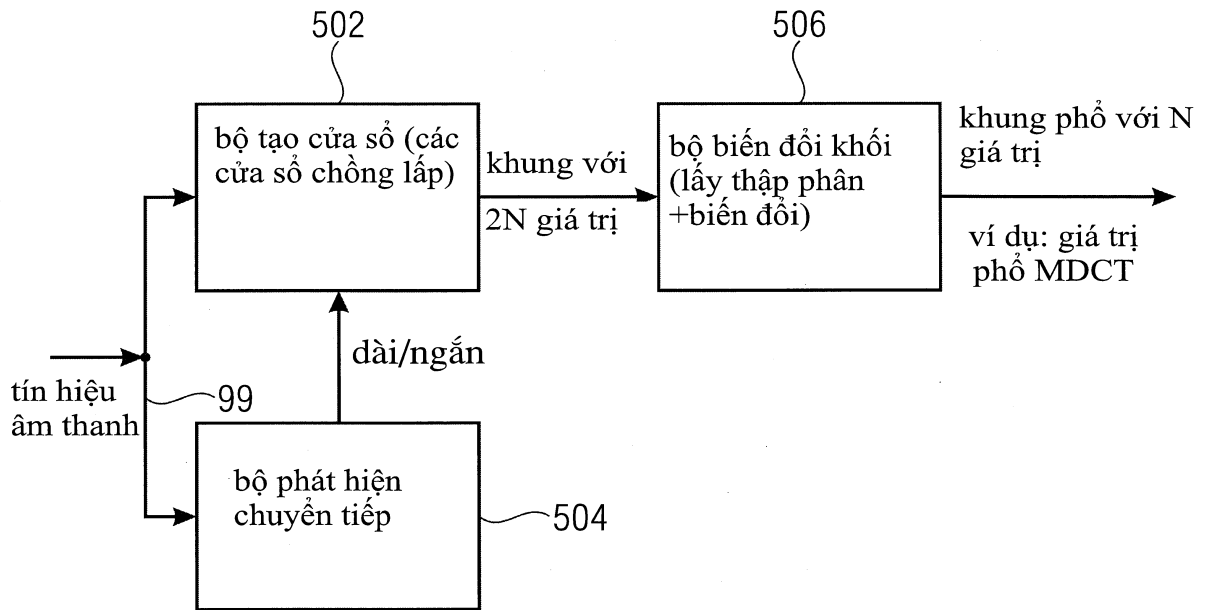


FIG 5A
(Các phần phổ khác)

