



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

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

(11)



1-0028922

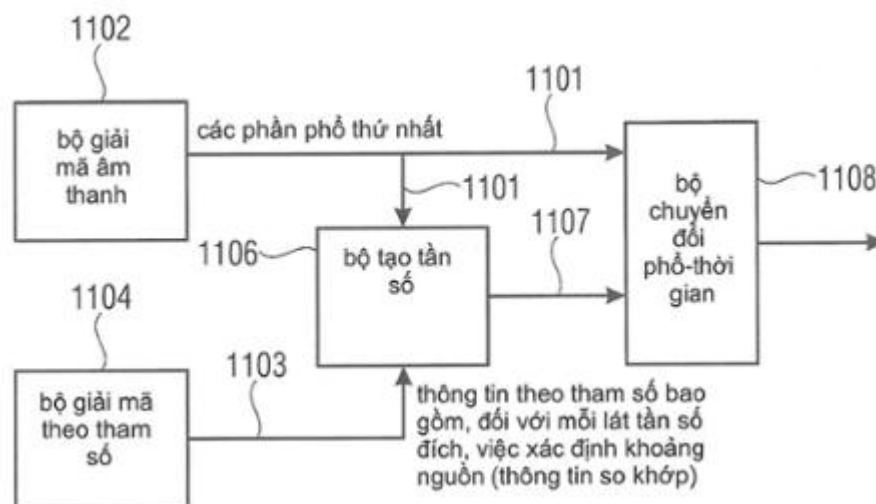
(51)<sup>7</sup> G10L 21/0388; G10L 19/02

(13) B

- (21) 1-2016-00588 (22) 15/07/2014  
(86) PCT/EP2014/065116 15/07/2014 (87) WO 2015/010952 29/01/2015  
(30) 13177346.7 22/07/2013 EP; 13177348.3 22/07/2013 EP; 13177350.9 22/07/2013 EP;  
13177353.3 22/07/2013 EP; 13189368.7 18/10/2013 EP  
(45) 25/07/2021 400 (43) 25/05/2016 338A  
(73) FRAUNHOFER-GESELLSCHAFT ZUR FOERDERUNG DER ANGEWANDTEN  
FORSCHUNG E.V. (DE)  
Hansastrasse 27c, 80686 Muenchen, Germany  
(72) NEUKAM, Christian (DE); DISCH, Sascha (DE); NAGEL, Frederik (DE);  
NIEDERMEIER, Andreas (DE); SCHMIDT, Konstantin (DE); THOSKHAHA,  
Balaji Nagendran (DE).  
(74) CÔNG TY LUẬT TRÁCH NHIỆM HỮU HẠN AMBYS HÀ NỘI (AMBYS  
HANOI)

(54) THIẾT BỊ VÀ PHƯƠNG PHÁP GIẢI MÃ VÀ MÃ HÓA TÍN HIỆU ÂM THANH

(57) Sáng chế đề cập đến thiết bị và phương pháp giải mã và mã hoá tín hiệu âm thanh sử dụng lựa chọn ô phổ thích ứng. Thiết bị giải mã tín hiệu được mã hoá, bao gồm: bộ giải mã âm thanh (1102) để giải mã phép biểu diễn được mã hoá của tập hợp thứ nhất của các phần phổ thứ nhất để thu được tập hợp thứ nhất được giải mã của các phần phổ thứ nhất (1101); bộ giải mã theo tham số (1104) để giải mã phép biểu diễn theo tham số được mã hoá của tập hợp thứ hai của các phần phổ thứ hai để thu được phép biểu diễn được giải mã của phép biểu diễn theo tham số (1103), trong đó thông tin theo tham số bao gồm, đối với mỗi lát tần số đích, việc nhận biết vùng nguồn làm thông tin so khớp; và bộ tái tạo tần số (1106) để tái tạo lát tần số đích sử dụng vùng nguồn từ tập hợp thứ nhất của các phần phổ thứ nhất (1101) được nhận biết bởi thông tin so khớp.



### **Lĩnh vực kỹ thuật được đề cập**

Sáng chế đề cập đến việc mã hóa/giải mã âm thanh, và cụ thể là đề cập đến việc mã hóa âm thanh sử dụng sự điền đầy khoảng trống thông minh.

### **Tình trạng kỹ thuật của sáng chế**

Mã hóa âm thanh là miền nén tín hiệu xử lý việc khai thác tính dư và tính không thích hợp trong các tín hiệu âm thanh bằng cách sử dụng kiến thức tâm thính học. Ngày nay, các bộ mã hóa-giải mã âm thanh thường cần khoảng 60 kbps/kênh để mã hóa trong suốt theo cảm quan hầu hết loại tín hiệu âm thanh bất kỳ. Các bộ mã hóa-giải mã mới hơn nhằm mục tiêu làm giảm tốc độ bit mã hóa bằng cách khai thác những điểm giống nhau về phổ trong tín hiệu sử dụng các kỹ thuật như mở rộng băng thông (bandwidth extension - BWE). Sơ đồ BWE sử dụng tập hợp tham số tốc độ bit thấp để biểu diễn các thành phần tần số cao (high frequency - HF) của tín hiệu âm thanh. Phổ HF được điền đầy với nội dung phổ từ các vùng tần số thấp (low frequency - LF) và hình 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 này cho phép các bộ mã hóa-giải mã âm thanh giữ được chất lượng tốt ở tốc độ bit thấp thậm chí bằng khoảng 24 kbps/kênh.

Việc lưu trữ hoặc truyền các tín hiệu âm thanh thường phải trải qua các ràng buộc tốc độ bit nghiêm ngặt. Trước đây, các bộ mã hóa bắt buộc phải giảm mạnh băng thông âm thanh được truyền khi chỉ có tốc độ bit rất thấp đã sẵn có.

Ngày nay, các bộ mã hóa-giải mã âm thanh hiện đại có thể mã hóa các tín hiệu dải rộng bằng cách sử dụng phương pháp mở rộng băng thông (bandwidth extension - BWE) trong tài liệu tham khảo [1]. Các thuật toán này dựa trên phép biểu diễn theo tham số của nội dung tần số cao (HF) – mà được tạo ra từ phần tần số thấp (LF) được mã hóa dạng sóng của tín hiệu được giải mã bằng cách chuyển vào trong vùng phổ HF (“vá”) và áp dụng việc xử lý sau theo tham số. Trong các sơ đồ BWE, sự khôi phục của vùng phổ HF cao hơn tần số giao chéo đã cho thường dựa trên phép vá phổ. Thông

thường, vùng HF gồm có nhiều bản vá liền kề và mỗi bản vá trong số các bản vá này được bắt nguồn từ các vùng thông dải (band-pass - BP) của phổ LF thấp hơn tần số giao chéo đã cho. Các hệ thống theo tình trạng kỹ thuật thực hiện một cách hữu hiệu việc vá trong phép biểu diễn giàn lọc, ví dụ giàn lọc gương cầu phương (Quadrature Mirror Filterbank - QMF), bằng cách sao chép tập hợp các hệ số băng con liền kề từ vùng nguồn đến vùng đích.

Một kỹ thuật khác được biết đến trong các bộ mã hóa-giải mã âm thanh ngày nay mà làm tăng hiệu quả nén và nhờ đó cho phép băng thông âm thanh mở rộng ở tốc độ bit thấp là sự thay thế tổng hợp theo tham số của các phần thích hợp của các phổ âm thanh. Ví dụ, các phần tín hiệu tương tự nhiễu âm của tín hiệu âm thanh ban đầu có thể được thay thế mà không bị hao tổn đáng kể đối với chất lượng đối tượng bởi nhiễu âm giả tạo được tạo ra trong bộ giải mã và được tạo tỷ lệ bởi các tham số thông tin phụ. Một ví dụ về công cụ thay thế nhiễu âm cảm quan (Perceptual Noise Substitution - PNS) có trong phép mã hóa âm thanh tiên tiến (Advanced Audio Coding - AAC) MPEG-4 trong tài liệu tham khảo [5].

Một đề xuất khác mà cũng cho phép băng thông âm thanh mở rộng ở tốc độ bit thấp là kỹ thuật điền đầy nhiễu âm được chứa trong mã hóa tiếng nói và âm thanh thống nhất (Unified Speech and Audio Coding - USAC) MPEG-D trong tài liệu tham khảo [7]. Các khoảng trống phổ (các điểm không) mà được suy ra bởi vùng chết của bộ lượng tử hoá do việc lượng tử hoá quá thô, tiếp theo được điền đầy với nhiễu âm giả trong bộ giải mã và được tạo tỷ lệ bằng việc xử lý sau theo tham số.

Một hệ thống theo tình trạng kỹ thuật khác được gọi là thay thế phổ chính xác (Accurate Spectral Replacement - ASR) trong tài liệu tham khảo [2-4]. Ngoài bộ mã hóa-giải mã dạng sóng, ASR dùng giai đoạn tổng hợp tín hiệu chuyên dụng mà phục hồi các phần hình sin quan trọng theo cảm quan của tín hiệu tại bộ giải mã. Ngoài ra, hệ thống được mô tả trong tài liệu tham khảo [5] dựa trên việc tạo mô hình hình sin trong vùng HF của bộ mã hóa dạng sóng để cho phép băng thông âm thanh mở rộng có chất lượng cảm quan tốt ở tốc độ bit thấp. Tất cả các phương pháp này bao gồm việc biến đổi dữ liệu thành phần miền thứ hai ngoài biến đổi cosin rời rạc cải biên

(Modified Discrete Cosine Transform - MDCT) và các giai đoạn phân tích/tổng hợp khá phức tạp để duy trì các thành phần hình sin HF.

Fig.13a minh họa biểu đồ dưới dạng sơ đồ của bộ mã hóa âm thanh đối với kỹ thuật mở rộng băng thông như, ví dụ, được sử dụng trong mã hóa âm thanh tiên tiến hiệu quả cao (High Efficiency Advanced Audio Coding - HE-AAC). Tín hiệu âm thanh ở đường dẫn 1300 là đầu vào vào trong hệ thống lọc bao gồm thông thấp 1302 và thông cao 1304. Đầu ra tín hiệu bởi bộ lọc thông cao 1304 là đầu vào vào trong bộ trích xuất/bộ mã hóa tham số 1306. Bộ trích xuất/bộ mã hóa tham số 1306 được tạo cấu hình để tính toán và mã hóa các tham số, ví dụ như tham số đường bao phổ, tham số cộng nhiễu âm, tham số sóng hài bị mất hoặc tham số lọc nghịch đảo. Các tham số trích xuất này được nhập vào trong bộ phân kênh bit 1308. Tín hiệu đầu ra thông thấp được nhập vào trong bộ xử lý thường bao gồm chức năng của bộ lấy mẫu giảm 1310 và bộ mã hóa lõi 1312. Thông thấp 1302 giới hạn băng thông sẽ được mã hóa đến băng thông nhỏ hơn một cách đáng kể so với xảy ra trong tín hiệu âm thanh đầu vào ban đầu trên đường dẫn 1300. Việc này cung cấp sự mã hóa đáng kể do thực tế rằng toàn bộ các chức năng xảy ra ở bộ mã hóa lõi chỉ phải hoạt động trên tín hiệu có băng thông giảm. Ví dụ, khi băng thông của tín hiệu âm thanh trên đường dẫn 1300 là 20kHz và ví dụ khi bộ lọc thông thấp 1302 có băng thông là 4kHz, để đáp ứng định lý lấy mẫu, theo lý thuyết, về mặt lý thuyết là đủ để tín hiệu theo sau bộ lấy mẫu giảm có tần số lấy mẫu bằng 8kHz, mà là sự giảm đáng kể cho tốc độ lấy mẫu cần thiết cho tín hiệu âm thanh 1300 vốn ít nhất phải là 40kHz.

Fig.13b minh họa giản đồ của bộ giải mã mở rộng băng thông tương ứng. Bộ giải mã bao gồm bộ phân kênh bit 1320. Bộ giải phân kênh bit 1320 trích xuất tín hiệu đầu vào đối với bộ giải mã lõi 1322 và tín hiệu đầu vào đối với bộ giải mã theo tham số 1324. Trong ví dụ trên đây, tín hiệu đầu ra bộ giải mã lõi có tốc độ lấy mẫu bằng 8kHz và do đó, băng thông bằng 4kHz trong khi đối với sự khôi phục băng thông hoàn toàn, tín hiệu đầu ra của bộ khôi phục tần số cao 1330 phải bằng 20kHz đòi hỏi tốc độ lấy mẫu ít nhất là 40kHz. Để việc này khả thi, cần phải có bộ xử lý bộ giải mã có chức năng của bộ lấy mẫu tăng 1325 và giàn lọc 1326. Sau đó, bộ khôi phục tần số cao 1330 nhận đầu vào tín hiệu tần số thấp được phân tích tần số bởi giàn lọc 1326 và khôi phục khoảng tần số được xác định bởi bộ lọc thông cao 1304 của Fig.13a bằng cách sử dụng



phép biểu diễn theo tham số của dải tần số cao. Bộ khôi phục tần số cao 1330 có vài chức năng như tái tạo khoảng tần số cao hơn bằng cách sử dụng khoảng nguồn trong khoảng tần số thấp, điều chỉnh đường bao phổ, chức năng cộng nhiều âm và chức năng đưa ra các sóng hài bị mất trong khoảng tần số cao hơn và, nếu được áp dụng và được tính toán trong bộ mã hoá của Fig.13a, thao tác lọc nghịch đảo để xét đến thực tế rằng khoảng tần số cao hơn thường không âm như khoảng tần số thấp. Trong HE-AAC, các sóng hài bị mất được tổng hợp lại ở phía bộ giải mã và đặt một cách chính xác vào khoảng giữa của dải khôi phục. Do đó, tất cả các đường dẫn sóng hài bị mất mà đã được xác định trong dải khôi phục nhất định không được đặt vào các giá trị tần số trong đó chúng được định vị trong tín hiệu ban đầu. Thay vào đó, các đường dẫn sóng hài bị mất này được đặt tại các tần số tại trung tâm của dải nhất định. Do đó, khi đường dẫn sóng hài bị mất trong tín hiệu ban đầu được đặt rất gần với đường biên dải khôi phục trong tín hiệu ban đầu, thì sai số về tần số được đưa ra bằng cách đặt đường dẫn sóng hài bị mất vào tín hiệu được khôi phục tại trung tâm của dải gần với 50% dải khôi phục riêng rẽ, cho các tham số đã được tạo ra và được truyền.