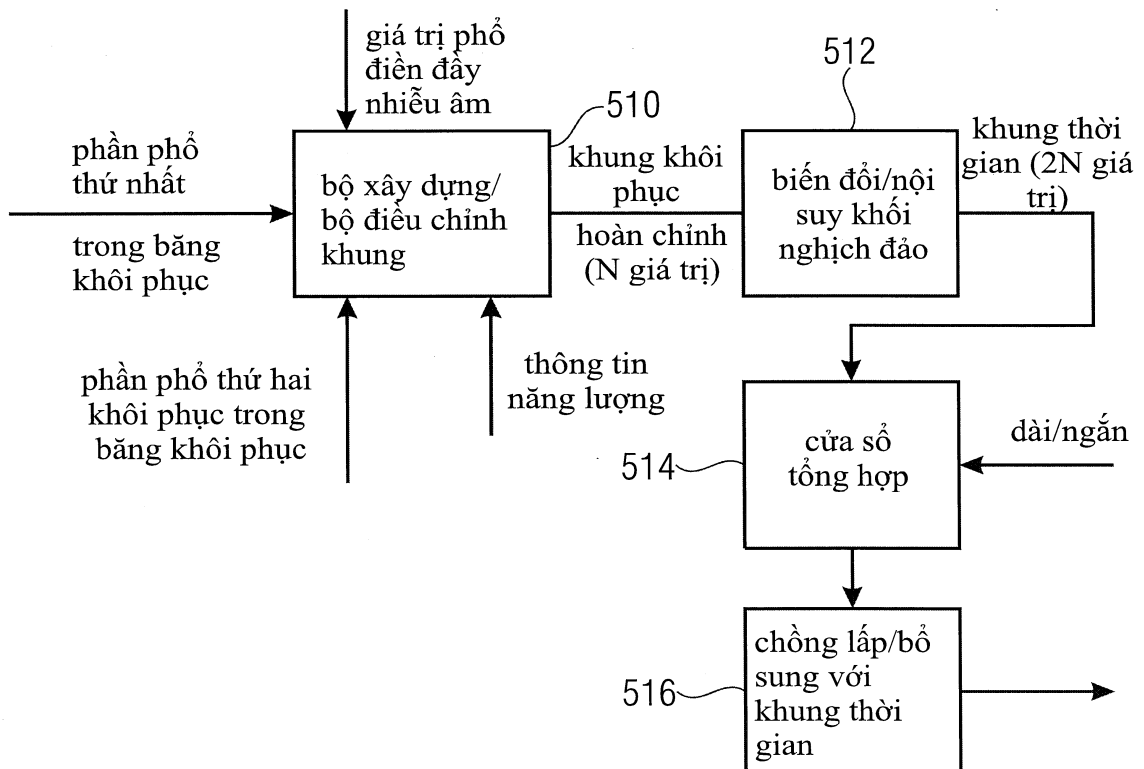


FIG 5B

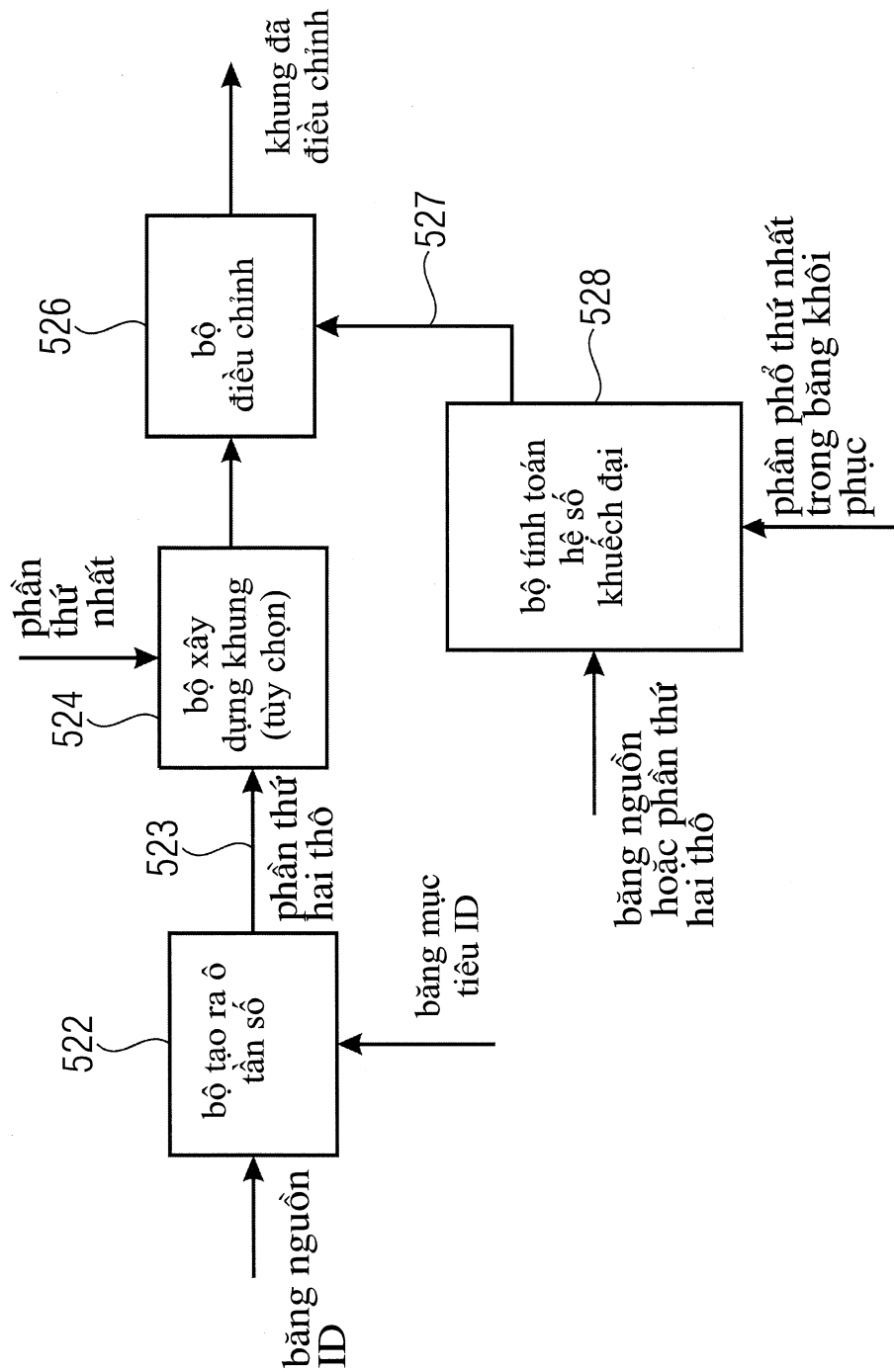


FIG 5C

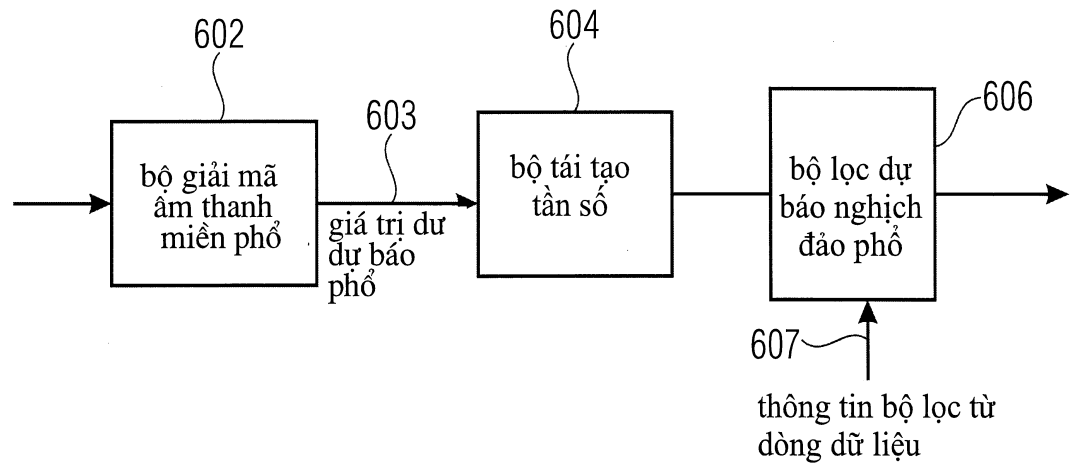


FIG 6A

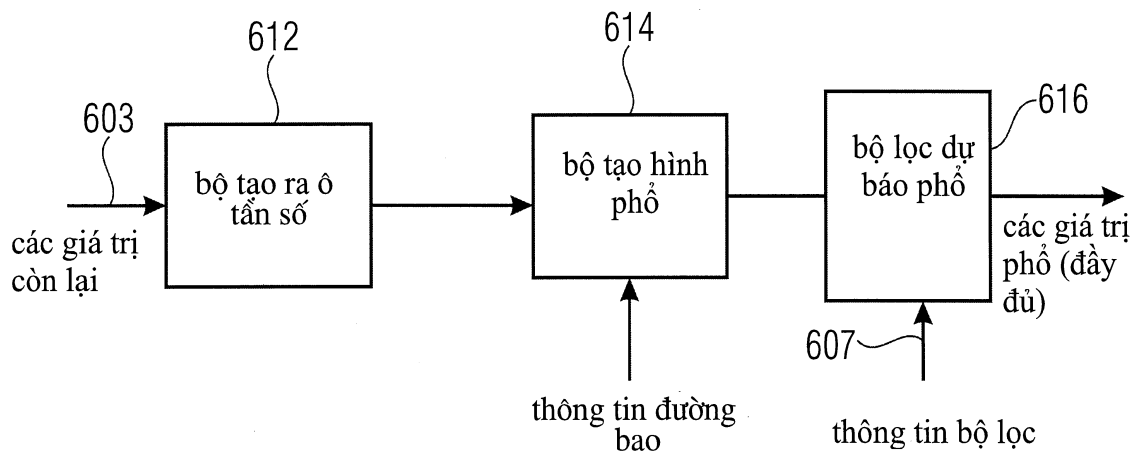


FIG 6B

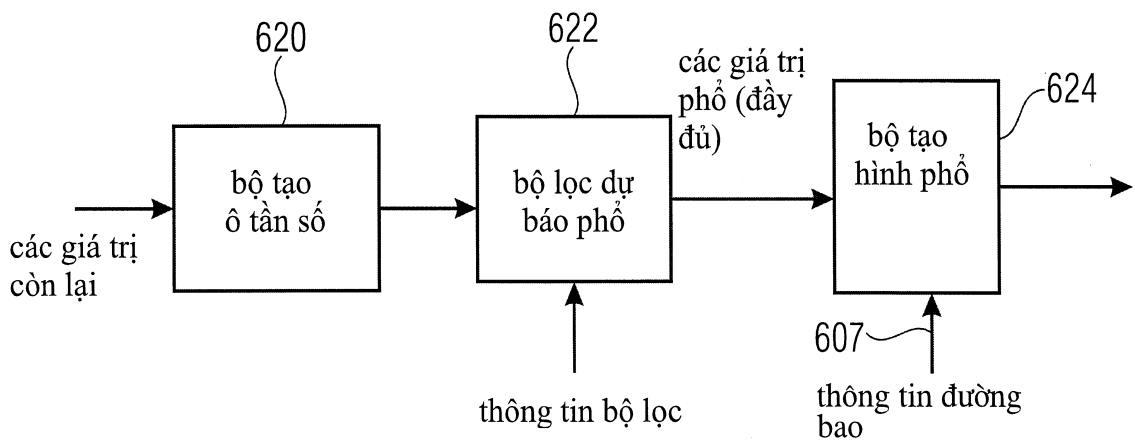


FIG 6C

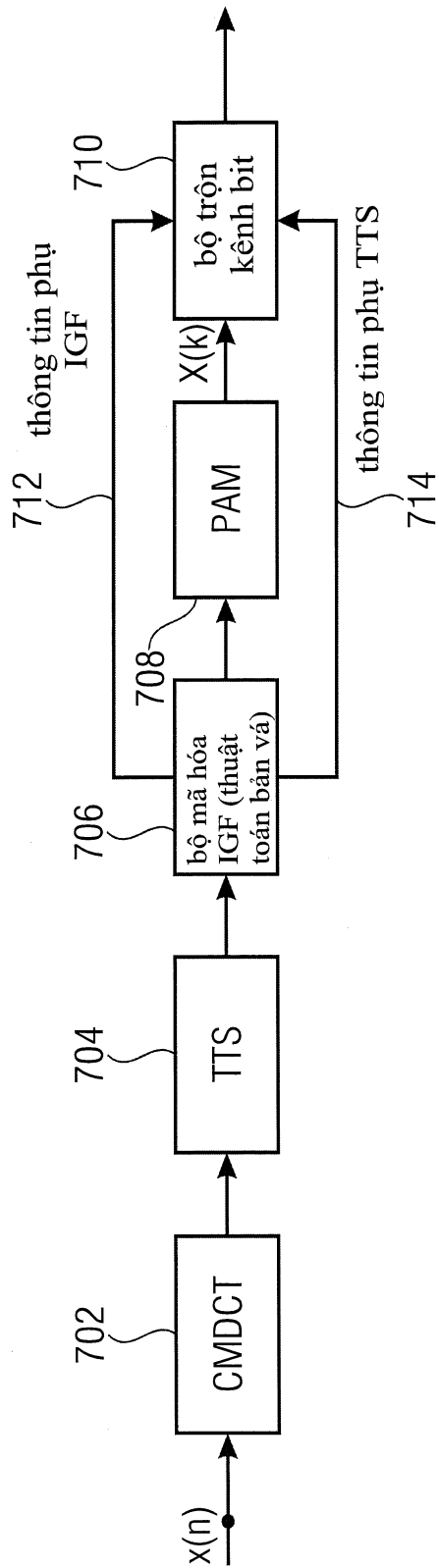


FIG 7A
(Bộ mã hóa)

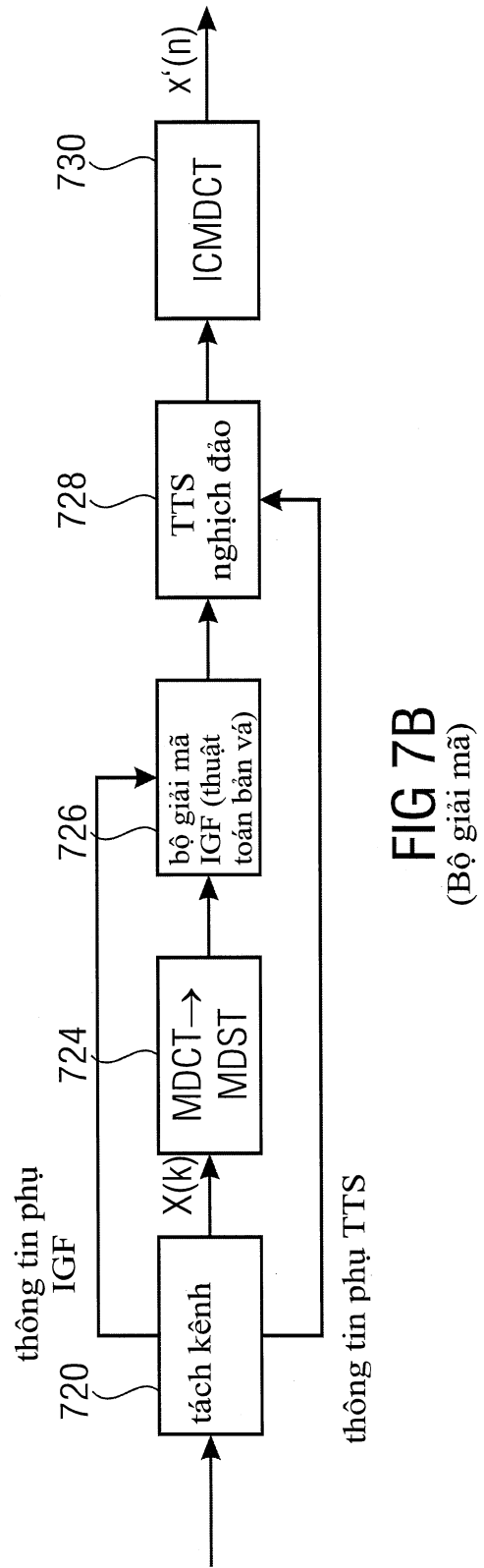
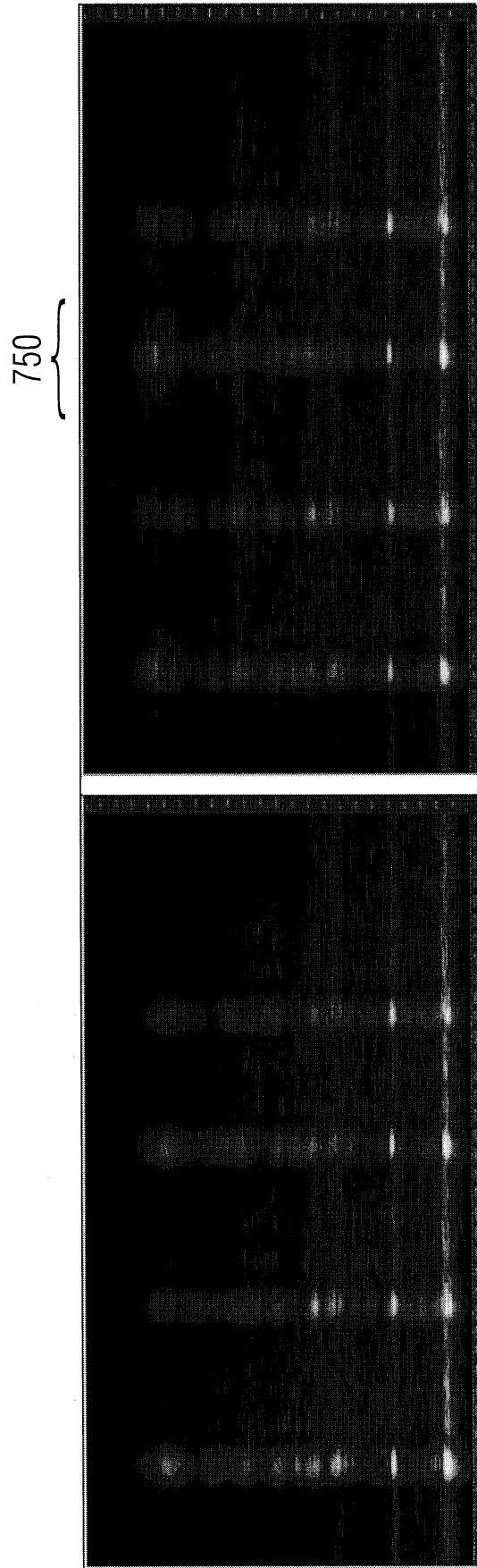
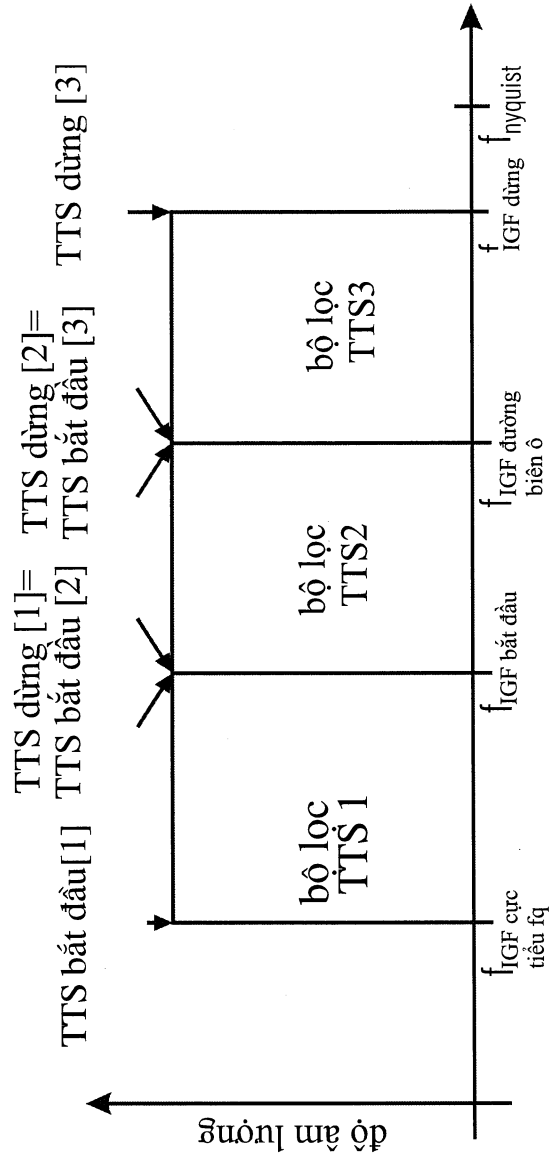


FIG 7B
(Bộ giải mã)



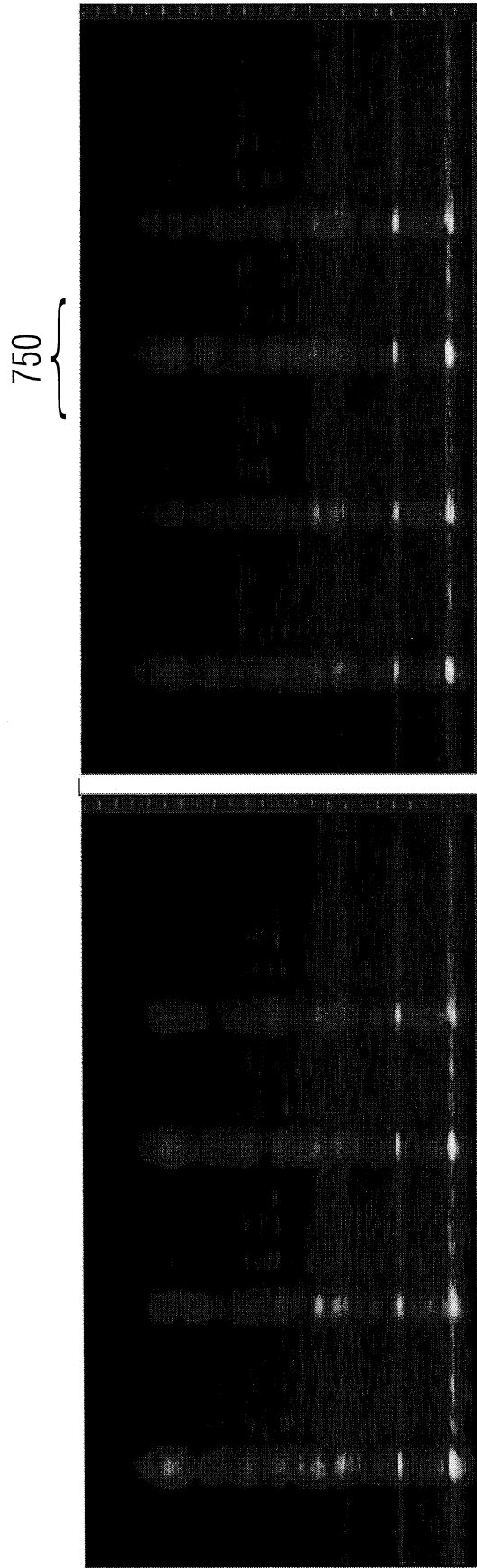
Ảnh phổ DFT của tín hiệu ban đầu (hình bên trái);
ảnh phổ DFT của tín hiệu mở rộng không có TTS (hình bên phải)

FIG 7C



Sơ đồ bộ lọc TTS cơ bản. Tầng hợp đánh dấu vùng phổ được che phủ bởi bộ lọc cụ thể

FIG 7D



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

FIG 7E

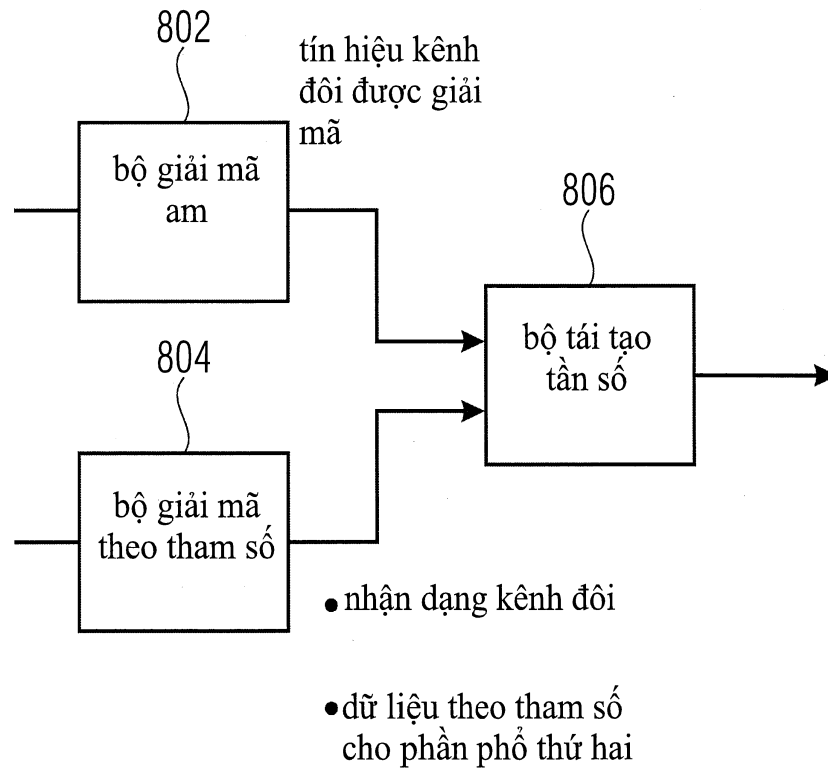


FIG 8A

khoảng nguồn	1	1	2	2	1: Phép biểu diễn kênh đôi thứ nhất (ví dụ L/R) 2: Phép biểu diễn kênh đôi thứ hai (ví dụ M/S)
khoảng đích	1	2	1	2	