Hơn nữa, ngay cả khi mặc dù các bộ mã hóa lỗi âm thanh điển hình hoạt động trong miền phổ, tuy nhiên bộ giải mã lỗi vẫn tạo ra tín hiệu miền thời gian mà sau đó một lần nữa chuyển đổi thành miền phổ bởi chức năng giàn lọc 1326. Việc này tạo ra các độ trễ xử lý bổ sung, có thể tạo ra các giả âm do xử lý song song với việc biến đổi thứ nhất từ miền phổ thành miền tần số và một lần nữa biến đổi thông thường thành miền tần số khác và tất nhiên, điều này cũng đòi hỏi độ phức tạp tính toán đáng kể và nhờ đó công suất điện đặc biệt là một vấn đề khi kỹ thuật mở rộng băng thông được áp dụng trong các thiết bị di động như điện thoại di động, máy tính bảng hoặc máy tính xách tay, v.v.

Các bộ mã hóa-giải mã âm thanh hiện tại thực hiện việc mã hóa âm thanh tốc độ bit thấp bằng cách sử dụng BWE như một phần không thể thiếu của sơ đồ mã hóa. Tuy nhiên, các kỹ thuật BWE bị hạn chế để chỉ thay thế nội dung tần số cao (high frequency - HF). Hơn nữa, các kỹ thuật này không cho phép nội dung quan trọng về mặt cảm quan cao hơn tần số chéo nhất định được mã hóa dạng sóng. Do đó, các bộ mã hóa-giải mã âm thanh hiện thời mất chi tiết hoặc âm sắc HF khi BWE được triển

khai, vì việc căn chỉnh chính xác các sóng hài âm của tín hiệu không được xét đến trong hầu hết các hệ thống.

Một nhược điểm khác của hiện trạng các hệ thống BWE trong kỹ thuật hiện nay là cần phải chuyển đổi tín hiệu âm thanh thành một miền mới để triển khai BWE (ví dụ chuyển đổi từ miền MDCT sang miền QMF). Điều này dẫn đến sự phức tạp của việc đồng bộ hoá, độ phức tạp tính toán bổ sung và tăng yêu cầu bộ nhớ.

Thông thường, các sơ đồ mở rộng băng thông sử dụng phép vá phổ với mục đích tái tạo lại vùng phổ tần số cao trên một tần số được gọi là tần số chéo nhất định. Vùng HF bao gồm nhiều bản vá liền kề và mỗi bản vá này được lấy từ cùng một vùng thông dải của phổ tần số thấp bên dưới tần số chéo đã cho. Trong phép biểu diễn giàn lọc của các tín hiệu, các hệ thống này sao chép tập hợp các hệ số băng con lân cận trong số phổ tần số thấp vào trong vùng HF. Các đường biên của các tập hợp được chọn thường phụ thuộc vào hệ thống và không phụ thuộc vào tín hiệu. Đối với một vài nội dung tín hiệu, việc lựa chọn bản vá tĩnh này có thể tạo ra âm sắc và sắc thái khó chịu của tín hiệu được khôi phục.

Các phương pháp khác truyền tín hiệu LF vào vùng HF qua điều biến dải phụ đơn lẻ (single side band – SSB) thích ứng tần số. Các phương pháp này có độ phức tạp tính toán cao so với các quy trình sao chép, vì các phương pháp này hoạt động ở tốc độ lấy mẫu cao đối với các tín hiệu miền thời gian.

Hơn nữa, việc vá có thể không ổn định cụ thể là đối với các tín hiệu không phải tiếng nói chẳng hạn như tiếng nói vô thanh. Do đó, các sơ đồ vá đã biết có thể tạo ra sự suy yếu trong tín hiệu âm thanh.

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. Ferreira<sup>1</sup> 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**

Mục đích của sáng chế là đề xuất thiết bị và phương pháp mã hoá/giải mã làm cho chất lượng âm thanh tốt hơn.

Mục đích này đạt được bởi thiết bị giải mã theo điểm 1, thiết bị mã hoá theo điểm 8, phương pháp giải mã âm thanh theo điểm 19 hoặc phương pháp giải mã âm thanh theo điểm 20 hoặc chương trình máy tính theo điểm 21.

Sáng chế dựa trên phát hiện rằng một vài suy giảm nhất định trong chất lượng âm thanh có thể được sửa chữa bằng cách áp dụng sơ đồ điền đầy lát tần số thích ứng tín hiệu. Với mục tiêu này, việc phân tích ở phía bộ mã hoá được thực hiện để tìm ra ứng viên vùng nguồn phù hợp tốt nhất đối với vùng đích nhất định. Thông tin so khớp nhận biết cho vùng đích vùng nguồn nhất định cùng với một vài thông tin bổ sung tùy ý được tạo ra và được truyền dưới dạng thông tin phụ tới bộ giải mã. Sau đó, bộ giải mã áp dụng thao tác điền đầy lát tần số bằng cách sử dụng thông tin so khớp. Với mục

tiêu này, bộ giải mã đọc thông tin so khớp từ dòng dữ liệu được truyền hoặc tệp dữ liệu và truy cập vào vùng nguồn được nhận biết đối với dải được khôi phục nhất định, và nếu được biểu thị trong thông tin so khớp, sẽ thực hiện bổ sung một vài việc xử lý của dữ liệu vùng nguồn này để tạo ra dữ liệu phổ thô cho dải khôi phục. Sau đó, kết quả của hoạt động điền đầy lát tần số, nghĩa là dữ liệu phổ thô đối với dải khôi phục, được tạo hình bằng cách sử dụng thông tin đường bao phổ để cuối cùng thu được dải khôi phục mà bao gồm các phần phổ thứ nhất như các phần tiếng nói. Tuy nhiên, các phần tiếng nói này không được tạo ra bởi sơ đồ điền đầy ô thích ứng, nhưng các phần phổ thứ nhất được xuất ra bởi bộ giải mã âm thanh hoặc bộ giải mã lời một cách trực tiếp.

Sơ đồ lựa chọn ô phổ thích ứng có thể hoạt động với độ chi tiết thấp. Theo phương án này, vùng nguồn thông thường được chia thành các vùng nguồn chồng lấp và vùng đích hoặc các dải khôi phục được đưa ra bởi các vùng đích tần số không chồng lấp. Sau đó, tính tương tự giữa mỗi vùng nguồn và mỗi vùng đích được xác định ở phía bộ mã hoá và cặp phù hợp tốt nhất của vùng nguồn và vùng đích được nhận biết bởi thông tin so khớp và ở phía bộ giải mã, vùng nguồn được nhận biết trong thông tin so khớp được sử dụng để tạo ra dữ liệu phổ thô đối với dải khôi phục.

Để đạt được mục tiêu thu được độ chi tiết cao hơn, mỗi vùng nguồn được cho phép chuyển dịch để thu được độ trễ nhất định trong đó tính tương tự là tối đa. Độ trễ này có thể là ngắn tần số mịn và cho phép sự phù hợp thậm chí tốt hơn giữa vùng nguồn và vùng đích.

Hơn nữa, ngoài việc chỉ nhận biết cặp phù hợp tốt nhất, độ trễ tương quan này cũng có thể được truyền trong thông tin so khớp và ngoài ra, ngay cả khi tín hiệu có thể được truyền. Khi tín hiệu được xác định là âm ở phía bộ mã hoá, khi đó ký hiệu tín hiệu tương ứng cũng được truyền trong thông tin so khớp và ở phía bộ giải mã, các giá trị phổ vùng nguồn được nhân với “-1” hoặc, trong phép biểu diễn phức, được “xoay”  $180^\circ$ .

Phương án khác của sáng chế áp dụng hoạt động làm trắng ô. Việc làm trắng phổ loại bỏ thông tin đường bao phổ thô và làm nổi bật cấu trúc mịn phổ mà là sự quan tâm hàng đầu để đánh giá tính tương tự ô. Do đó, một mặt lát tần số và/hoặc mặt khác

tín hiệu nguồn được làm trắng trước khi tính toán phép đo tương quan chéo. Khi chỉ ô được làm trắng bằng cách sử dụng quy trình được định trước, ký hiệu làm trắng được truyền biểu thị cho bộ giải mã mà quy trình làm trắng định trước tương tự sẽ được áp dụng cho lát tần số trong IGF.

Đối với việc lựa chọn ô, ưu tiên sử dụng độ trễ tương quan để chuyển dịch theo phổ được tái tạo lại bởi số nguyên của các ngăn biến đổi. Phụ thuộc vào việc biến đổi cơ sở, việc chuyển dịch phổ có thể đòi hỏi việc hiệu chỉnh bổ sung. Trong trường hợp độ trễ lẻ, ô được điều biến bổ sung qua phép nhân theo chuỗi theo thời gian luân phiên  $-1/1$  để bù cho phép biểu diễn đảo ngược tần số của mỗi dải khác trong MDCT. Hơn nữa, tín hiệu của kết quả tương quan được áp dụng khi tạo ra lát tần số.

Hơn nữa, ưu tiên sử dụng sự chia nhỏ và ổn định hoá ô để đảm bảo tránh được các giả âm được tạo ra bởi việc thay đổi nhanh các vùng nguồn cho cùng vùng khôi phục hoặc vùng đích. Với mục tiêu này, phép phân tích tương tự trong số các vùng nguồn được nhận biết khác nhau được thực hiện và khi lát nguồn là tương tự với các lát nguồn khác với tính tương tự cao hơn ngưỡng, khi đó lát nguồn này có thể rơi từ tập hợp các lát nguồn tiềm năng do nó có tương quan ở mức cao với các lát nguồn khác. Hơn nữa, đối với sự ổn định hóa lựa chọn ô, ưu tiên duy trì thứ tự ô từ khung trước nếu lát nguồn trong khung hiện thời không tương quan (tốt hơn so với ngưỡng đã nêu) với các lát đích trong khung hiện thời.

Khía cạnh khác dựa trên phát hiện rằng chất lượng âm thanh của tín hiệu được khôi phục có thể được cải thiện qua IGF do toàn bộ phổ có thể truy cập vào bộ mã hoá lõi sao cho, ví dụ, các phần tiếng nói quan trọng theo cảm quan trong khoảng phổ ở mức cao vẫn có thể được mã hoá bởi bộ mã hóa lõi thay vì thay thế theo tham số. Ngoài ra, hoạt động điền đầy khoảng trống bằng cách sử dụng các lát tần số từ tập hợp thứ nhất của các phần phổ thứ nhất mà, ví dụ, là tập hợp các phần tiếng nói thường là từ khoảng tần số thấp, nhưng cũng từ khoảng tần số cao hơn nếu sẵn có, được thực hiện. Tuy nhiên, đối với việc điều chỉnh đường bao phổ ở phía bộ giải mã, các phần phổ từ tập hợp thứ nhất của các phần phổ được định vị ở dải khôi phục không được xử lý sau thêm nữa bằng, ví dụ, sự điều chỉnh đường bao phổ. Chỉ có các giá trị phổ còn lại trong dải khôi phục mà không bắt nguồn từ bộ giải mã lõi sẽ được điều chỉnh

đường bao bằng cách sử dụng thông tin đường bao. Tốt hơn là, thông tin đường bao là thông tin đường bao đầy đủ khi xét đến năng lượng của tập hợp thứ nhất của các phần phổ thứ nhất trong dải khôi phục và tập hợp thứ hai của các phần phổ thứ hai trong cùng dải khôi phục, trong đó các giá trị phổ sau trong tập hợp thứ hai của các phần phổ thứ hai được biểu thị bằng không và do đó, không được mã hoá bởi bộ mã hoá lỗi, nhưng được mã hóa theo tham số với thông tin năng lượng phân giải thấp.

Các tác giả sáng chế đã phát hiện ra rằng các giá trị năng lượng tuyệt đối, được chuẩn hoá đối với băng thông của dải tương ứng hoặc không được chuẩn hoá, là hữu ích và rất hữu hiệu khi áp dụng ở phía bộ giải mã. Điều này áp dụng một cách đặc biệt khi các hệ số thu được phải được tính toán dựa trên năng lượng dư trong dải khôi phục, năng lượng mất trong dải khôi phục và thông tin lát tần số trong dải khôi phục.