FIG 8B

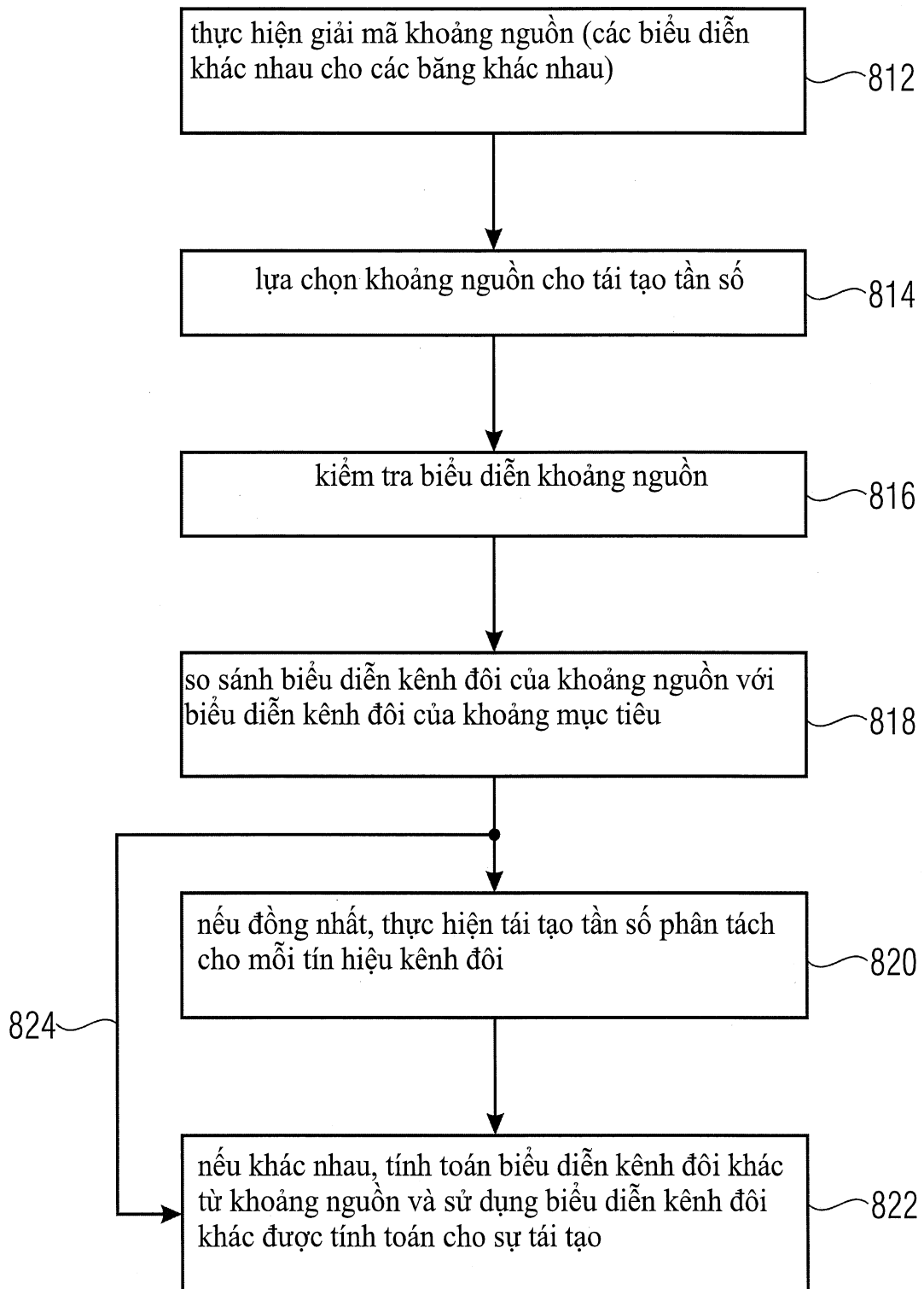
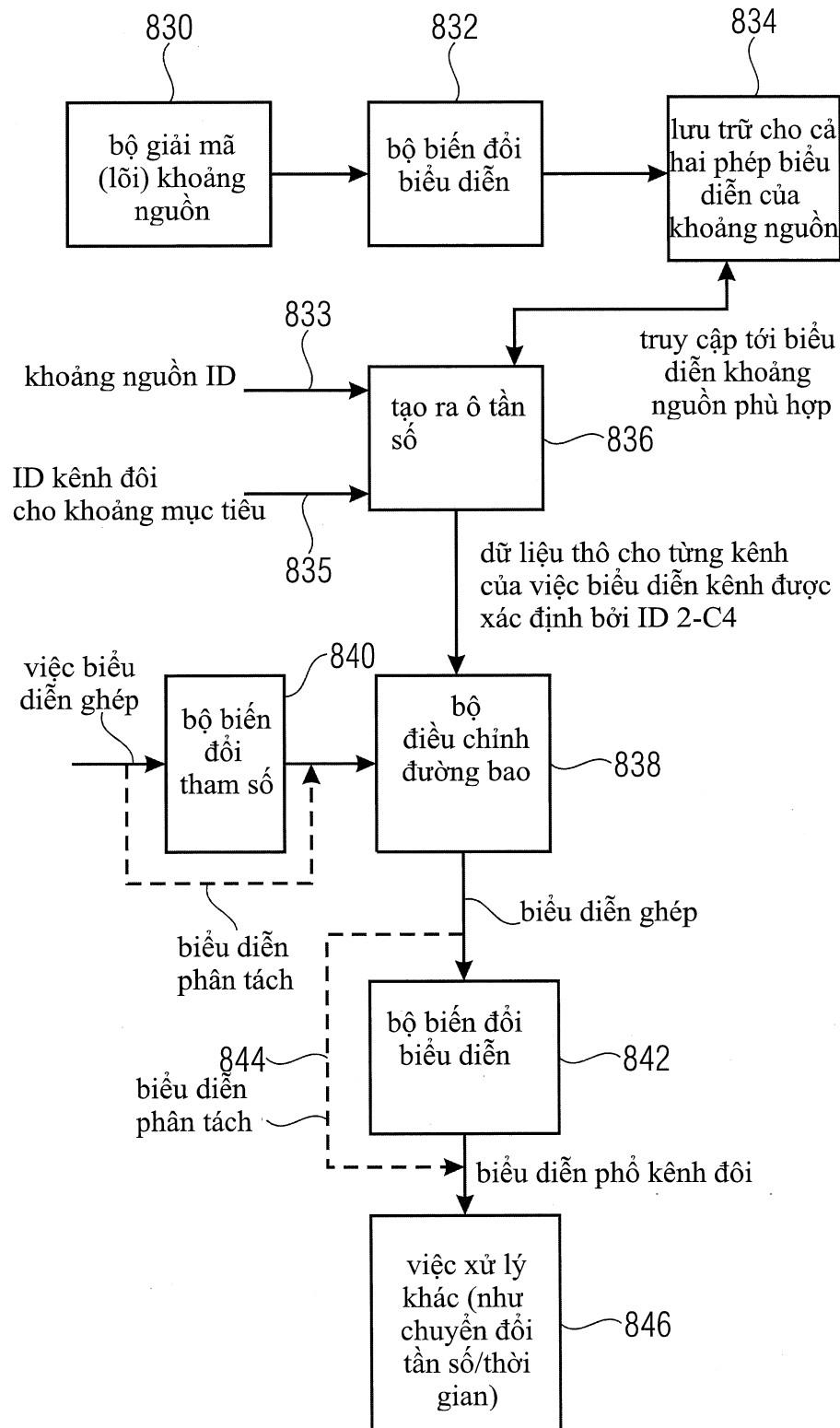


FIG 8C



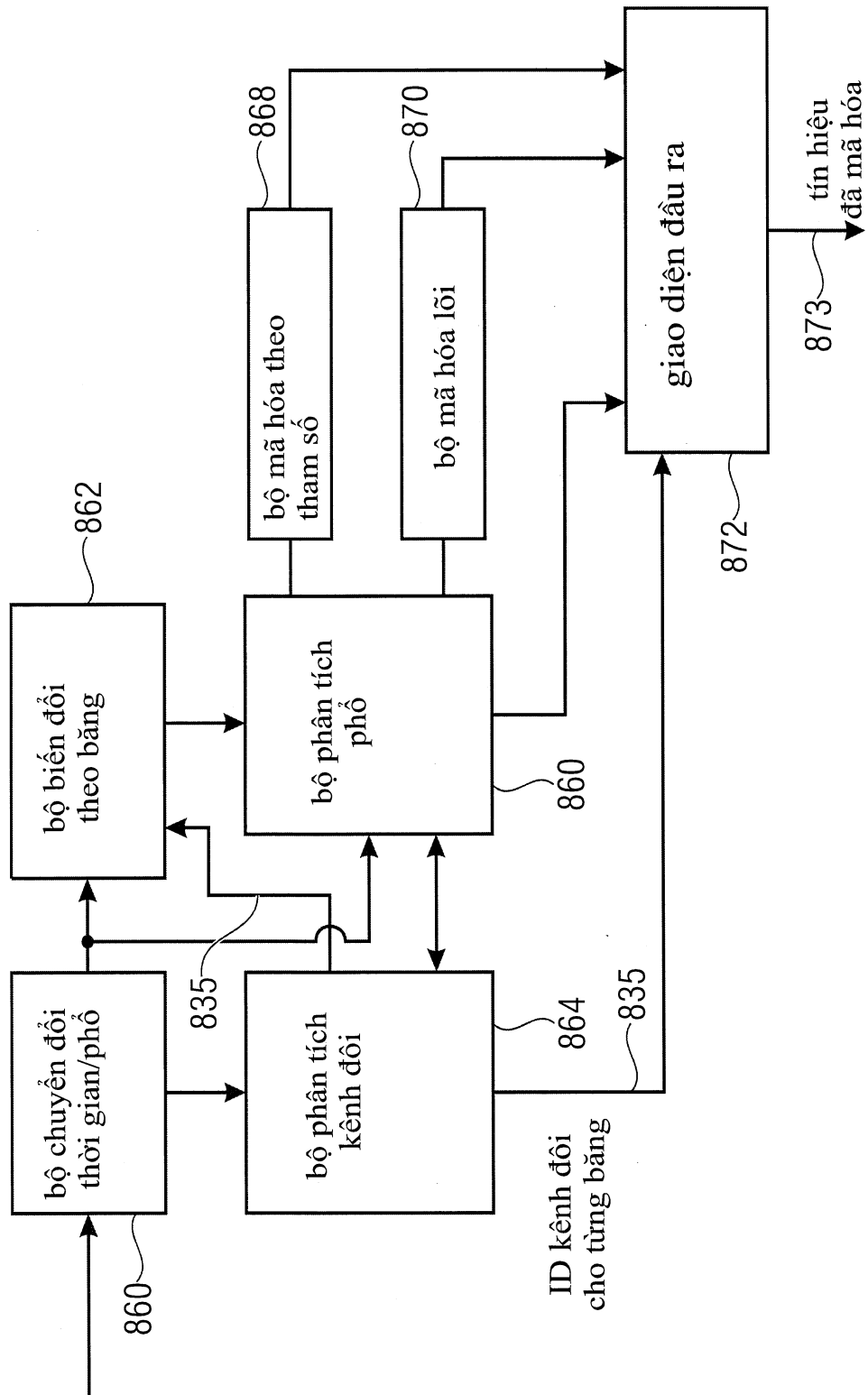


FIG 8E

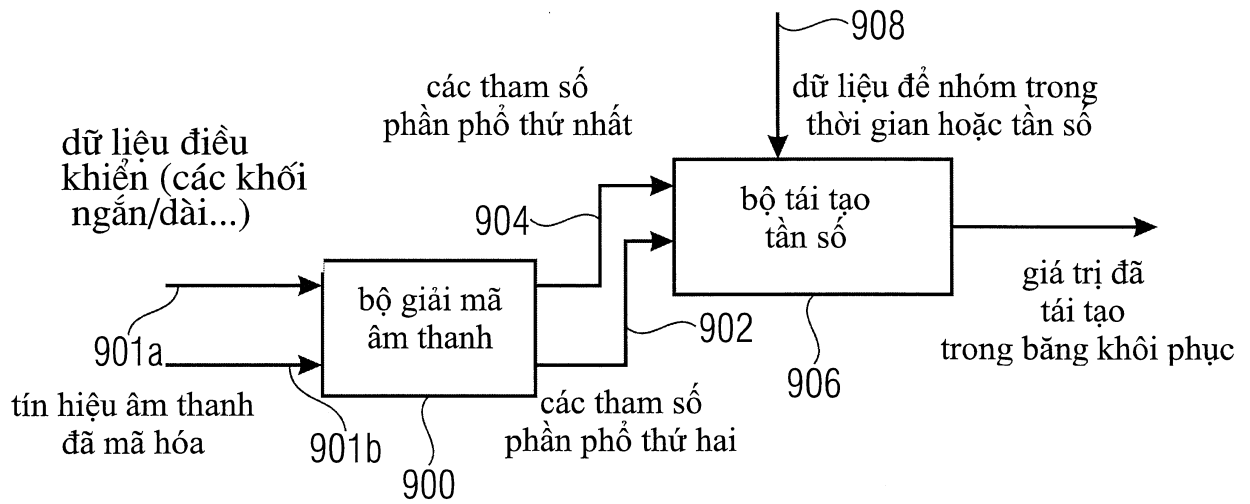


FIG 9A

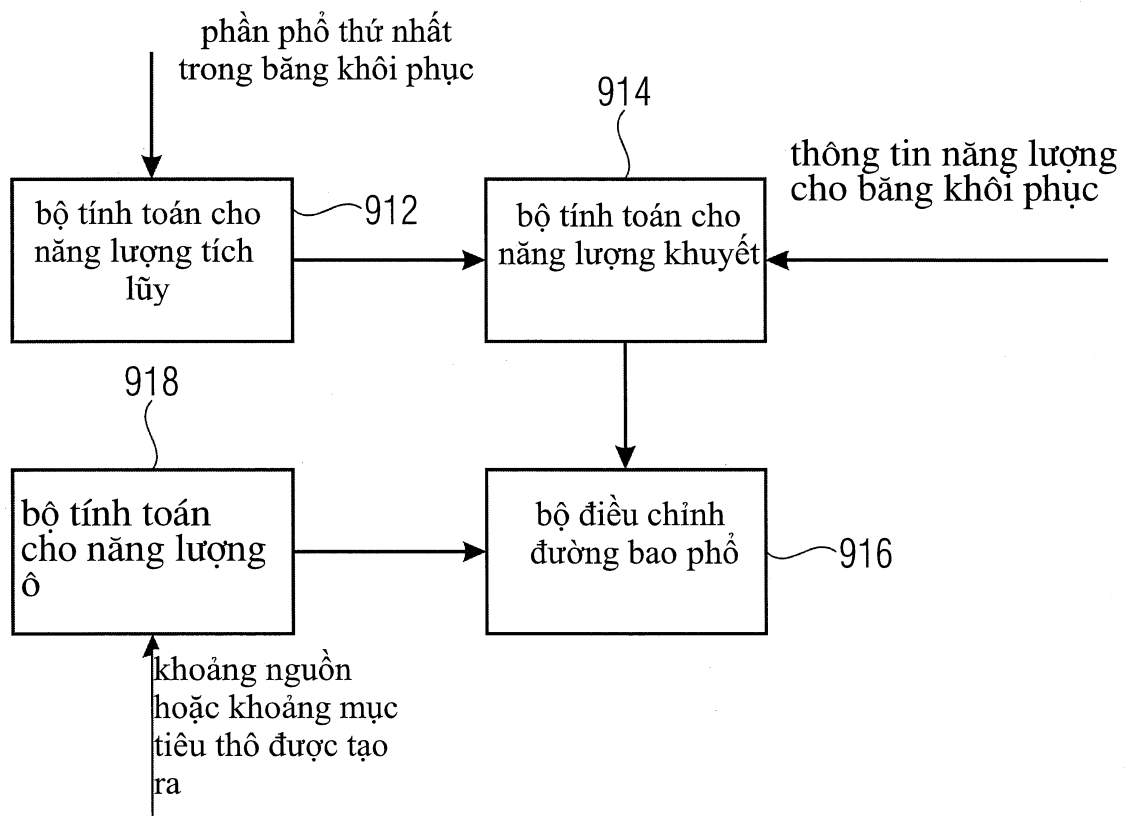
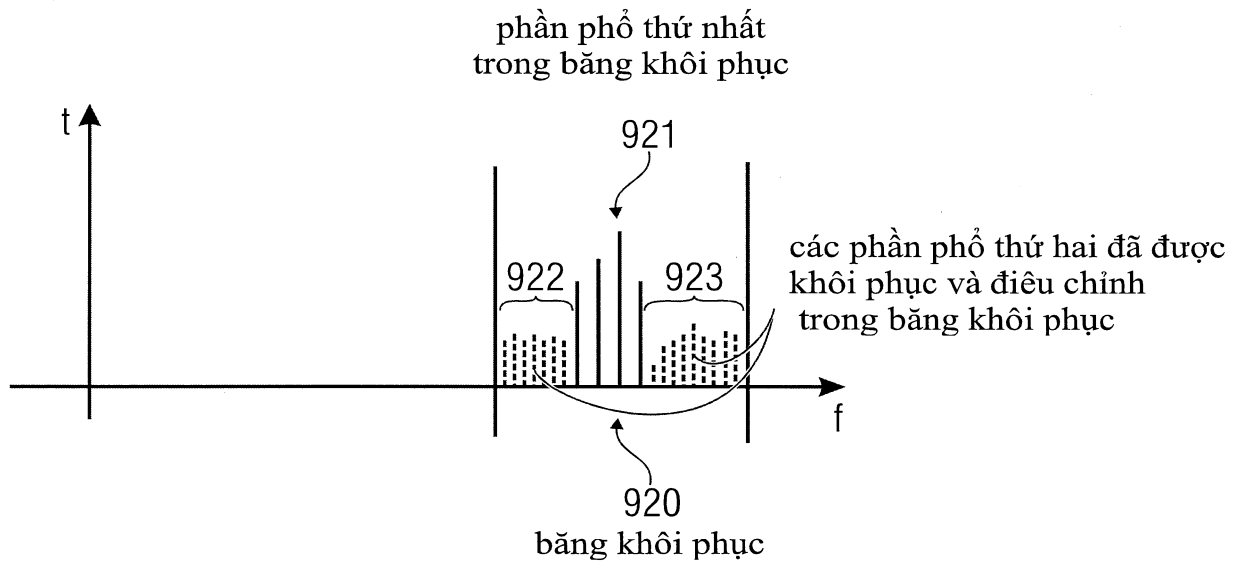


FIG 9B



- vd
- năng lượng tích lũy: 5 bộ
 - giá trị năng lượng cho băng khôi phục: 10 bộ (bao trùm phần phổ thứ nhất và thứ hai trong băng khôi phục)
 - năng lượng của dữ liệu khoảng nguồn hoặc dữ liệu khoảng mục tiêu: 8 bộ
 - năng lượng khuyết: 5 bộ
 - 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 được xác định bằng việc điều chỉnh đường bao