Hơn nữa, tốt hơn là dòng bit được mã hoá không chỉ bao phủ thông tin năng lượng cho các dải khôi phục, mà ngoài ra, các hệ số tỷ lệ dùng cho các dải hệ số tỷ lệ mở rộng đến tần số tối đa. Việc này đảm bảo rằng đối với mỗi dải khôi phục, mà phần tiếng nói nhất định, nghĩa là phần phổ thứ nhất là sẵn có, tập hợp thứ nhất này của phần phổ thứ nhất thực tế có thể được giải mã với biên độ bên phải. Hơn nữa, ngoài hệ số tỷ lệ đối với mỗi dải khôi phục, năng lượng dùng cho dải khôi phục được tạo ra trong bộ mã hoá và được truyền đến bộ giải mã. Hơn nữa, tốt hơn là các dải khôi phục trùng khớp với các dải hệ số tỷ lệ hoặc trong trường hợp phân nhóm năng lượng, ít nhất các đường biên của dải khôi phục trùng khớp với các đường biên của các dải hệ số tỷ lệ.

Khía cạnh khác dựa trên phát hiện rằng các vấn đề một mặt liên quan đến việc chia tách cho sự mở rộng băng thông và mặt khác liên quan đến việc mã hóa lỗi có thể được giải quyết và khắc phục bằng cách thực hiện việc mở rộng băng thông trong cùng miền phổ trong đó bộ giải mã lỗi hoạt động. Do đó, bộ giải mã lỗi tốc độ đầy đủ được đề xuất mà mã hoá và giải mã khoảng tín hiệu âm thanh đầy đủ. Điều này không cần bộ lấy mẫu giảm ở phía bộ mã hoá và bộ lấy mẫu tăng ở phía bộ giải mã. Thay vào đó, toàn bộ việc xử lý được thực hiện với tốc độ lấy mẫu đầy đủ hoặc miền băng thông đầy đủ. Để mã hóa ở mức cao, tín hiệu âm thanh được phân tích để tìm ra tập hợp thứ nhất của các phần phổ thứ nhất mà sẽ được mã hoá với độ phân giải cao, trong đó tập

hợp thứ nhất này của các phần phổ thứ nhất có thể bao gồm, theo phương án, các phần tiếng nói của tín hiệu âm thanh. Mặt khác, các thành phần vô thanh hoặc thành phần nhiều âm trong tín hiệu âm thanh cấu thành tập hợp thứ hai của các phần phổ thứ hai được mã hoá theo tham số với độ phân giải phổ thấp. Sau đó, tín hiệu âm thanh được mã hoá chỉ đòi hỏi tập hợp thứ nhất của các phần phổ thứ nhất được mã hoá theo cách bảo toàn dạng sóng với độ phân giải phổ cao và ngoài ra, tập hợp thứ hai của các phần phổ thứ hai được mã hoá theo tham số với độ phân giải thấp sử dụng “lát” tần số được bắt nguồn từ tập hợp thứ nhất. Ở phía bộ giải mã, bộ giải mã lõi, mà là bộ giải mã dải đầy đủ, khôi phục tập hợp thứ nhất của các phần phổ thứ nhất theo cách bảo toàn dạng sóng, nghĩa là, không biết rằng có bất kỳ sự khôi phục tần số bổ sung. Tuy nhiên, phổ được tạo ra này có nhiều khoảng trống phổ. Các khoảng trống này được điền đầy một cách tuần tự bằng kỹ thuật điền đầy khoảng trống thông minh (IGF) của sáng chế bằng cách sử dụng, một mặt, tái tạo tần số áp dụng các dữ liệu theo tham số và mặt khác sử dụng khoảng phổ nguồn, nghĩa là, các phần phổ thứ nhất được khôi phục bởi bộ giải mã âm thanh tốc độ đầy đủ.

Theo phương án khác, các phần phổ, mà được khôi phục bằng việc điền đầy nhiều âm chỉ tốt hơn việc tái tạo bằng thông hoặc điền đầy lát tần số, cấu thành tập hợp thứ ba của các phần phổ thứ ba. Do thực tế là, một mặt, khái niệm mã hóa hoạt động trong miền đơn lẻ để mã hóa lõi/giải mã và mặt khác tái tạo tần số, IGF không chỉ bị giới hạn ở việc điền đầy khoảng tần số cao hơn mà có thể điền đầy các khoảng tần số thấp, bằng việc điền đầy nhiều âm mà không tái tạo tần số hoặc tái tạo tần số bằng cách sử dụng lát tần số ở khoảng tần số khác nhau.

Hơn nữa, nhấn mạnh rằng thông tin đối với các năng lượng phổ, thông tin đối với các năng lượng riêng rẽ hoặc thông tin năng lượng riêng rẽ, thông tin đối với năng lượng còn lại hoặc thông tin năng lượng còn lại, thông tin đối với năng lượng ô hoặc thông tin năng lượng ô hoặc thông tin đối với năng lượng mất hoặc thông tin năng lượng mất có thể không chỉ bao gồm giá trị năng lượng, mà còn cả giá trị biên độ (ví dụ tuyệt đối), giá trị mức hoặc giá trị khác bất kỳ, mà từ đó giá trị năng lượng cuối cùng có thể được suy ra. Do đó, thông tin đối với năng lượng có thể, ví dụ bao gồm bản thân giá trị năng lượng, và/hoặc giá trị mức và/hoặc biên độ và/hoặc biên độ tuyệt đối.



Khía cạnh khác dựa trên phát hiện rằng trạng thái tương quan không chỉ quan trọng đối với khoảng nguồn, mà cũng quan trọng đối với khoảng đích. Hơn nữa, sáng chế xác nhận trạng thái mà các trạng thái tương quan khác nhau có thể xảy ra trong khoảng nguồn và khoảng đích. Ví dụ, khi tín hiệu tiếng nói có nhiều âm tần số cao được xét đến, trạng thái có thể là dải tần số thấp bao gồm tín hiệu tiếng nói với số lượng nhỏ âm bội có tương quan ở mức cao ở kênh phía bên trái và kênh phía bên phải, khi loa được đặt ở giữa. Tuy nhiên, phần tần số cao có thể không tương quan nhiều do thực tế rằng có thể có nhiều âm tần số cao khác ở phía bên trái so với nhiều âm tần số cao khác hoặc không có nhiều âm tần số cao ở phía bên phải. Do đó, khi hoạt động điền đầy khoảng trống thẳng có thể được thực hiện mà bỏ qua trạng thái này, khi đó phần tần số cao cũng sẽ tương quan, và việc này có thể tạo ra các giả âm chia tách không gian nghiêm trọng trong tín hiệu khôi phục. Để giải quyết vấn đề này, dữ liệu tham số đối với dải khôi phục hoặc thông thường, cho tập hợp thứ hai của các phần phổ thứ hai mà phải được khôi phục bằng cách sử dụng tập hợp thứ nhất của các phần phổ thứ nhất được tính toán để nhận biết phép biểu diễn kênh đôi khác nhau thứ nhất hoặc thứ hai cho phần phổ thứ hai hoặc, đưa ra theo cách khác nhau, đối với dải khôi phục. Ở phía bộ mã hoá, việc nhận biết kênh đôi do đó được tính toán đối với các phần phổ thứ hai, nghĩa là đối với các phần, mà ngoài ra, thông tin năng lượng cho dải khôi phục được tính toán. Sau đó, bộ tái tạo tần số ở phía bộ giải mã tái tạo phần phổ thứ hai phụ thuộc vào phần thứ nhất của tập hợp thứ nhất của các phần phổ thứ nhất, nghĩa là khoảng nguồn và dữ liệu theo tham số đối với phần thứ hai như thông tin năng lượng đường bao phổ hoặc dữ liệu đường bao phổ khác bất kỳ và ngoài ra, phụ thuộc vào việc nhận biết kênh đôi đối với phần thứ hai, nghĩa là đối với dải khôi phục này khi xem xét lại.

Việc nhận biết kênh đôi tốt hơn là nên được truyền dưới dạng ký hiệu đối với mỗi dải khôi phục và dữ liệu này được truyền từ bộ mã hoá đến bộ giải mã và bộ giải mã sau đó giải mã tín hiệu lời như được biểu thị bởi các ký hiệu được tính toán tốt hơn cho các dải lời. Sau đó, trong quá trình triển khai, tín hiệu lời được lưu trữ trong cả hai phép biểu diễn âm thanh nổi (ví dụ trái/phải và giữa/bên) và để điền đầy lát tần số IGF, phép biểu diễn lát nguồn được chọn để phù hợp với phép biểu diễn lát đích như được

biểu thị bởi các ký hiệu nhận biết kênh đôi cho việc điền đầy khoảng trống thông minh hoặc các dải khôi phục, nghĩa là đối với khoảng đích.

Nhấn mạnh rằng quy trình này không có tác dụng đối với các tín hiệu âm thanh nổi, nghĩa là đối với kênh bên trái và kênh bên phải, mà cũng hoạt động đối với các tín hiệu đa kênh. Trong trường hợp các tín hiệu đa kênh, vài cặp kênh khác nhau có thể được xử lý theo cách đó như kênh bên trái và bên phải là cặp thứ nhất, kênh bao quanh bên trái và bao quanh bên phải là cặp thứ hai và kênh trung tâm và kênh LFE là cặp thứ ba. Các cặp khác có thể được xác định đối với các dạng kênh đầu ra cao hơn như 7,1, 11,1, v.v.

Khía cạnh khác dựa trên phát hiện rằng chất lượng được cải thiện và tốc độ bit giảm rõ ràng đối với các tín hiệu bao gồm các phần chuyển tiếp do chúng xảy ra rất thường xuyên trong các tín hiệu âm thanh 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 sự khôi phục tần số cao. TNS/TTS xử lý ở phía bộ mã hoá sẽ được triển khai nhờ dự báo qua tần số khôi phục đường bao thời gian của tín hiệu âm thanh. Phụ thuộc vào phương án, nghĩa là khi bộ lọc tạo hình nhiễu âm theo thời gian được xác định trong khoảng tần số không chỉ bao phủ khoảng tần số nguồn, mà cũng bao gồm khoảng tần số mục tiêu sẽ được khôi phục trong bộ giải mã tái tạo tần số, đường bao theo thời gian không chỉ được áp dụng cho tín hiệu âm thanh lỗi đối với tần số bắt đầu điền đầy khoảng trống, mà cả đường bao theo thời gian cũng được áp dụng cho các khoảng phổ của các phần phổ thứ hai được khôi phục. Do đó, tiếng vang trước hoặc tiếng vang sau mà có thể xảy ra mà không có sự tạo hình theo thời gian được giảm hoặc được loại trừ. Việc này được hoàn thành bằng cách áp dụng sự dự báo nghịch đảo theo tần số không chỉ trong khoảng tần số lỗi đối với tần số bắt đầu điền đầy khoảng trống nhất định mà còn cả trong khoảng tần số cao hơn khoảng tần số lỗi. Với mục tiêu này, việc tái tạo tần số hoặc tái tạo lát tần số được thực hiện ở phía bộ giải mã trước khi áp dụng sự dự báo theo tần số. Tuy nhiên, việc dự báo theo tần số có thể được áp dụng trước hoặc sau việc tạo hình đường bao phổ phụ thuộc vào việc tính toán thông tin năng lượng đã được thực hiện đối với các giá trị dư phổ sau khi lọc hoặc các giá trị phổ (đầy đủ) trước khi tạo hình đường bao.

Việc xử lý TTS qua một hoặc nhiều lát tần số còn thiết lập bổ sung tính liên tục tương quan giữa khoảng nguồn và khoảng khôi phục hoặc trong hai khoảng khôi phục hoặc các lát tần số liên kề.

Theo phương án, ưu tiên sử dụng việc lọc TNS/TTS phức hợp. Nhờ đó, tránh được các giả âm răng cưa (theo thời gian) của phép biểu diễn thực được lấy mẫu tới hạn, tương tự MDCT. Bộ lọc TNS phức hợp có thể được tính toán ở phía bộ mã hoá bằng cách áp dụng không chỉ biến đổi cosin rời rạc cải biên mà ngoài ra còn cả biến đổi sin rời rạc cải biên để thu được biến đổi cải biên phức hợp. Tuy nhiên, chỉ có các giá trị biến đổi cosin rời rạc cải biên, nghĩa là, phần thực của biến đổi phức hợp được truyền. Tuy nhiên, ở phía bộ giải mã, có thể ước lượng phần ảo của biến đổi bằng cách sử dụng các phổ MDCT của các khung trước hoặc sau để, ở phía bộ giải mã, bộ lọc phức hợp có thể được áp dụng một lần nữa trong việc dự báo nghịch đảo theo tần số và cụ thể là, sự dự báo qua đường biên giữa khoảng nguồn và khoảng khôi phục và cả qua đường biên giữa các lát tần số liên kề tần số trong khoảng khôi phục.

Hệ thống mã hóa âm thanh của sáng chế mã hóa một cách hữu hiệu các tín hiệu âm thanh tùy ý ở khoảng rộng của các tốc độ bit. Trong khi đó, đối với các tốc độ bit cao, hệ thống của sáng chế đồng quy tới độ trong suốt, đối với tốc độ bit thấp sự khó chịu cảm quan được giảm thiểu. Do đó, phần chung chính của tốc độ bit sẵn có được sử dụng để mã hóa dạng sóng chỉ cho cấu trúc liên quan nhất theo cảm quan của tín hiệu trong bộ mã hoá và các khoảng trống phổ thu được được điền đầy trong bộ giải mã với nội dung tín hiệu mà gần bằng phổ ban đầu. Ngân sách bit rất bị giới hạn được giả định điều khiển việc gọi là điền đầy khoảng trống thông minh (IGF) theo phổ bằng tham số bởi thông tin phụ chuyên dụng được truyền từ bộ mã hoá đến bộ giải mã.

### **Mô tả vắn tắt các hình vẽ**

Các phương án được ưu tiên của sáng chế sẽ được mô tả sau đây kết hợp với các hình vẽ, trong đó:

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

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

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

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

Fig.3a minh họa phép biểu diễn bằng sơ đồ của phổ như được tạo ra bởi bộ giải mã miền phổ của Fig.1b;

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

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

Fig.4b minh họa việc thực hiện về chức năng của Fig.4a;

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

Fig.5b minh họa chức năng của bộ giải mã với kỹ thuật MDCT;

Fig.5c minh họa phương án 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 tạo hình nhiễu âm theo thời gian/chức năng tạo hình ô theo thời gian với thứ tự khác nhau của bộ lọc dự báo phổ và bộ tạo hình phổ;

Fig.7a minh họa phương án về chức năng tạo hình ô theo thời gian (TTS);

Fig.7b minh họa phương án bộ giải mã phù hợp với phương án bộ mã hoá của Fig.7a;

Fig.7c minh họa ảnh phổ của tín hiệu ban đầu và tín hiệu mở rộng mà không có TTS;

Fig.7d minh họa phép biểu diễn tần số minh họa sự tương ứng 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 ban đầu và tín hiệu 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 các phép biểu diễn và các khoảng nguồn/nơi đến;

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

Fig.8d minh họa phương án chi tiết hơn của bộ giải mã của Fig.8a;

Fig.8e minh họa phương án của bộ giải mã xử lý kênh đôi sẽ được giải mã bởi bộ giải mã của Fig.8a:

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

Fig.9b minh họa phương án chi tiết hơn của bộ tái tạo tần số của Fig.9a;

Fig.9c minh họa việc minh họa theo biểu đồ chức năng của Fig.9b;

Fig.9d minh họa phương án khác của bộ giải mã của Fig.9a;

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

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

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

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

Fig.11a minh họa bộ giải mã khác có sự nhận biết khoảng nguồn cụ thể đối với hoạt động tạo ô phổ trong bộ giải mã;

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

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

Fig.11d minh họa sơ đồ khối của phương án của bộ tính toán tham số của Fig.11c;

Fig.12a và Fig.12b minh họa lược đồ tần số để minh họa khoảng nguồn và khoảng đích;

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

Fig.13a minh họa bộ mã hoá kỹ thuật ưu tiên với sự mở rộng băng thông; và

Fig.13b minh họa bộ giải mã kỹ thuật ưu tiên với sự mở rộng băng thông.

### **Mô tả chi tiết sáng chế**

Fig.1a minh họa thiết bị để mã hoá tín hiệu âm thanh 99. Tín hiệu âm thanh 99 được đưa vào trong bộ chuyển đổi phổ theo thời gian 100 để chuyển đổi tín hiệu âm thanh có tốc độ lấy mẫu thành phép biểu diễn phổ 101 được xuất ra bởi bộ chuyển đổi phổ theo thời gian. Phổ 101 được đưa vào trong 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 sẽ được mã hoá với độ phân giải phổ thứ nhất và tập hợp thứ hai khác của các phần phổ thứ hai 105 sẽ được mã hoá với độ phân giải phổ thứ hai. Độ phân giải phổ thứ hai nhỏ hơn so với độ phân giải phổ thứ nhất. Tập hợp thứ hai của các phần phổ thứ hai 105 được đưa vào trong bộ tính toán 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 hoá thứ nhất 107 của tập hợp thứ nhất của các phần phổ thứ nhất có độ phân giải phổ thứ nhất. Hơn nữa, bộ tính toán tham số/bộ mã hóa theo tham số 104 được tạo cấu hình để tạo ra phép biểu diễn được mã hoá thứ hai 109 của tập hợp thứ hai của các phần phổ thứ hai. Phép biểu diễn được mã hoá thứ nhất 107 và phép biểu diễn được mã hoá 109 thứ hai được đưa vào trong bộ phân kênh bit hoặc bộ tạo mẫu dòng bit 108 và khối 108 cuối cùng xuất ra tín hiệu âm thanh được mã hoá để truyền hoặc lưu trữ trên thiết bị lưu trữ.

Thông thường, phần phổ thứ nhất chẳng hạn như 306 của Fig.3a sẽ được bao quanh bởi hai phần phổ thứ hai chẳng hạn như 307a, 307b. Đây không phải là trường hợp trong HE AAC, trong đó khoảng tần số bộ mã hóa lõi là dải giới hạn.

Fig.1b minh họa bộ giải mã phù hợp với bộ mã hoá của Fig.1a. Phép biểu diễn được mã hoá thứ nhất 107 được đưa vào trong bộ giải mã âm thanh miền phổ 112 để tạo ra phép 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. Hơn nữa, phép biểu diễn được mã hoá thứ hai 109 được đưa vào trong 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 so với 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 bằng cách sử dụng phần phổ thứ nhất. Bộ tái tạo tần số 116 thực hiện hoạt động điền đầy ô, nghĩa 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 vào trong khoảng khôi phục hoặc dải khôi phục có phần phổ thứ hai và thường thực hiện việc tạo hình đường bao phổ hoặc hoạt động khác như được biểu thị bởi đầu ra phép biểu diễn thứ hai được giải mã bởi bộ giải mã theo tham số 114, nghĩa là bằng cách sử dụng thông tin đối với 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 ở đầu ra của bộ tái tạo tần số 116 trên đường dẫn 117 được đưa vào trong bộ chuyển đổi phổ theo thời gian 118 được tạo cấu hình để chuyển đổi phép 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 có tốc độ lấy mẫu cao nhất định.

Fig.2b minh họa phương án của bộ mã hoá của Fig.1a. Tín hiệu đầu ra âm thanh 99 được đưa vào trong giàn lọc phân tích 220 tương ứng với bộ chuyển đổi phổ theo thời gian 100 của Fig.1a. Sau đó, hoạt động tạo hình nhiễu âm theo thời gian được thực hiện trong khối TNS 222. Do đó, đầu vào trong bộ phân tích phổ 102 của Fig.1a tương ứng với màn chắn âm khối 226 của Fig.2b có thể là các giá trị phổ đầy đủ, khi việc hoạt động tạo hình nhiễu âm theo thời gian/tạo hình ô theo thời gian không được áp dụng hoặc có thể là các giá trị dư phổ, khi hoạt động TNS như được minh họa trên Fig.2b, khối 222 được áp dụng. Đối với các tín hiệu kênh đôi hoặc các tín hiệu đa kênh, việc mã hóa kênh ghép 228 còn có thể được thực hiện, sao cho bộ mã hoá miền phổ 106 của Fig.1a có thể bao gồm khối mã hóa kênh ghép 228. Hơn nữa, bộ mã hóa

entropy 232 để thực hiện việc nén dữ liệu không hao tổn được đề xuất mà cũng là một phần của bộ mã hoá miền phổ 106 của Fig.1a.

Bộ phân tích phổ/màn chắn âm 226 chia tách đầu vào của khối TNS 222 thành dải lõi và các thành phần âm tương ứng với tập hợp thứ nhất của các phần phổ thứ nhất 103 và các thành phần dư tương ứng với tập hợp thứ hai của các phần phổ thứ hai 105 của Fig.1a. Khối 224 được biểu thị dưới dạng mã hóa trích xuất theo tham số IGF tương ứng với bộ mã hóa theo tham số 104 của Fig.1a và bộ phân kênh bit 230 tương ứng với bộ phân kênh bit 108 của Fig.1a.

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

Bộ phân tích phổ 226 ưu tiên áp dụng màn chắn âm. Giai đoạn ước lượng màn chắn âm này được sử dụng để chia tách các thành phần âm từ các thành phần tương tự nhiều âm trong tín hiệu. Việc 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 này có thể được thực hiện theo nhiều cách khác nhau và được ưu tiên thực hiện tương tự với chức năng của nó với giai đoạn ước lượng rãnh ghi hình sin được sử dụng trong việc tạo mô hình sin và nhiễu âm để mã hóa tiếng nói/âm thanh trong tài liệu tham khảo [8, 9] hoặc mô hình HILN dựa trên bộ mã hóa âm thanh được mô tả trong tài liệu tham khảo [10]. Tốt hơn là sử dụng phương án mà dễ dàng triển khai mà không cần phải duy trì các quỹ đạo sinh-tử, nhưng bộ phát hiện tiếng nói hoặc nhiễu âm khác bất kỳ cũng có thể được sử dụng.

Mô đun IGF tính toán sự tương tự mà tồn tại giữa vùng nguồn và vùng đích. Vùng đích sẽ được biểu diễn bởi phổ từ vùng nguồn. Phép đo sự tương tự giữa vùng nguồn và vùng đích được thực hiện bằng cách sử dụng phương pháp tương quan chéo. Vùng đích được tách thành các lát tần số không chồng lấp  $nTar$ . Đối với mỗi lát trong vùng đích, các lát nguồn  $nSrc$  được tạo ra từ tần số bắt đầu cố định. Các lát nguồn này chồng lấp bởi hệ số nằm trong khoảng từ 0 đến 1, trong đó 0 có nghĩa là 0% chồng lấp và 1 có nghĩa là 100% chồng lấp. Mỗi lát trong số các lát nguồn này tương quan với lát



đích ở các độ trễ khác nhau để phát hiện lát nguồn mà phù hợp tốt nhất với lát đích. Số lượng lát phù hợp tốt nhất được lưu trữ trong *tileNum[idx\_tar]*, độ trễ mà trong đó tương quan tốt nhất với mục tiêu được lưu trữ trong *xcorr\_lag[idx\_tar][idx\_src]* và tín hiệu tương quan được lưu trữ trong *xcorr\_lag[idx\_tar][idx\_src]*. Trong trường hợp sự tương quan là âm ở mức cao, lát nguồn cần phải được nhân với -1 trước quy trình điền đầy ô ở bộ giải mã. Mô đun IGF cũng chịu trách nhiệm không ghi đè các thành phần âm trong phổ do các thành phần âm được bảo toàn bằng cách sử dụng màn chắn âm. Tham số năng lượng theo dải được sử dụng để lưu trữ năng lượng của vùng đích cho phép khôi phục phổ một cách chính xác.

Phương pháp này có các ưu điểm nhất định so với SBR cổ điển trong tài liệu tham khảo [1] trong đó lưới sóng hài của tín hiệu đa âm sắc được bảo toàn bởi bộ mã hóa lỗi trong khi chỉ các khoảng trống giữa các đường sin được điền đầy với “nhiều âm được tạo hình” phù hợp tốt nhất từ vùng nguồn. Ưu điểm khác của hệ thống này so với ASR (thay thế phổ chính xác) trong các tài liệu tham khảo [2-4] là sự vắng mặt của giai đoạn tổng hợp tín hiệu mà tạo ra các phần quan trọng của tín hiệu ở bộ giải mã. Thay vào đó, nhiệm vụ này được thực hiện bởi bộ mã hóa lỗi, cho phép bảo toàn các thành phần quan trọng của phổ. Ưu điểm khác của hệ thống được đề xuất là khả năng tạo tỷ lệ liên tục mà các đặc tính tạo ra. Chỉ sử dụng *tileNum[idx\_tar]*, đối với mỗi lát được gọi là so khớp tổng độ kết hạt và có thể được sử dụng cho tốc độ bit thấp trong khi sử dụng biến *xcorr\_lag* đối với mỗi ô cho phép khớp phổ đích và nguồn tốt hơn.

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

Trong trường hợp các cặp kênh âm thanh nổi, việc xử lý âm thanh nổi ghép bổ sung được áp dụng. Việc này là cần thiết, bởi vì đối với khoảng đích nhất định, tín hiệu có thể tương quan ở mức cao với nguồn âm thanh theo nhịp. Trong trường hợp các vùng nguồn được chọn cho vùng cụ thể này không được tương quan tốt, mặc dù các năng lượng được khớp đối với các vùng đích, hình ảnh không gian có thể bị ảnh hưởng do các vùng nguồn không tương quan. Bộ mã hoá phân tích mỗi dải năng lượng vùng đích, thường thực hiện tương quan chéo của các giá trị phổ và nếu ngưỡng nhất định bị vượt quá, sẽ thiết lập ký hiệu cho dải năng lượng này. Trong bộ giải mã, các dải năng

lượng kênh bên trái và bên phải được xử lý theo cách riêng rẽ nếu ký hiệu âm thanh nổi này không được thiết lập. Trong trường hợp ký hiệu âm thanh nổi ghép được thiết lập, cả các năng lượng và việc vá được thực hiện trong miền âm thanh nổi ghép. Thông tin âm thanh nổi ghép cho các vùng IGF được tín hiệu hóa tương tự thông tin âm thanh nổi ghép để mã hóa lỗi, bao gồm ký hiệu biểu thị trong trường hợp dự báo nếu việc định của dự báo là từ trộn giảm đến dư hoặc ngược lại.