FIG 9C

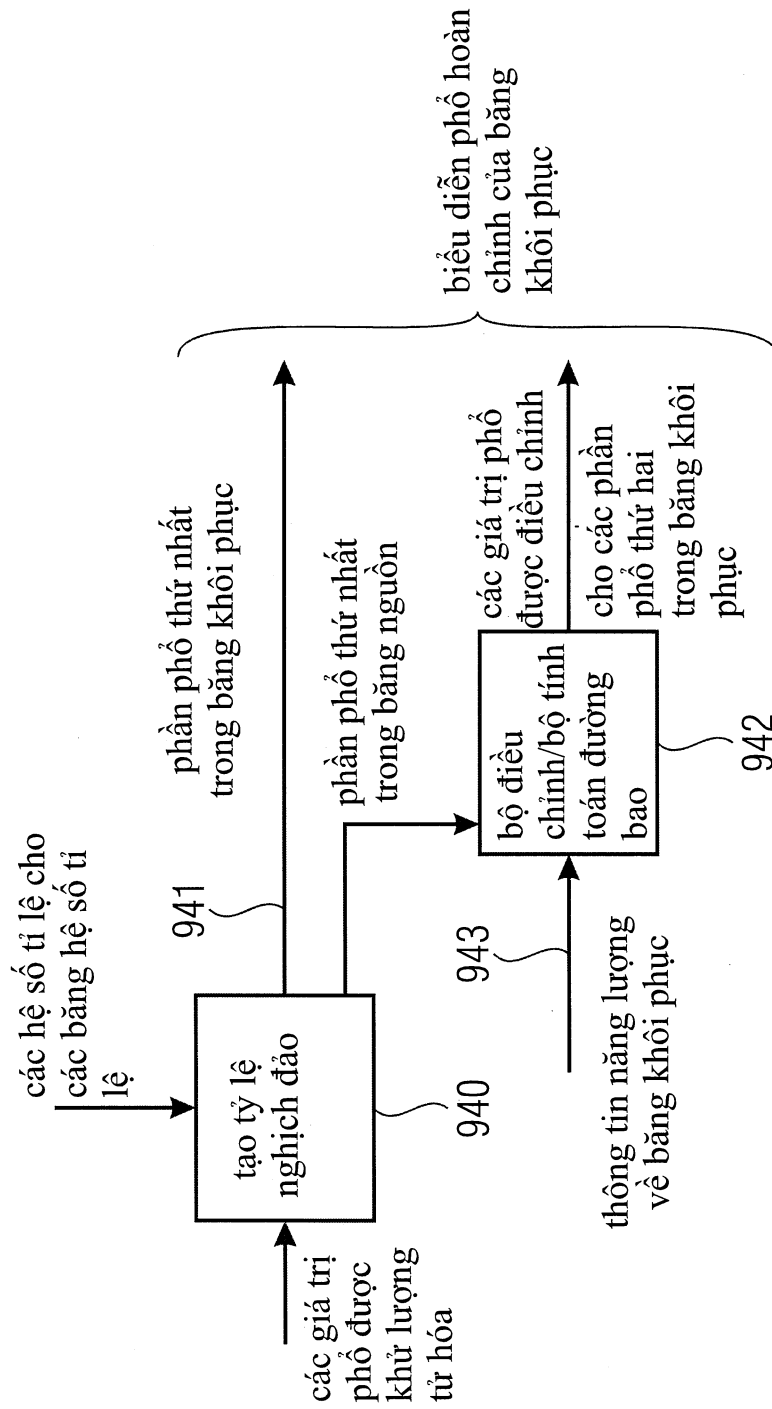


FIG 9D

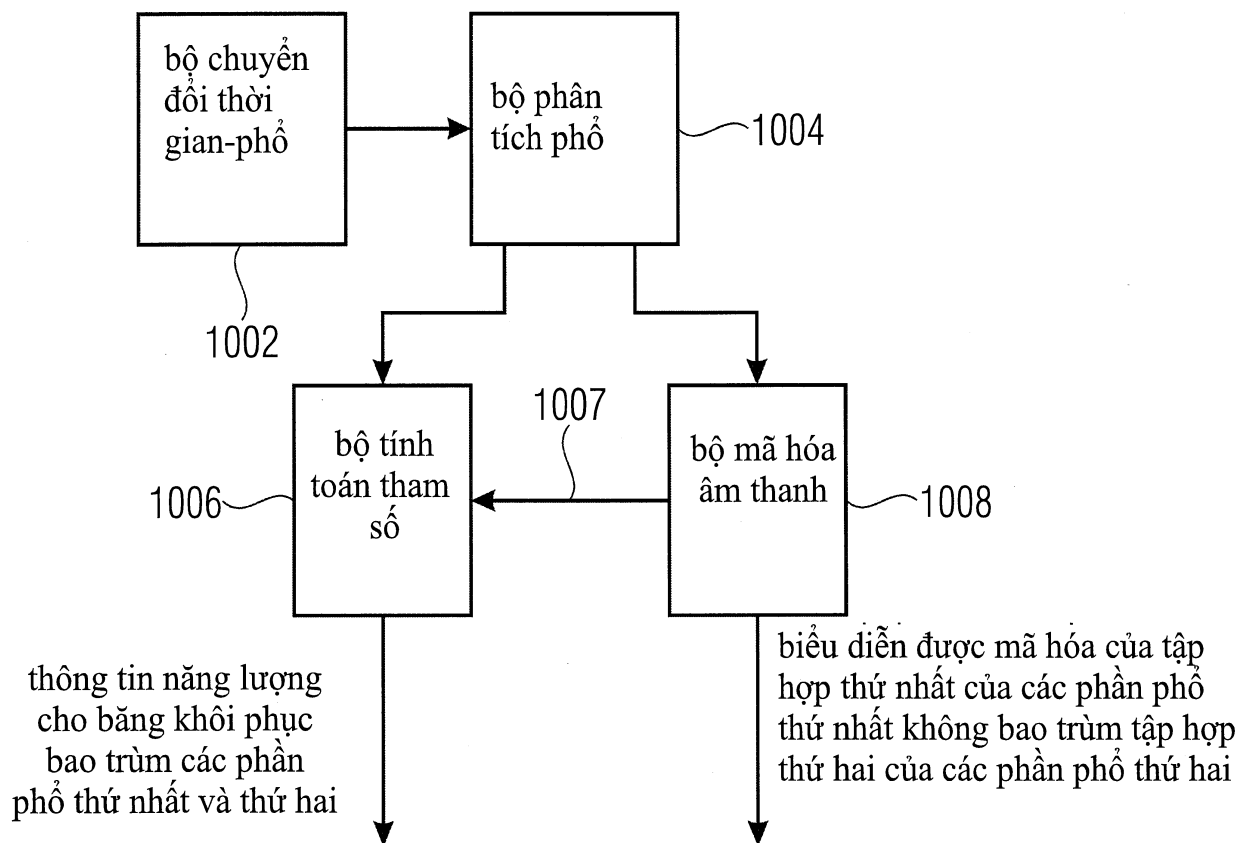


FIG 10A

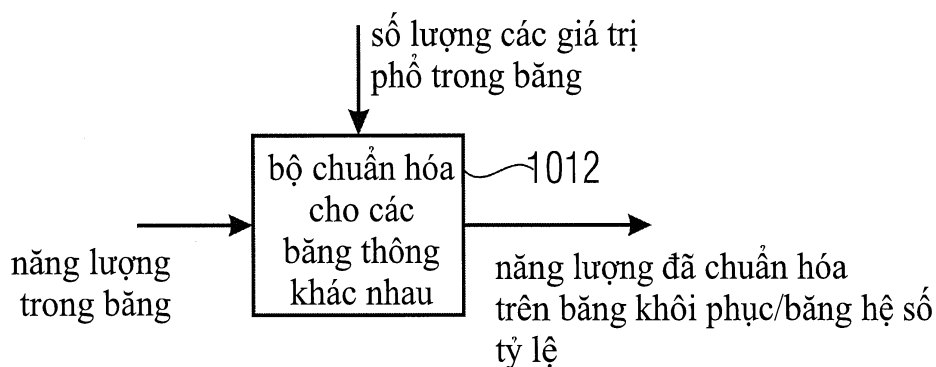


FIG 10B

thông tin điều khiển từ bộ mã hóa lỗi

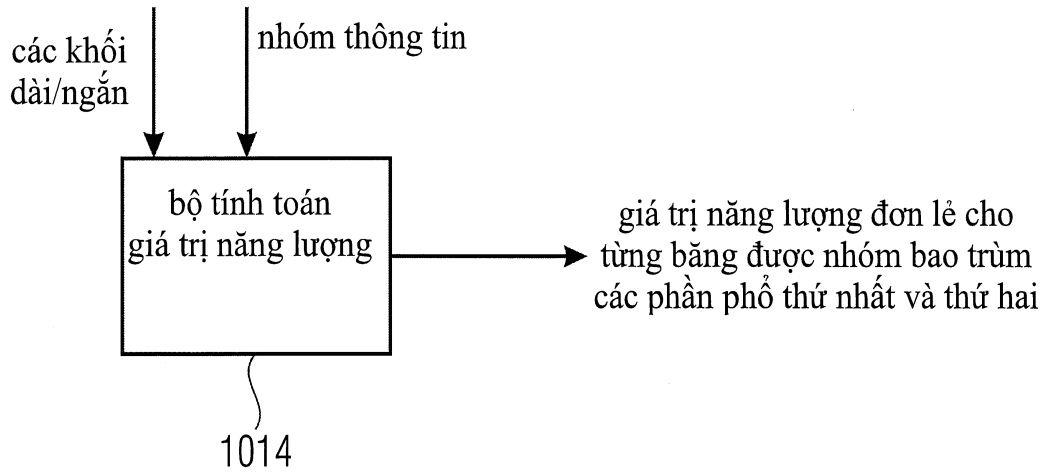


FIG 10C

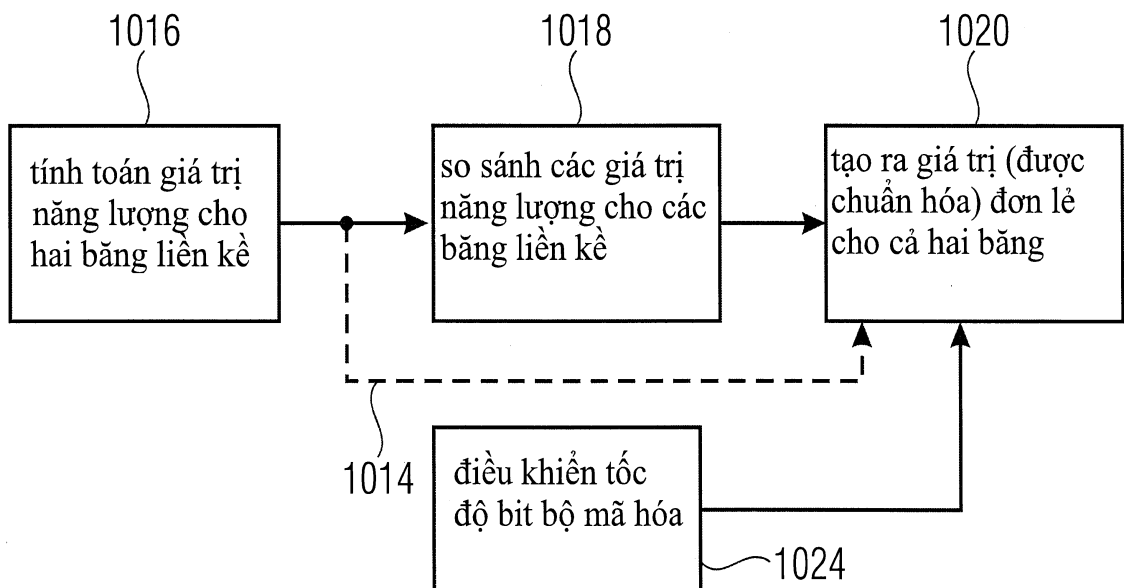