Các năng lượng có thể được tính toán từ các năng lượng 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.

Một giải pháp khác để tính toán và truyền các năng lượng một cách trực tiếp trong miền âm thanh nổi ghép cho các dải trong đó âm thanh nổi ghép được kích hoạt, do đó không cần sự biến đổi năng lượng bổ sung nào ở phía bộ giải mã.

Các lát nguồn luôn 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 thanh nổi ghép -> biến đổi LR:

Nếu không có tham số dự báo bổ sung nào 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 hóa từ giữa đến bên:

$$\begin{aligned}
sideTile[k] &= sideTile[k] - predictionCoeff \cdot midTile[k] \\
leftTile[k] &= midTile[k] + sideTile[k] \\
rightTile[k] &= midTile[k] - sideTile[k]
\end{aligned}$$

Nếu hướng tín hiệu hóa là từ bên đến giữa:

$$\begin{aligned}
midTile1[k] &= midTile[k] - predictionCoeff \cdot sideTile[k] \\
leftTile[k] &= midTile1[k] - sideTile[k] \\
rightTile[k] &= midTile1[k] + sideTile[k]
\end{aligned}$$

Việc xử lý này đảm bảo rằng từ các lát được sử dụng để tái tạo các vùng đến tương quan ở mức cao và các vùng đến theo nhịp, các kênh bên trái và bên phải thu được vẫn biểu diễn nguồn âm thanh tương quan và theo nhịp ngay cả nếu các vùng nguồn không tương quan, bảo toàn hình ảnh âm thanh nổi đối với các vùng này.

Nói cách khác, trong dòng bit, các ký hiệu âm thanh nổi ghép được truyền mà biểu thị L/R hoặc M/S như ví dụ đối với việc mã hóa âm thanh nổi ghép chung sẽ được sử dụng hay không. Trong bộ giải mã, trước tiên, tín hiệu lỗi được giải mã như được biểu thị bởi các cờ âm thanh nổi ghép cho các ô lỗi. Thứ hai, tín hiệu lỗi được lưu trữ trong phép biểu diễn L/R và M/S. Để điền đầy ô IGF, phép biểu diễn lát nguồn được chọn để khớp với phép biểu diễn lát đích như được biểu thị bởi thông tin âm thanh nổi ghép đối với các ô IGF.

Tạo hình nhiễu âm theo thời gian (TNS) là kỹ thuật tiêu chuẩn và là bộ phận của AAC trong các tài liệu tham khảo [11 - 13]. TNS có thể được xét đến như phần mở rộng sơ đồ cơ bản của bộ mã hóa theo cảm quan, chèn vào bước xử lý tùy ý giữa khâu lọc và giai đoạn lượng tử hoá. Nhiệm vụ chính của mô đun TNS là để che giấu nhiễu âm lượng tử hoá được tạo ra trong vùng màn chắn âm theo thời gian của các tín hiệu tương tự chuyển tiếp và do đó mô đun này tạo ra sơ đồ mã hóa hữu hiệu hơn. Trước tiên, TNS tính toán tập hợp các hệ số dự báo bằng cách sử dụng “dự báo chuyển tiếp” trong miền biến đổi, ví dụ MDCT. Các hệ số này sau đó được sử dụng để làm phẳng đường bao theo thời gian của tín hiệu. Do việc lượng tử hoá tác động đến phổ được lọc TNS, nhiễu âm lượng tử hoá cũng phẳng theo thời gian. Bằng cách áp dụng việc lọc TNS nghịch đảo ở phía bộ giải mã, nhiễu âm lượng tử hoá được tạo hình theo đường

bao theo thời gian của bộ lọc TNS và do đó nhiều âm lượng tử hoá được che giấu bằng sự chuyển tiếp.

IGF được dựa trên phép biểu diễn MDCT. Để mã hóa hữu hiệu, ưu tiên các khối dài khoảng 20ms phải được sử dụng. Nếu tín hiệu trong khối dài chứa các chuyển tiếp, âm vang trước và sau có thể nghe được xảy ra trong các dải phổ IGF do việc điền đầy ô. Fig.7c thể hiện tác dụng âm vang trước điển hình trước sự bắt đầu chuyển tiếp do IGF. Ở phía bên trái, ảnh phổ của tín hiệu ban đầu được thể hiện và ở phía bên phải, ảnh phổ của tín hiệu mở rộng bằng thông mà không có lọc TNS được thể hiện.

Hiệu ứng âm vang trước này được giảm bằng cách sử dụng TNS trong ngữ cảnh IGF. Ở đây, TNS được sử dụng làm công cụ tạo hình ô theo thời gian (TTS) do phép biểu diễn phổ trong bộ giải mã được thực hiện đối với tín hiệu dư TNS. Các hệ số dự báo TTS cần phải được tính toán và được áp dụng bằng cách sử dụng phổ đầy đủ ở phía bộ mã hoá như thông thường. Các tần số bắt đầu và kết thúc TNS/TTS không bị tác động bởi tần số bắt đầu IGF  $f_{IGFstart}$  của công cụ IGF. So với TNS truyền thống, tần số kết thúc TTS tăng đến tần số kết thúc của công cụ IGF, mà cao hơn so với  $f_{IGFstart}$ . Ở phía bộ giải mã, các hệ số TNS/TTS được áp dụng đối với phổ đầy đủ một lần nữa, nghĩa là phổ lỗi cộng với phổ tái tạo cộng với các thành phần âm từ bản đồ âm (xem Fig.7e). Việc áp dụng TTS là cần thiết để tạo thành đường bao theo thời gian của phổ tái tạo để phù hợp đường bao của tín hiệu ban đầu một lần nữa. Do đó, âm vang trước được thể hiện được giảm. Ngoài ra, vẫn tạo hình nhiều âm lượng tử hoá trong tín hiệu dưới  $f_{IGFstart}$  như thông thường với TNS.

Trong các bộ giải mã truyền thống, việc vá phổ đối với tín hiệu âm thanh làm sai lệch tương quan phổ ở các đường biên bản vá và do đó làm suy yếu đường bao theo thời gian của tín hiệu âm thanh bằng cách đưa đến độ tán sắc. Do đó, lợi ích khác của việc thực hiện điền đầy ô IGF đối với tín hiệu dư là ở chỗ, sau khi áp dụng bộ lọc tạo hình, các đường biên ô có tương quan một cách liên mạch, dẫn đến việc tái tạo theo thời gian trung thực hơn của tín hiệu.

Trong bộ mã hoá của sáng chế, phổ trải qua việc lọc TNS/TTS, việc xử lý màn chắn âm và ước lượng tham số IGF không có tín hiệu bất kỳ cao hơn tần số bắt đầu IGF ngoại trừ các thành phần âm. Sau đó, phổ thừa này được mã hóa bởi bộ mã hóa lỗi

bằng cách sử dụng các nguyên lý mã hóa và mã hóa dự báo số học. Các thành phần mã hóa này cùng với các bit tín hiệu hóa tạo thành dòng bit của âm thanh.

Fig.2a minh họa việc triển khai bộ giải mã tương ứng. Dòng bit trên Fig.2a tương ứng với tín hiệu âm thanh được mã hoá được đưa vào trong bộ khử trộn kênh/bộ giải mã mà sẽ được kết nối, so với Fig.1b, với các khối 112 và 114. Bộ khử phân kênh bit chia tách tín hiệu âm thanh đầu vào thành phép biểu diễn được mã hoá thứ nhất 107 của Fig.1b và phép biểu diễn được mã hoá thứ hai 109 của Fig.1b. Phép biểu diễn được mã hoá thứ nhất có tập hợp thứ nhất của các phần phổ thứ nhất được đưa vào trong khối giải mã kênh ghép 204 tương ứng với bộ giải mã miền phổ 112 của Fig.1b. Phép biểu diễn được mã hoá thứ hai được đưa vào trong bộ giải mã theo tham số 114 không được minh họa trên Fig.2a và sau đó được đưa vào trong khối IGF 202 tương ứng với bộ tái tạo tần số 116 của Fig.1b. Tập hợp thứ nhất của các phần phổ thứ nhất đòi hỏi việc tái tạo tần số được đưa vào trong khối IGF 202 qua đường dẫn 203. Hơn nữa, sau khi giải mã kênh ghép 204, việc giải mã lõi đặc hiệu được áp dụng trong khối màn chắn âm 206 sao cho đầu ra của màn chắn âm 206 tương ứng với đầu ra của bộ giải mã miền phổ 112. Khi đó, sự tổ hợp bởi bộ tổ hợp 208 được thực hiện, nghĩa là xây dựng khung trong đó đầu ra của bộ tổ hợp 208 giờ đây có phổ khoảng đầy đủ, nhưng vẫn trong miền được lọc TNS/TTS. Sau đó, trong khối 210, hoạt động TNS/TTS nghịch đảo được thực hiện bằng cách sử dụng thông tin bộ lọc TNS/TTS được cung cấp qua đường dẫn 109, nghĩa là thông tin phụ TTS ưu tiên được bao gồm trong phép biểu diễn được mã hoá thứ nhất được tạo ra bởi bộ mã hoá miền phổ 106 mà có thể, ví dụ, là AAC thẳng hoặc bộ mã hoá lõi USAC hoặc cũng có thể được bao gồm trong phép biểu diễn được mã hoá thứ hai. Ở đầu ra của khối 210, phổ hoàn thiện cho đến khi tần số tối đa được cung cấp mà là tần số khoảng đầy đủ được xác định bởi tốc độ lấy mẫu của tín hiệu đầu vào ban đầu. Sau đó, sự chuyển đổi phổ/thời gian được thực hiện trong giàn lọc tổng hợp 212 để cuối cùng thu được tín hiệu đầu ra âm thanh.

Fig.3a minh họa phép biểu diễn theo sơ đồ của phổ. Phổ được chia theo các dải hệ số tỷ lệ SCB trong đó có bảy dải hệ số tỷ lệ từ SCB1 đến SCB7 trong ví dụ được minh họa của Fig.3a. Các dải hệ số tỷ lệ có thể là các dải hệ số tỷ lệ AAC mà được xác định theo tiêu chuẩn AAC và có băng thông tăng dần đến các tần số cao như được minh họa trên Fig.3a theo sơ đồ. Tốt hơn là thực hiện việc điền đầy khoảng trống

thông minh không từ đầu của phổ, nghĩa là tại các tần số thấp, nhưng để bắt đầu hoạt động IGF tại tần số bắt đầu IGF được minh họa ở 309. Do đó, dải tần số lõi mở rộng từ tần số thấp nhất đến tần số bắt đầu IGF. Cao hơn tần số bắt đầu IGF, phép phân tích phổ được áp dụng để chia 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 được nhập ví dụ đưa vào bộ mã hoá miền phổ 106 hoặc bộ lập kênh ghép 228, nghĩa là bộ mã hoá lõi hoạt động trong khoảng đầy đủ, nhưng mã hoá số lượng đáng kể của các giá trị phổ bằng 0, nghĩa là, các giá trị phổ bằng 0 này được lượng tử hoá đến 0 hoặc được đặt thành 0 trước khi lượng tử hoá hoặc sau khi lượng tử hoá. Theo cách bất kỳ, bộ mã hoá lõi hoạt động trong khoảng đầy đủ, nghĩa là khi phổ như được minh họa, nghĩa là bộ giải mã lõi không nhất thiết phải nhận biết về việc điền đầy khoảng trống thông minh bất kỳ hoặc mã hoá của tập hợp thứ hai của các phần phổ thứ hai với độ phân giải phổ thấp hơn.

Tốt hơn là, độ phân giải cao được xác định bằng cách mã hóa theo đường dẫn của các đường dẫn phổ như các đường dẫn MDCT, trong khi độ phân giải thứ hai hoặc độ phân giải thấp được xác định, ví dụ, chỉ bằng cách tính toán giá trị phổ đơn lẻ cho mỗi dải hệ số tỷ lệ, trong đó dải hệ số tỷ lệ bao phủ vài đường dẫn tần số. Do đó, độ phân giải thấp thứ hai, so với độ phân giải phổ của nó, thấp hơn nhiều so với độ phân giải thứ nhất hoặc cao được xác định bằng cách mã hóa theo đường dẫn thường được áp dụng bởi bộ mã hoá lõi như bộ mã hoá lõi AAC hoặc USAC.

Liên quan đến hệ số tỷ lệ hoặc việc tính toán năng lượng, trạng thái được minh họa trên Fig.3b. Do thực tế rằng bộ mã hoá là bộ mã hoá lõi và do thực tế rằng có thể, nhưng không nhất thiết phải, có các thành phần của tập hợp thứ nhất của các phần phổ trong mỗi dải, bộ mã hoá lõi tính toán hệ số tỷ lệ đối với mỗi dải không chỉ trong khoảng lõi thấp hơn tần số bắt đầu IGF 309, mà còn cao hơn tần số bắt đầu IGF cho tới tần số tối đa  $f_{IGF_{stop}}$  mà nhỏ hơn hoặc bằng một nửa tần số lấy mẫu, nghĩa là,  $f_s/2$ . Do đó, các phần âm được mã hoá 302, 304, 305, 306, 307 của Fig.3a và theo phương án này cùng với các hệ số tỷ lệ từ SCB1 đến SCB7 tương ứng với dữ liệu phổ có 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 các giá trị thông tin năng lượng  $E_1, E_2, E_3, E_4$ , mà được truyền cùng với các hệ số tỷ lệ từ SF4 đến SF7.

Đặc biệt, khi bộ mã hoá lỗi trong điều kiện dòng bit thấp, hoạt động điền đầy nhiều âm bổ sung trong ô lỗi, nghĩa là thấp hơn trong tần số so với tần số bắt đầu IGF, nghĩa là trong các dải hệ số tỷ lệ từ SCB1 đến SCB3 có thể được áp dụng bổ sung. Trong điền đầy nhiều âm, có tồn tại vài đường dẫn phổ liên kề mà đã được lượng tử hoá đến 0. Ở phía bộ giải mã, có các giá trị phổ được lượng tử hoá đến 0 này được tổng hợp lại và các giá trị phổ được tổng hợp lại được điều chỉnh trong biên độ của chúng bằng cách sử dụng năng lượng điền đầy nhiều âm như  $NF_2$  được minh họa ở 308 trên Fig.3b. Năng lượng điền đầy nhiều âm, mà có thể được đưa ra trong các thuật ngữ tuyệt đối hoặc trong các thuật ngữ tương đối đặc biệt đối với hệ số tỷ lệ như trong USAC tương ứng với năng lượng của tập hợp các giá trị phổ được lượng tử hoá đến 0. Các đường dẫn phổ điền đầy nhiều âm này cũng có thể được xét đến là tập hợp thứ ba của các phần phổ thứ ba mà được tái tạo bằng tổng hợp điền đầy nhiều âm thẳng mà không có hoạt động IGF bất kỳ dựa trên tái tạo tần số bằng cách sử dụng các lát tần số từ các tần số khác để khôi phục các lát tần số bằng cách sử dụng các giá trị phổ từ khoảng nguồn và thông tin năng lượng  $E_1, E_2, E_3, E_4$ .

Tốt hơn là, các dải mà thông tin năng lượng được tính toán trùng khớp với các dải hệ số tỷ lệ. Theo các phương án khác, việc nhóm giá trị thông tin năng lượng được áp dụng sao cho, ví dụ, đối với các dải hệ số tỷ lệ 4 và 5, chỉ có giá trị thông tin năng lượng đơn lẻ được truyền, nhưng ngay cả theo phương án này, các đường biên của các dải khôi phục được nhóm trùng khớp với các đường biên của các dải hệ số tỷ lệ. Nếu các phép chia tách dải khác nhau được áp dụng, thì các phép tính toán lại nhất định hoặc các tính toán đồng bộ hoá có thể được áp dụng và điều này có thể có ý nghĩa phụ thuộc vào việc triển khai nhất định.

Tốt hơn là, bộ mã hoá miền phổ 106 của Fig.1a là bộ mã hoá dựa trên tâm thính học như được minh họa trên Fig.4a. Thông thường, đối với ví dụ được minh họa trong tiêu chuẩn MPEG2/4 AAC hoặc MPEG1/2, tiêu chuẩn Lớp 3, sẽ là tín hiệu âm thanh được mã hoá sau khi đã được biến đổi thành khoảng phổ (401 trên Fig.4a) được chuyển tiếp đến bộ tính toán hệ số tỷ lệ 400. Bộ tính toán hệ số tỷ lệ được điều khiển

bởi mô hình tâm thính học còn nhận tín hiệu âm thanh được lượng tử hoá hoặc nhận, như trong tiêu chuẩn MPEG1/2 lớp 3 hoặc MPEG AAC, phép biểu diễn phổ phức hợp của tín hiệu âm thanh. Mô hình tâm thính học tính toán, đối với mỗi dải hệ số tỷ lệ, hệ số tỷ lệ biểu diễn ngưỡng tâm thính học. Ngoài ra, các hệ số tỷ lệ sau đó, nhờ sự hợp tác của các vòng lặp bên trong hoặc bên ngoài đã biết hoặc bằng quy trình giải mã thích hợp khác bất kỳ được điều chỉnh sao cho các điều kiện tốc độ bit nhất định được đáp ứng hoàn toàn. Sau đó, một mặt, các giá trị phổ được lượng tử hoá và mặt khác các hệ số tỷ lệ được tính toán được đưa vào trong bộ xử lý lượng tử hoá 404. Trong hoạt động bộ mã hoá âm thanh thẳng, các giá trị phổ lượng tử hoá được gán trọng số theo các hệ số tỷ lệ và sau đó các giá trị phổ trọng số được đưa vào trong bộ lượng tử hoá cố định thường có chức năng nén tới các khoảng biên độ phía trên. Sau đó, tại đầu ra của bộ xử lý lượng tử hoá, có các chỉ số lượng tử hoá tồn tại mà sau đó được chuyển tiếp vào trong bộ mã hoá entropy thường có sự mã hóa đặc hiệu và rất hữu hiệu đối với tập hợp của các chỉ số lượng tử hoá tới không đối với các giá trị tần số liền kề hoặc ngoài ra như được kể đến trong lĩnh vực này, “tiền trình” của các giá trị bằng không.

Tuy nhiên, trong bộ mã hoá âm thanh của Fig.1a, bộ xử lý bộ lượng tử hoá thường nhận thông tin đối với các phần phổ thứ hai từ bộ phân tích phổ. Do đó, bộ xử lý bộ lượng tử hoá 404 đảm bảo rằng, trong đầu ra của bộ xử lý bộ lượng tử hoá 404, các phần phổ thứ hai như được nhận biết bởi bộ phân tích phổ 102 là bằng 0 hoặc có phép biểu diễn được bộ mã hóa hoặc bộ giải mã thừa nhận là phép biểu diễn 0 mà có thể được mã hóa rất hiệu quả, đặc biệt là khi tồn tại các “tiền trình” của các giá trị 0 tồn tại trong phổ.

Fig.4b minh họa việc triển khai bộ xử lý bộ lượng tử hoá. Các giá trị phổ MDCT có thể được đưa vào trong tập hợp đến khối bằng 0 410. Sau đó, các phần phổ thứ hai được đặt thành 0 trước khi lấy trọng số theo các hệ số tỷ lệ trong khối 412 được thực hiện. Theo phương án bổ sung, khối 410 không được cung cấp, nhưng việc hợp tác đặt thành 0 được thực hiện trong khối 418 tiếp theo khối trọng số 412. Theo phương án khác, hoạt động đặt thành không cũng có thể được thực hiện trong khối đặt thành không 422 tiếp theo sự lượng tử hoá trong khối bộ lượng tử hoá 420. Theo phương án này, các khối 410 và 418 sẽ không có mặt. Thông thường, ít nhất một trong các khối 410, 418, 422 được cung cấp phụ thuộc vào việc triển khai đặc hiệu.



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

Các khối đặt thành không 410, 418, 422, mà được cung cấp theo một cách khác cùng nhau hoặc theo cách song song được điều khiển bởi bộ phân tích phổ 424. Bộ phân tích phổ ưu tiên bao gồm việc thực thi bất kỳ của bộ phát hiện theo âm điệu đã được biết đến đến hoặc bao gồm loại khác bất kỳ của bộ phát hiện hoạt động để tách phổ thành các thành phần sẽ được mã hoá với độ phân giải cao và các thành phần sẽ được mã hoá với độ phân giải thấp. Các thuật toán khác như vậy được thực thi trong bộ phân tích phổ có thể là bộ phát hiện hoạt động giọng nói, bộ phát hiện nhiều âm, bộ phát hiện tiếng nói hoặc bộ phát hiện khác bất kỳ quyết định, phụ thuộc vào thông tin phổ hoặc siêu dữ liệu ghép đối với các yêu cầu độ phân giải đối với các phần phổ khác nhau.

Fig.5a minh họa phương án được ưu tiên của bộ chuyển đổi phổ thời gian 100 của Fig.1a như, ví dụ, được triển khai trong AAC hoặc USAC. Bộ chuyển đổi phổ thời gian 100 bao gồm bộ tạo cửa sổ 502 được điều khiển bởi bộ phát hiện chuyển tiếp 504. Khi bộ phát hiện chuyển tiếp 504 phát hiện sự chuyển tiếp, thì sự luân phiên từ các cửa sổ dài thành các cửa sổ ngắn được tín hiệu hóa đến bộ tạo cửa sổ. Sau đó, bộ tạo cửa sổ 502 tính toán, đối với các khối chồng lấp, các khung cửa sổ, trong đó mỗi khung cửa sổ thường có hai giá trị N như các giá trị 2048. Sau đó, việc biến đổi trong bộ biến đổi khối 506 được thực hiện, và bộ biến đổi khối này thường cung cấp tiếp cho sự tiêu hao, sao cho sự tiêu hao/biến đổi tổ hợp được thực hiện để thu được khung phổ với các giá trị N như các giá trị phổ MDCT. Do đó, đối với hoạt động cửa sổ dài, khung ở đầu ra của khối 506 bao gồm hai giá trị N như các giá trị 2048 và sau đó khung phổ có các giá trị 1024. Tuy nhiên, sau đó việc chuyển đổi được thực hiện đến các khối ngắn, khi tám khối ngắn được thực hiện trong đó mỗi khối ngắn có 1/8 các giá trị miền thời gian cửa sổ so với cửa sổ dài và mỗi khối phổ có 1/8 các giá trị phổ so với khối dài. Do đó, khi sự tiêu hao này được tổ hợp với 50% hoạt động chồng lấp của bộ tạo cửa sổ, phổ là phiên bản lấy mẫu tới hạn của tín hiệu âm thanh miền thời gian 99.

Sau đây, sự tham khảo được thực hiện đến Fig.5b minh họa việc triển khai đặc hiệu của bộ tái tạo tần số 116 và bộ chuyển đổi phổ-thời gian 118 của Fig.1b, hoặc của hoạt động tổ hợp của các khối 208, 212 của Fig.2a. Trên Fig.5b, dải khôi phục đặc hiệu được xét đến như dải hệ số tỷ lệ 6 của Fig.3a. Phần phổ thứ nhất trong dải khôi phục này, nghĩa là phần phổ thứ nhất 306 của Fig.3a được đưa vào trong bộ tích lũy/bộ điều chỉnh khung 510. Hơn nữa, phần phổ thứ hai được khôi phục đối với dải hệ số tỷ lệ 6 cũng được đưa vào trong bộ tích lũy/bộ điều chỉnh khung 510. Hơn nữa, thông tin năng lượng như  $E_3$  của Fig.3b đối với dải hệ số tỷ lệ 6 cũng được đưa vào trong khối 510. Phần phổ thứ hai được khôi phục trong dải khôi phục đã được tạo ra bằng việc điền đầy lát tần số bằng cách sử dụng khoảng nguồn và khi đó dải khôi phục tương ứng với khoảng đích. Giờ đây, việc điều chỉnh năng lượng của khung được thực hiện để sau đó cuối cùng thu được khung khôi phục hoàn toàn có các giá trị N như, ví dụ, thu được ở đầu ra của bộ tổ hợp 208 của Fig.2a. Sau đó, trong khối 512, sự biến đổi/phép nội suy khối nghịch đảo được thực hiện để thu được các giá trị miền thời gian 248 đối với, ví dụ, các giá trị phổ 124 ở đầu vào của khối 512. Sau đó, hoạt động của sổ tổng hợp được thực hiện trong khối 514 mà một lần nữa được điều khiển bởi biểu thị của sổ dài/cửa sổ ngắn được truyền như thông tin phụ trong tín hiệu âm thanh được mã hoá. Sau đó, trong khối 516, hoạt động chồng lấp/bổ sung với khung thời gian trước được thực hiện. Tốt hơn là, MDCT áp dụng 50% chồng lấp, sao cho đối với mỗi khung thời gian mới có các giá trị  $2N$ , các giá trị miền thời gian N cuối cùng là đầu ra. 50% chồng lấp được ưu tiên hơn do thực tế rằng sự chồng lấp này cung cấp việc lấy mẫu tới hạn và giao nhau liên tiếp từ một khung đến khung tiếp theo do hoạt động chồng lấp/bổ sung trong khối 516.

Như được minh họa ở 301 trên Fig.3a, hoạt động điền đầy nhiễu âm còn có thể được áp dụng không chỉ thấp hơn tần số bắt đầu IGF, mà còn trên tần số bắt đầu IGF như đối với dải khôi phục được dự định trùng khớp với dải hệ số tỷ lệ 6 của Fig.3a. Sau đó, các giá trị phổ điền đầy nhiễu âm cũng có thể được đưa vào bộ tích lũy/bộ điều chỉnh khung 510 và sự điều chỉnh các giá trị phổ điền đầy nhiễu âm cũng có thể được áp dụng trong khối này hoặc các giá trị phổ điền đầy nhiễu âm có thể được điều chỉnh bằng cách sử dụng năng lượng điền đầy nhiễu âm trước khi được đưa vào trong bộ tích lũy/bộ điều chỉnh khung 510.