FIG 10D

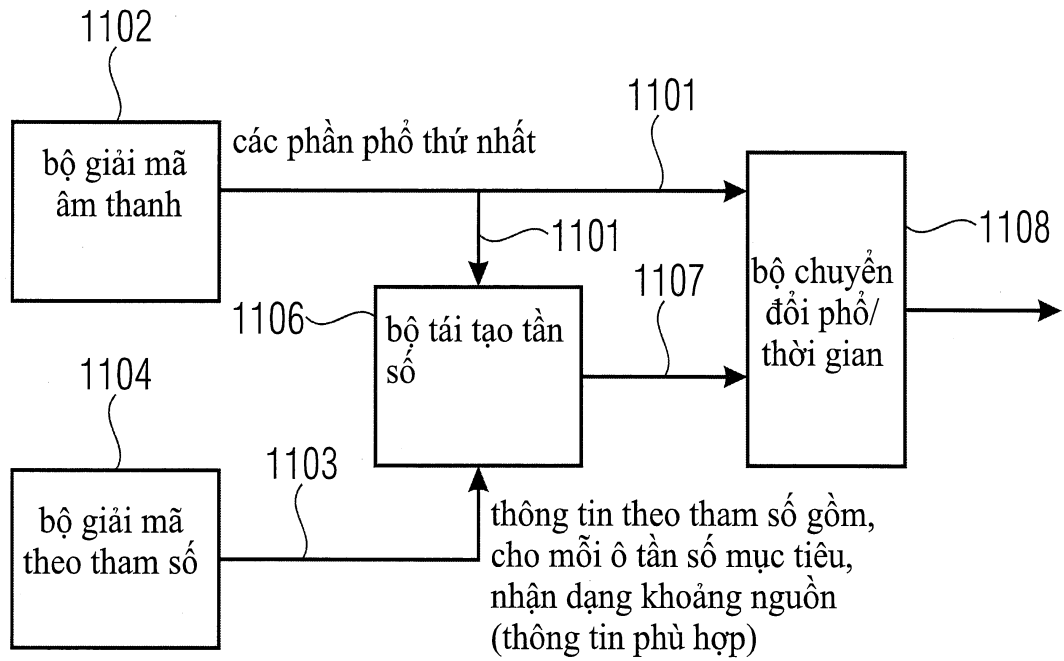


FIG 11A

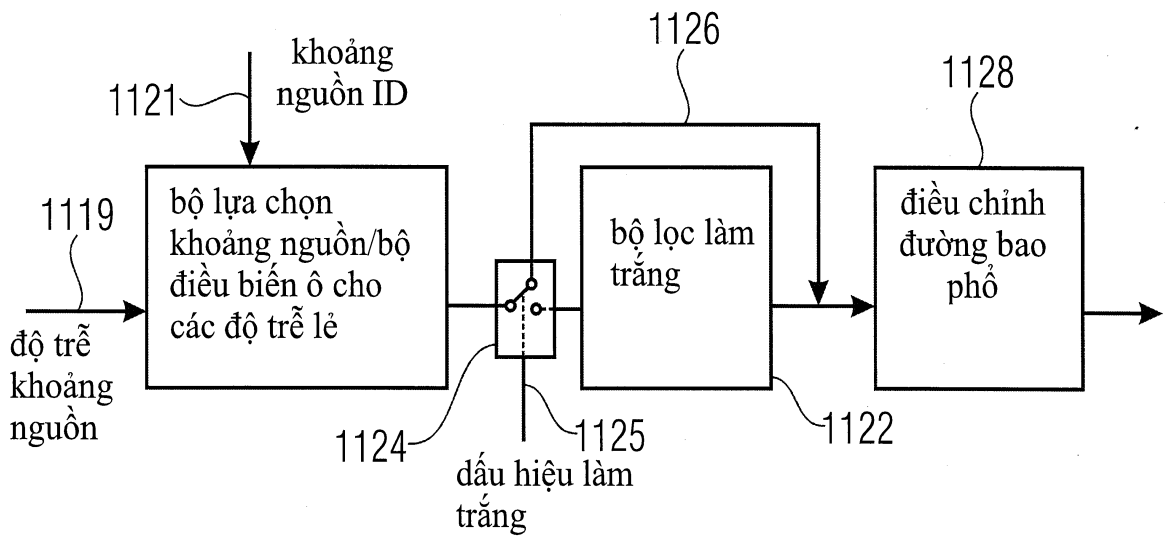


FIG 11B

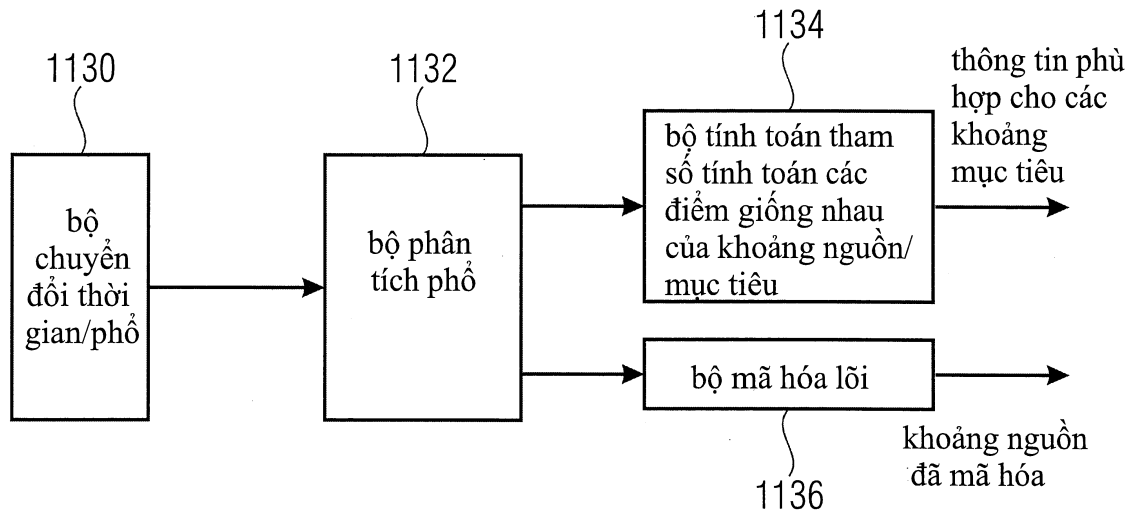


FIG 11C

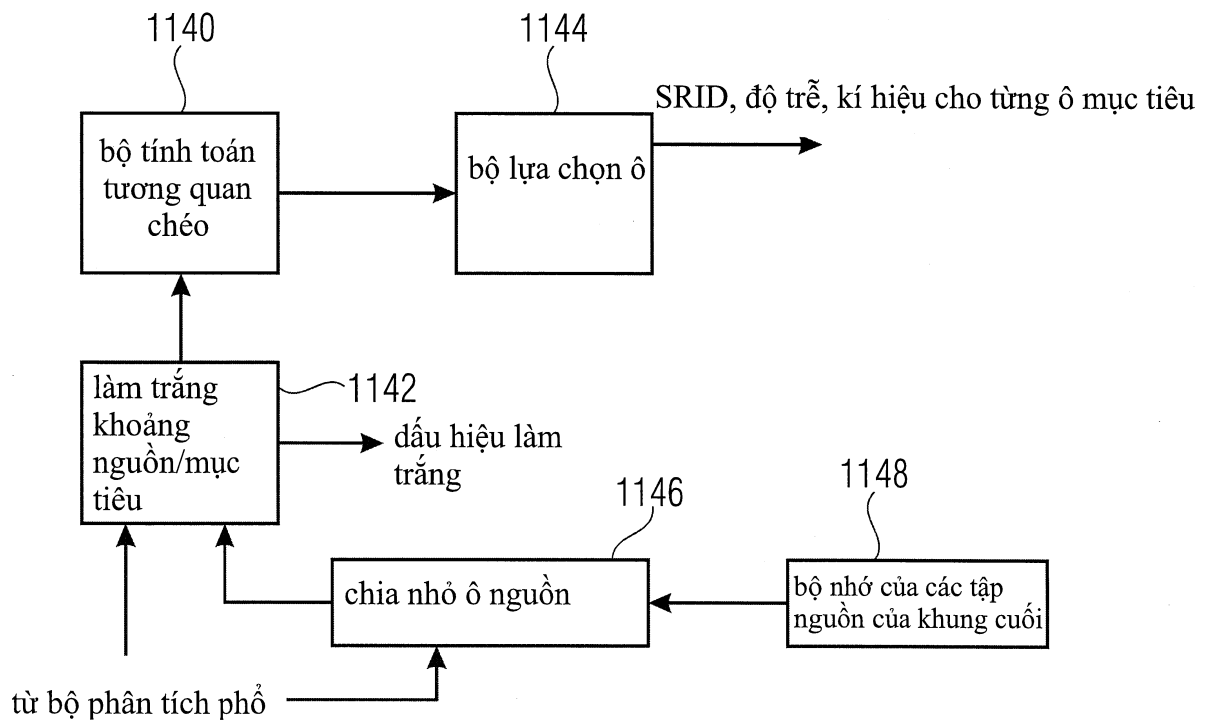


FIG 11D

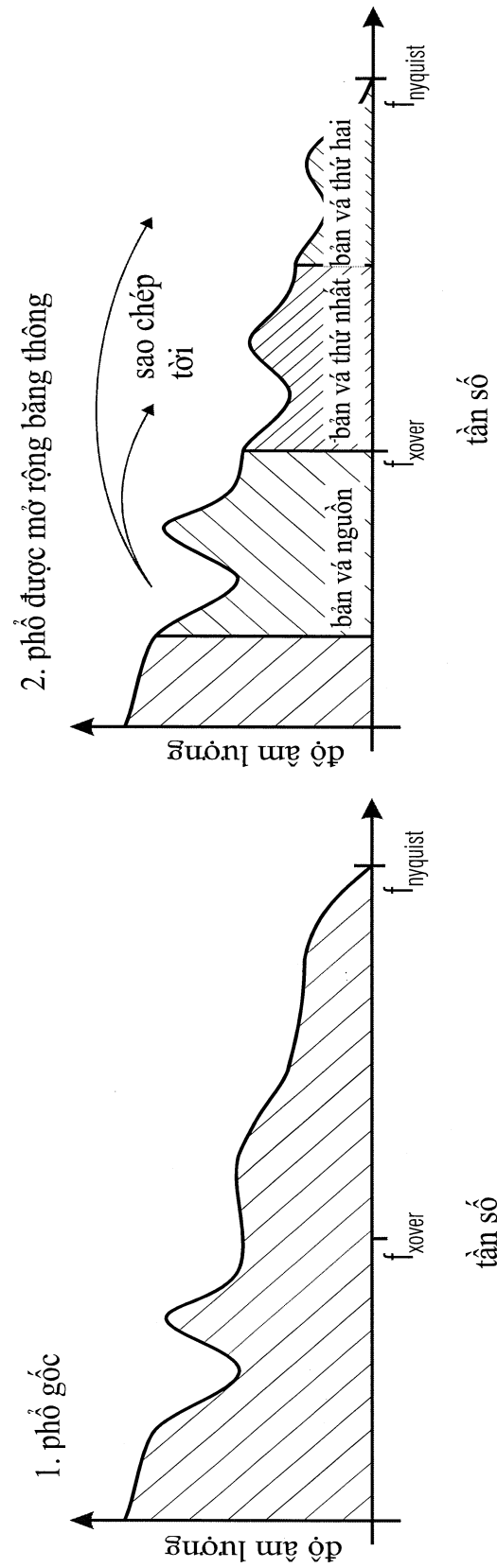


FIG 12A