Tốt hơn là, hoạt động IGF, nghĩa là hoạt động điền đầy lát tần số bằng cách sử dụng các giá trị phổ từ các phần khác có thể được áp dụng trong phổ hoàn thiện. Do đó, hoạt động điền đầy ô phổ có thể không chỉ được áp dụng trong dải cao hơn tần số bắt đầu IGF mà cũng có thể được áp dụng trong dải thấp. Hơn nữa, việc điền đầy nhiều âm mà không có điền đầy lát tần số cũng có thể được áp dụng không chỉ thấp hơn tần số bắt đầu IGF mà còn cao hơn tần số bắt đầu IGF. Tuy nhiên, các tác giả sáng chế đã phát hiện ra rằng, việc mã hoá âm thanh chất lượng cao và hữu hiệu ở mức cao có thể thu được khi hoạt động điền đầy nhiều âm bị giới hạn ở khoảng tần số thấp hơn tần số bắt đầu IGF và khi hoạt động điền đầy lát tần số được giới hạn trong khoảng tần số cao hơn tần số bắt đầu IGF như được minh họa trên Fig.3a.

Tốt hơn là, các lát đích (target tiles - TT) (có các tần số lớn hơn tần số bắt đầu IGF) được liên kết với các đường biên dải hệ số tỷ lệ của bộ mã hóa tốc độ đầy đủ. Các lát nguồn (source tiles - ST), mà từ đó thông tin thu được, nghĩa là, đối với các tần số thấp hơn so với tần số bắt đầu IGF không được liên kết bởi các đường biên dải hệ số tỷ lệ. Kích cỡ của ST sẽ tương ứng với kích cỡ của TT ghép. Việc này được minh họa bằng cách sử dụng ví dụ sau đây. TT[0] có chiều dài 10MDCT Ngăn. Kích cỡ này tương ứng một cách chính xác với chiều dài của hai SCB nối tiếp (như  $4 + 6$ ). Khi đó, tất cả các ST có thể mà tương quan với TT[0], cũng có chiều dài 10 ngăn. Lát đích thứ hai TT[1] liên kết với TT[0] có chiều dài 15 ngăn 1 (SCB có chiều dài  $7 + 8$ ). Khi đó, ST có chiều dài 15 ngăn chứ không phải 10 ngăn như đối với TT[0].

Trong trường hợp phát sinh rằng có thể không phát hiện ra TT đối với ST có chiều dài của lát đích (khi ví dụ chiều dài của TT lớn hơn so với khoảng nguồn sẵn có), khi đó sự tương quan không được tính toán và khoảng nguồn được sao chép một số lần vào trong TT này (việc sao chép này được thực hiện lần lượt sao cho đường dẫn tần số đối với tần số thấp nhất của bản sao thứ hai tiếp theo ngay sau – theo tần số – đường dẫn tần số đối với tần số cao nhất của bản sao thứ nhất), cho đến khi lát đích TT được điền đầy hoàn toàn.

Tiếp theo, việc tham khảo được thực hiện đến Fig.5c minh họa phương án được ưu tiên khác của bộ tái tạo tần số 116 của Fig.1b hoặc khối IGF 202 của Fig.2a. Khối 522 là bộ tạo lát tần số nhận, không chỉ lát đích ID, mà còn nhận lát nguồn ID. Ví dụ,

các tác giả sáng chế đã xác định dựa trên phía bộ mã hoá rằng dải hệ số tỷ lệ 3 của Fig.3a là rất thích hợp để khôi phục dải hệ số tỷ lệ 7. Do đó, lát nguồn ID sẽ bằng 2 và lát đích ID sẽ bằng 7. Dựa trên thông tin này, bộ tạo lát tần số 522 áp dụng sao chép hoặc hoạt động điền đầy ô sóng hài hoặc hoạt động điền đầy ô khác bất kỳ để tạo ra phần thứ hai thô của các thành phần phổ 523. Phần thứ hai thô của các thành phần phổ có độ phân giải tần số giống với độ phân giải tần số được bao gồm trong tập hợp thứ nhất của các phần phổ thứ nhất.

Khi đó, phần phổ thứ nhất của dải khôi phục như 307 của Fig.3a được đưa vào trong bộ tích lũy khung 524 và phần thứ hai thô 523 cũng được đưa vào trong bộ tích lũy khung 524. Sau đó, khung khôi phục được điều chỉnh bằng bộ điều chỉnh 526 bằng cách sử dụng hệ số khuếch đại đối với dải khôi phục được tính toán bằng bộ tính toán hệ số khuếch đại 528. Tuy nhiên, quan trọng là, phần phổ thứ nhất trong khung không bị tác động bởi bộ điều chỉnh 526, nhưng chỉ phần thứ hai thô đối với khung khôi phục bị tác động bởi độ điều chỉnh 526. Với mục tiêu này, bộ tính toán hệ số khuếch đại 528 phân tích dải nguồn hoặc phần thứ hai thô 523 và còn phân tích phần phổ thứ nhất trong dải khôi phục để cuối cùng tìm thấy hệ số khuếch đại 527 chính xác sao cho năng lượng của khung điều chỉnh được xuất ra bởi bộ điều chỉnh 526 có năng lượng  $E_4$  khi dải hệ số tỷ lệ 7 được dự định.

Ở đây, quan trọng là đánh giá độ chính xác khôi phục tần số cao của sáng chế so với HE-AAC. Việc này được giải thích đối với dải hệ số tỷ lệ 7 trên Fig.3a. Các tác giả sáng chế đã giả định rằng bộ mã hoá của kỹ thuật ưu tiên như được minh họa trên Fig.13a sẽ phát hiện phần phổ 307 sẽ được mã hoá với độ phân giải cao như “sóng hài bị mất”. Khi đó, năng lượng của thành phần phổ này sẽ được truyền cùng với thông tin đường bao phổ đối với dải khôi phục như dải hệ số tỷ lệ 7 vào bộ giải mã. Sau đó, bộ giải mã sẽ tái tạo sóng hài bị mất. Tuy nhiên, giá trị phổ, mà ở đó sóng hài bị mất 307 sẽ được khôi phục bởi bộ giải mã của kỹ thuật ưu tiên của Fig.13b sẽ ở giữa của dải 7 ở tần số được biểu thị bởi tần số khôi phục 390. Do đó, sáng chế tránh được sai số tần số 391 mà có thể được tạo ra bởi bộ giải mã của kỹ thuật ưu tiên của Fig.13d.

Theo phương án, bộ phân tích phổ cũng được triển khai để tính toán tính tương tự giữa các phần phổ thứ nhất và các phần phổ thứ hai và để xác định, dựa trên tính

tương tự được tính toán, đối với phần phổ thứ hai trong khoảng khôi phục phần phổ thứ nhất phù hợp với phần phổ thứ hai nhiều nhất có thể. Khi đó, khi triển khai khoảng nguồn biến đổi/khoảng điểm đến, bộ mã hóa theo tham số sẽ còn đưa đến trong phép biểu diễn được mã hoá thứ hai thông tin so khớp biểu thị đối với mỗi khoảng điểm đến khoảng nguồn phù hợp. Ở phía bộ giải mã, thông tin này sau đó được sử dụng bởi bộ tạo lát tần số 522 của Fig.5c minh họa việc tạo ra phần thứ hai thô 523 dựa trên dải nguồn ID và dải đích ID.

Hơn nữa, như được minh họa trên Fig.3a, bộ phân tích phổ được tạo cấu hình để phân tích phép biểu diễn phổ đến tần số phân tích tối đa chỉ là lượng nhỏ thấp hơn một nửa tần số lấy mẫu và ưu tiên bằng ít nhất một phần tư tần số lấy mẫu hoặc thường là cao hơn.

Như được minh họa, bộ mã hoá hoạt động mà không lấy mẫu giảm và bộ giải mã hoạt động mà không lấy mẫu tăng. Theo cách khác, bộ mã hoá 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.

Hơn nữa, như được minh họa trên Fig.3a, bộ phân tích phổ được tạo cấu hình để phân tích phép biểu diễn phổ bắt đầu với tần số bắt đầu điền đầy khoảng trống và kết thúc với tần số tối đa được biểu diễn bởi tần số tối đa được bao gồm trong phép biểu diễn phổ, trong đó phần phổ mở rộng từ tần số tối thiểu đến tần số bắt đầu điền đầy khoảng trống thuộc tập hợp thứ nhất của các phần phổ và trong đó phần phổ khác như 304, 305, 306, 307 có các giá trị tần số cao hơn tần số điền đầy khoảng trống còn được bao gồm trong tập hợp thứ nhất của các phần phổ thứ nhất.

Như được mô tả, bộ giải mã âm thanh miền phổ 112 được tạo cấu hình sao cho tần số tối đa được biểu diễn bởi giá trị phổ trong phép biểu diễn được giải mã thứ nhất bằng với tần số tối đa được bao gồm trong phép biểu diễn thời gian có tốc độ lấy mẫu trong đó giá trị phổ đối với tần số tối đa trong tập hợp thứ nhất của các phần phổ thứ nhất là bằng 0 hoặc khác 0. Dù sao, đối với tần số tối đa này trong tập hợp thứ nhất của các thành phần phổ, hệ số tỷ lệ đối với dải hệ số tỷ lệ tồn tại, mà được tạo ra và được truyền không kể tất cả giá trị phổ trong dải hệ số tỷ lệ này được đặt thành 0 hay không hoặc không như được mô tả trong nội dung của các hình vẽ Fig.3a và Fig.3b.

Do đó, sáng chế có ưu điểm so với các kỹ thuật tham số khác là làm tăng hệ số nén, ví dụ, sự thay thế nhiều âm và điền đầy nhiều âm (các kỹ thuật này dành riêng cho phép biểu diễn hữu hiệu của nhiều âm tương tự nội dung tín hiệu cục bộ) sáng chế cho phép tái tạo tần số chính xác của các thành phần âm. Cho đến nay, không có kỹ thuật theo tình trạng kỹ thuật nào giải quyết phép biểu diễn theo tham số hữu hiệu của nội dung tín hiệu tùy ý bằng việc điền đầy khoảng trống phổ mà không hạn chế sự phân chia trước cố định trong dải thấp (LF) và dải cao (HF).

Các phương án của hệ thống của sáng chế cải thiện các phương pháp theo tình trạng kỹ thuật và do đó tạo ra hiệu quả nén cao, không có hoặc chỉ có sự khó chịu cảm quan nhỏ và băng thông âm thanh đầy đủ thậm chí đối với tốc độ bit thấp.

Hệ thống thông thường gồm có

- mã hóa lõi dải đầy đủ
- điền đầy khoảng trống thông minh (điền đầy ô hoặc điền đầy nhiều âm)
- các phần tiếng nói thừa trong lõi được chọn bởi màn chắn âm
- mã hóa cặp âm thanh nổi ghép đối với dải đầy đủ, bao gồm điền đầy ô
- TNS trên ô
- làm trắng phổ trong khoảng IGF

Bước thứ nhất để hệ thống hữu hiệu hơn là để loại bỏ nhu cầu để biến đổi các dữ liệu phổ vào trong miền biến đổi thứ hai khác với một trong các bộ mã hóa lõi. Như đã số các bộ mã hóa-giải mã âm thanh, như AAC ví dụ, sử dụng MDCT làm biến đổi cơ sở, sẽ hữu ích để thực hiện BWE trong miền MDCT. Yêu cầu thứ hai đối với hệ thống BWE là nhu cầu bảo toàn lưới âm nhờ đó ngay cả các thành phần âm HF cũng được bảo toàn và do đó chất lượng âm thanh được mã hóa tốt hơn so với các hệ thống đang tồn tại. Để đáp ứng được các yêu cầu được kể đến trên đây đối với sơ đồ BWE, hệ thống mới được đề xuất được gọi là điền đầy khoảng trống thông minh (IGF). Fig.2b thể hiện biểu đồ khối của hệ thống được đề xuất ở phía bộ mã hoá và Fig.2a thể hiện hệ thống ở phía bộ giải mã.

Fig.6a minh họa thiết bị để giải mã tín hiệu âm thanh được mã hoá theo phương án 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 như bộ tái tạo tần số 604 được kết nối dòng giảm của bộ giải mã âm thanh miền phổ 602 để tạo ra phần phổ thứ hai được khôi phục bằng cách 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 trong phần phổ thứ hai là các giá trị dự báo phổ. Để biến đổi các giá trị dự báo phổ này thành phép biểu diễn phổ đầy đủ, bộ lọc dự báo phổ 606 được đề xuất. Bộ lọc dự báo nghịch đảo này được tạo cấu hình để thực hiện việc dự báo nghịch đảo trong tần số sử dụng các giá trị dự phổ đối với tập hợp thứ nhất của tần số thứ nhất và khôi phục các phần phổ thứ hai. Bộ lọc dự báo nghịch đảo phổ 606 được tạo cấu hình bởi thông tin bộ lọc được bao gồm trong tín hiệu âm thanh được mã hoá. Fig.6b minh họa việc triển khai chi tiết hơn của phương án theo Fig.6a. Giá trị dự báo phổ 603 được đưa vào trong bộ tạo lát tần số 612 tạo ra các giá trị phổ thô đối với dải khôi phục hoặc đối với phần tần số thứ hai nhất định và dữ liệu thô này giờ đây có cùng độ phân giải khi phép biểu diễn phổ độ phân giải cao được đưa vào trong bộ tạo hình phổ 614. Bộ tạo hình phổ giờ đây tạo hình phổ bằng cách sử dụng thông tin đường bao được truyền trong dòng bit và sau đó dữ liệu được tạo hình theo phổ được áp dụng vào bộ lọc dự báo phổ 616 cuối cùng tạo ra khung chứa các giá trị phổ đầy đủ bằng cách sử dụng thông tin bộ lọc 607 được truyền từ bộ mã hoá đến bộ giải mã qua dòng bit.