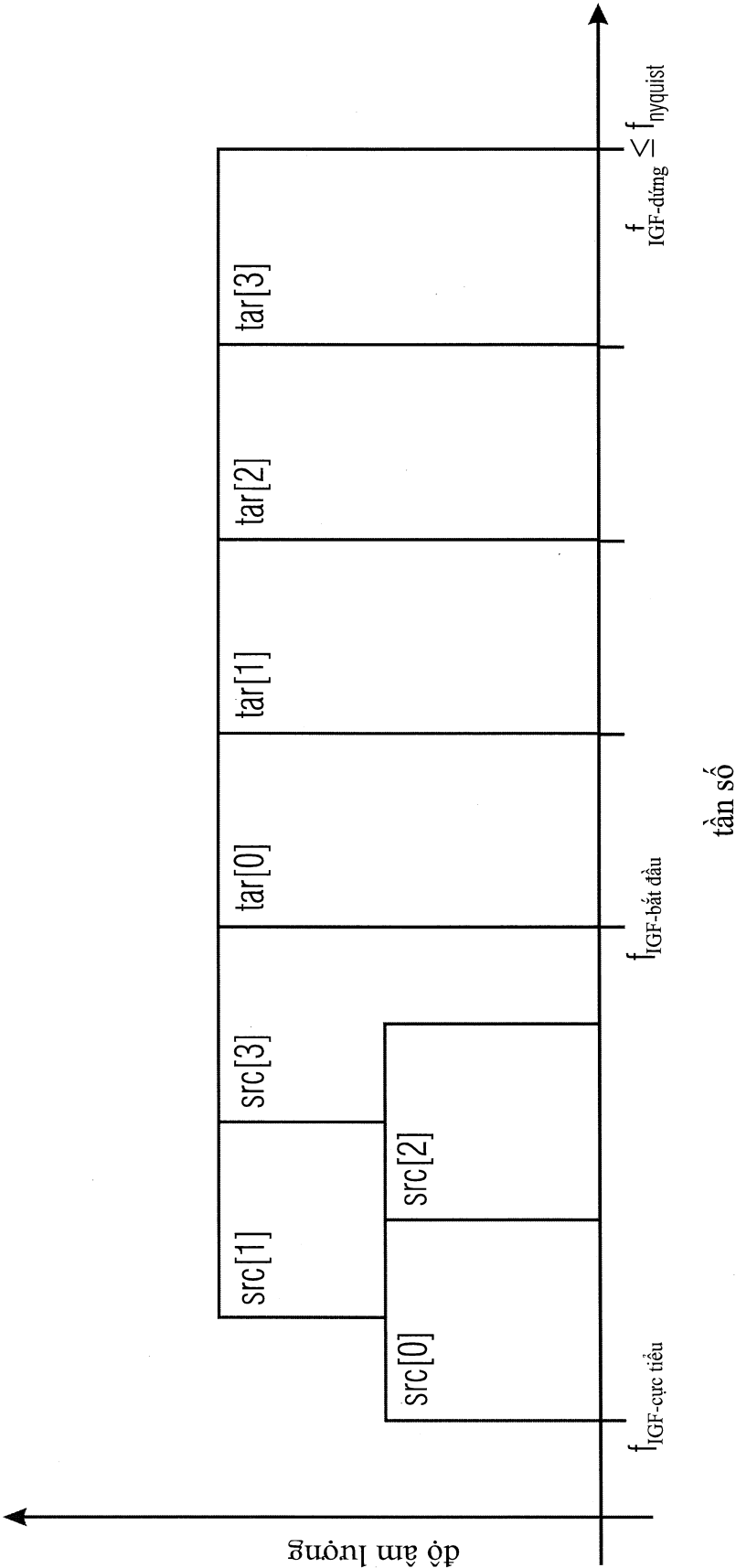


FIG 12B

ví dụ độ tương quan của hai tín hiệu

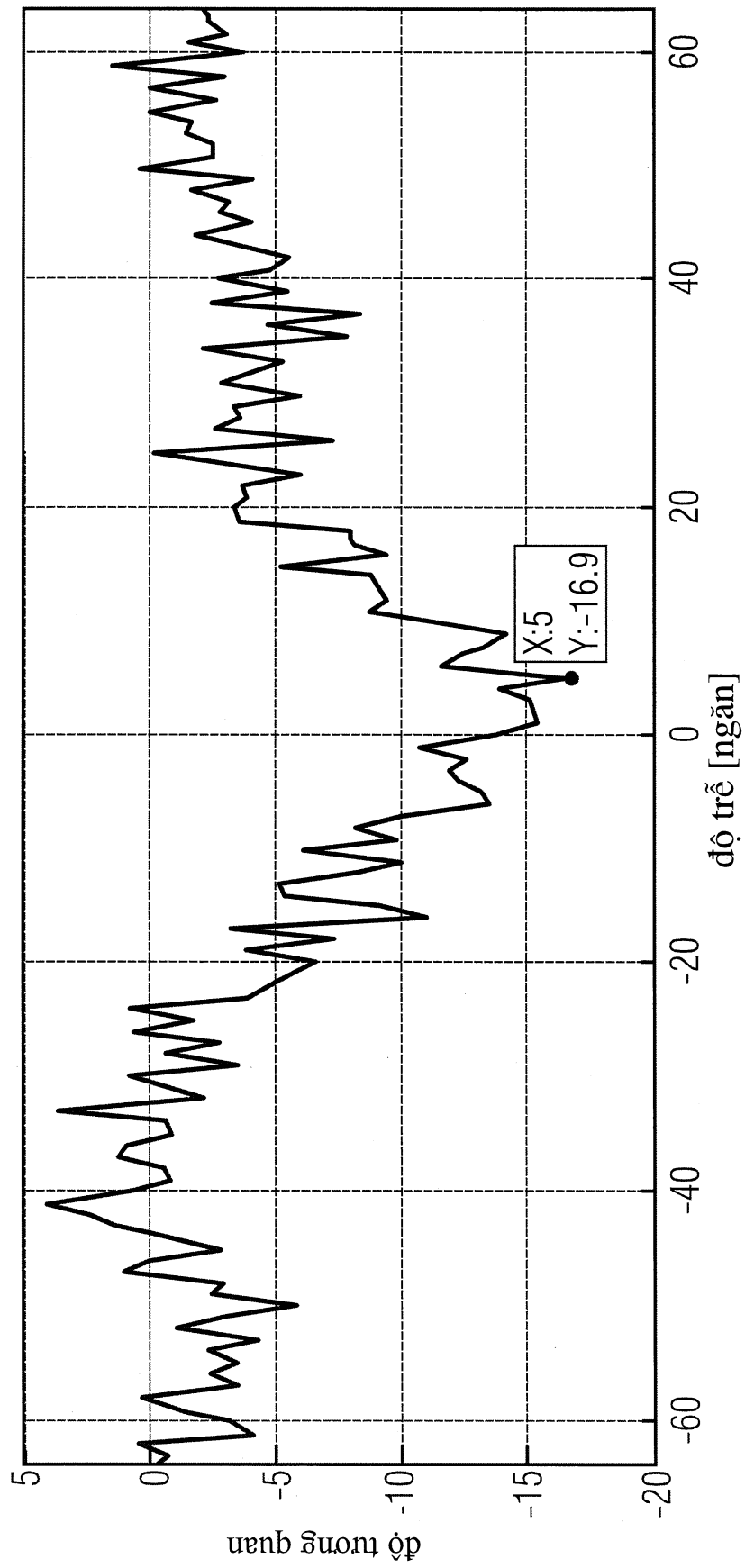


FIG 12C

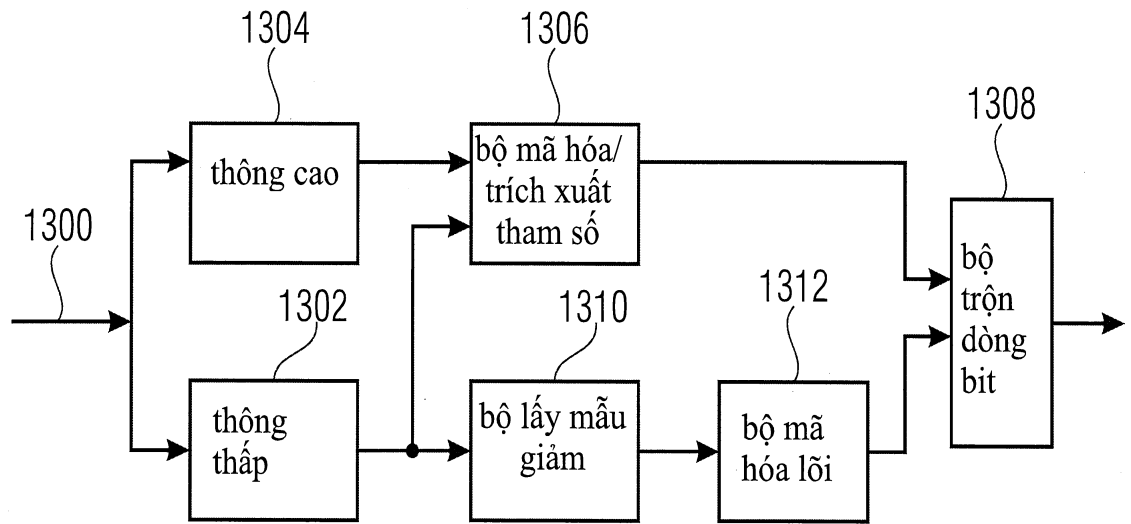


FIG 13A

(tình trạng kỹ thuật)

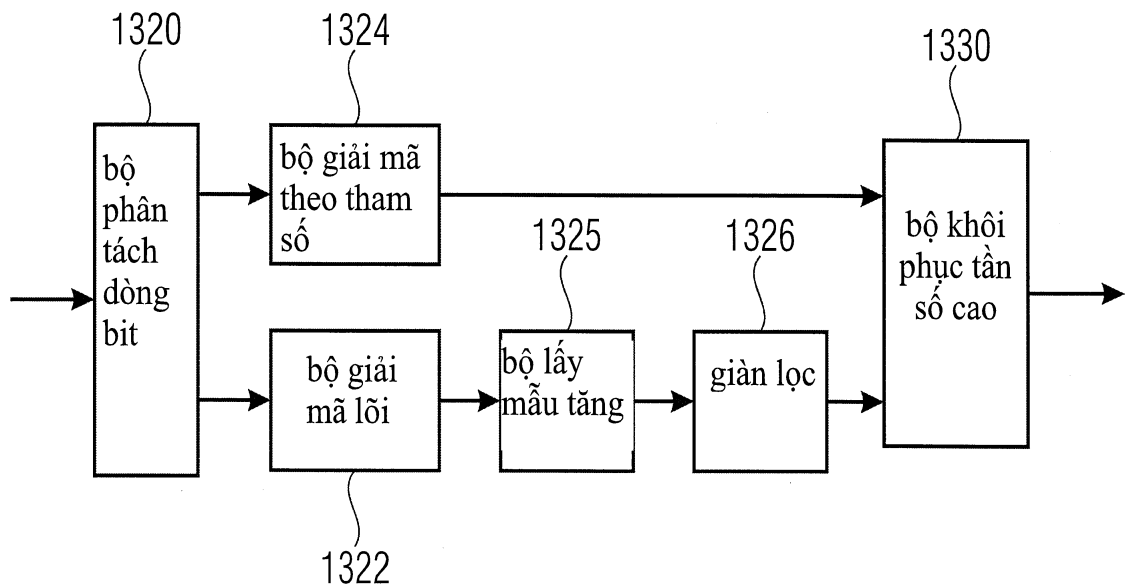


FIG 13B

(tình trạng kỹ thuật)