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

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

Ở phía bộ giải mã, dữ liệu mã hoá được đưa vào trong bộ khử trộn kênh 720 để chia tách một mặt, thông tin phụ IGF, mặt khác thông tin phụ TTS và phép biểu diễn được mã hoá 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 các phổ giá trị thực. Sau đó, cả hai phổ giá trị thực và phức hợp được đưa vào trong khối 726 để tạo ra các giá trị tần số khôi phục trong tập hợp thứ hai của các phần phổ thứ hai đối với dải khôi phục. Sau đó, ở khung thu được đầy đủ và khung dải đầy đủ được điền đầy ô, hoạt động TTS nghịch đảo 728 được thực hiện và, ở phía bộ giải mã, hoạt động MDCT phức hợp nghịch đảo cuối cùng được thực hiện trong khối 730. Do đó, việc sử dụng thông tin bộ lọc TNS phức hợp cho phép, khi được áp dụng không chỉ trong dải lõi hoặc trong các dải ô riêng biệt mà còn được áp dụng qua các đường biên lõi/ô hoặc các đường biên ô/ô một cách tự động tạo ra xử lý đường biên ô, mà cuối cùng, lại đưa đến sự tương quan phổ giữa các ô. Sự tương quan phổ này qua các đường biên ô không thu được chỉ bằng cách tạo ra các lát tần số và thực hiện điều chỉnh đường bao phổ đối với dữ liệu thô này của các lát tần số.

Fig.7c minh họa phép so sánh tín hiệu ban đầu (dải bên trái) và tín hiệu mở rộng mà không có TTS. Có thể thấy được rằng có các giả âm mạnh được minh họa bởi các phần mở rộng trong khoảng tần số trên được minh họa ở 750. Tuy nhiên, việc này không xảy ra trên Fig.7e khi phần phổ tương tự ở 750 được so sánh với thành phần liên quan đến giả âm 750 của Fig.7c.



Các phương án hoặc hệ thống mã hóa âm thanh của sáng chế sử dụng phần chung chủ yếu của tốc độ bit sẵn có để mã hóa dạng sóng chỉ cấu trúc liên quan nhất theo cảm quan của tín hiệu trong bộ mã hoá và các khoảng trống phổ thu được được điền đầy trong bộ giải mã với nội dung tín hiệu mà gần bằng phổ ban đầu. Ngân sách bit rất bị giới hạn được dùng để điều khiển việc gọi là điền đầy khoảng trống thông minh (IGF) phổ theo tham số bởi thông tin phụ chuyên dụng được truyền từ bộ mã hoá đến bộ giải mã.

Việc lưu trữ hoặc truyền các tín hiệu âm thanh thường phải chịu các ràng buộc tốc độ bit nghiêm ngặt. Trước đây, các bộ mã hóa buộc phải giảm mạnh băng thông âm thanh được truyền khi chỉ có sẵn tốc độ bit rất thấp. Ngày nay, các bộ mã hóa-giải mã âm thanh hiện đại có thể mã hóa các tín hiệu dải rộng bằng cách sử dụng phương pháp mở rộng băng thông (BWE) tương tự sao chép băng thông phổ (SBR) [1]. Các thuật toán này dựa trên phép biểu diễn tham số của nội dung tần số cao (HF) - mà được tạo ra từ phần tần số thấp (LF) được mã hóa dạng sóng của tín hiệu được giải mã bằng phương pháp chuyển vị vào trong vùng phổ HF (“vá”) và áp dụng việc xử lý sau bằng tham số. Trong các sơ đồ BWE, sự khôi phục vùng phổ HF cao hơn tần số chéo đã cho thường được dựa trên phép vá phổ. Thông thường, vùng HF gồm có nhiều bản vá liền kề và mỗi trong số các bản vá này có nguồn gốc từ các vùng thông dải (BP) của phổ LF thấp hơn tần số chéo đã cho. Các hệ thống theo tình trạng kỹ thuật thực hiện một cách hữu hiệu việc vá trong phép biểu diễn giàn lọc bằng cách sao chép tập hợp các hệ số băng con liền kề từ vùng nguồn đến vùng đích.

Nếu hệ thống BWE được triển khai trong giàn lọc hoặc miền biến đổi thời gian-tần số, chỉ có khả năng hạn chế để điều khiển dạng theo thời gian của tín hiệu mở rộng băng thông. Thông thường, độ kết hạt theo thời gian bị giới hạn bởi kích cỡ nhảy được sử dụng giữa các cửa sổ biến đổi liền kề. Việc này có thể dẫn đến các âm vang trước hoặc sau không mong muốn trong khoảng phổ BWE.

Dựa vào việc mã hóa âm thanh theo cảm quan, các tác giả sáng chế đã biết được rằng dạng đường bao theo thời gian của tín hiệu âm thanh có thể được khôi phục bằng cách sử dụng các kỹ thuật lọc phổ tương tự tạo hình đường bao theo thời gian (TNS) [14]. Tuy nhiên, bộ lọc TNS theo tình trạng kỹ thuật được biết đến là bộ lọc giá trị

thực trên phổ giá trị thực. Bộ lọc giá trị thực này trên các phổ giá trị thực có thể bị suy yếu nghiêm trọng bởi các giả âm bí danh, đặc biệt là nếu việc biến đổi thực cơ bản là biến đổi cosin rời rạc cải biên (MDCT).

Tạo hình ô đường bao theo thời gian áp dụng việc lọc phức hợp trên các phổ giá trị phức hợp, giống như thu được từ, ví dụ, biến đổi cosin rời rạc cải biên phức hợp (CMDCT). Do đó, các giả âm bí danh được ngăn ngừa.

Việc tạo hình ô theo thời gian gồm có

- ước lượng hệ số bộ lọc phức hợp và áp dụng bộ lọc phẳng trên phổ tín hiệu ban đầu ở bộ mã hoá
- truyền các hệ số bộ lọc vào 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 theo tình trạng kỹ thuật đã biết từ việc mã hóa biến đổi âm thanh, đặc biệt là tạo hình nhiều âm theo thời gian (TNS) bằng sự dự báo tuyến tính dọc theo hướng tần số, để sử dụng theo cách cải biến trong nội dung của mở rộng băng thông.

Hơn nữa, thuật toán mở rộng băng thông của sáng chế dựa trên phép điền đầy khoảng trống thông minh (IGF), nhưng dùng biến đổi giá trị phức hợp quá mẫu (CMDCT), như trái với cấu hình tiêu chuẩn IGF mà dựa trên phép biểu diễn MDCT được lấy mẫu tới hạn giá trị thực của tín hiệu. CMDCT có thể được coi như tổ hợp của các hệ số MDCT trong phần thực và các hệ số MDST trong phần ảo của mỗi hệ số phổ giá trị phức hợp.

Mặc dù phương pháp mới được mô tả trong nội dung của IGF, việc xử lý của sáng chế có thể được sử dụng trong tổ hợp với phương pháp BWE bất kỳ mà dựa trên phép biểu diễn gián lọc của tín hiệu âm thanh.

Trong phần mô tả mới này, việc dự báo tuyến tính dọc theo hướng tần số không được sử dụng như sự tạo hình nhiều âm theo thời gian, mà là kỹ thuật tạo hình ô theo thời gian (TTS). Việc đổi tên được căn chỉnh bởi thực tế rằng các thành phần tín hiệu được điền đầy ô được tạo hình theo thời gian bởi TTS như trái với việc tạo hình nhiều

âm lượng tử hoá bởi TNS trong các bộ mã hóa-giải mã biến đổi theo cảm quan theo tình trạng kỹ thuật.

Fig.7a thể hiện sơ đồ khối của bộ mã hoá BWE sử dụng IGF và phương pháp TTS mới.

Do đó, giản đồ mã hoá cơ bản hoạt động như sau:

- tính toán CMDCT của tín hiệu miền thời gian  $x(n)$  để thu được tín hiệu miền tần số  $X(k)$
- tính toán bộ lọc TTS giá trị phức hợp
- thu được thông tin phụ đối với BWE và loại bỏ thông tin phổ mà phải được thay thế bằng bộ giải mã
- áp dụng việc lượng tử hoá bằng cách sử dụng mô đun tâm thính học (psycho acoustic module - PAM)
- lưu trữ/truyền dữ liệu, chỉ các hệ số MDCT giá trị thực được truyền

Fig.7b thể hiện bộ giải mã tương ứng. Sơ đồ này chủ yếu đảo ngược các bước được thực hiện trong bộ mã hoá.

Ở đây, giản đồ giải mã cơ bản hoạt động như sau:

- ước lượng các hệ số MDST từ các giá trị MDCT (việc xử lý này bổ sung một độ trễ bộ giải mã khối) và tổ hợp các hệ số MDCT và MDST vào trong các hệ số CMDCT giá trị phức hợp
- thực hiện việc điền đầy ô với việc xử lý sau của công tác này
- áp dụng việc lọc TTS nghịch đảo với các hệ số bộ lọc TTS được truyền
- tính toán CMDCT nghịch đảo

Theo cách khác, lưu ý rằng thứ tự tổng hợp TTS và xử lý sau IGF cũng có thể được đảo ngược trong bộ giải mã nếu việc phân tích TTS và ước lượng tham số IGF được đảo ngược một cách nhất quán trong bộ mã hoá.

Đối với việc mã hóa biến đổi hữu hiệu, ưu tiên cái được gọi là “các khối dài” bằng khoảng 20ms phải được sử dụng để đạt được khuếch đại biến đổi chấp nhận

được. Nếu tín hiệu trong khối dải này chứa các chuyển tiếp, các âm vang trước và sau có thể nghe được xảy ra trong các dải phổ khôi phục do việc điền đầy ô. Fig.7c thể hiện các tác dụng âm vang trước và sau điền hình mà làm suy yếu các chuyển tiếp do IGF. Ở dải bên trái của Fig.7c, ảnh phổ của tín hiệu ban đầu được thể hiện và ở dải bên phải, ảnh phổ của tín hiệu điền đầy ô mà không lọc TTS của sáng chế được thể hiện. Trong ví dụ này, tần số bắt đầu IGF  $f_{IGFstart}$  hoặc  $f_{split}$  giữa dải lõi và dải điền đầy ô được chọn là  $f_s/4$ . Trong dải bên phải của Fig.7c, các âm vang trước và sau khác biệt có thể nhìn thấy bao quanh các chuyển tiếp, đặc biệt nổi bật tại đầu phổ trên của vùng tần số được sao chép.

Nhiệm vụ chính của mô đun TTS là xác nhận các thành phần tín hiệu không mong muốn này trong vùng lân cận kín xung quanh chuyển tiếp và do đó che giấu các phép biến đổi trong vùng theo thời gian được chi phối bởi tác dụng màn chắn âm theo thời gian của nhận thức của con người. Do đó, các hệ số dự báo TTS đòi hỏi được tính toán và được áp dụng bằng cách sử dụng “sự dự báo chuyển tiếp” trong miền CMDCT.

Theo phương án mà tổ hợp TTS và IGF vào trong bộ mã hóa-giải mã, quan trọng để sắp hàng các tham số TTS và các tham số IGF nhất định sao cho ô IGF được lọc toàn bộ bởi một bộ lọc TTS (bộ lọc phẳng hoặc tạo hình) hoặc không. Do đó, tất cả các tần số TTSstart[...] hoặc TTSstop[...] sẽ không được bao gồm trong ô IGF, nhưng sẽ được sắp hàng với các tần số  $f_{IGF...}$  tương ứng. Fig.7d thể hiện ví dụ về các khu vực hoạt động TTS và IGF đối với tập hợp của ba bộ lọc TTS.

Tần số kết thúc TTS được điều chỉnh đến tần số kết thúc của công cụ IGF, mà cao hơn  $f_{IGFstart}$ . Nếu TTS sử dụng nhiều hơn một bộ lọc, phải đảm bảo rằng tần số chéo giữa hai bộ lọc TTS phải phù hợp với tần số chia tách IGF. Mặt khác, một bộ lọc phụ TTS sẽ xung đột với  $f_{IGFstart}$  dẫn đến các giả âm không mong muốn tương tự tạo hình quá.

Theo biến thể của phương án được mô tả trên Fig.7a và Fig.7b, phải chú ý hơn rằng trong bộ giải mã đó các năng lượng IGF được điều chỉnh một cách thích hợp. Đặc biệt là, đây là trường hợp nếu trong quá trình xử lý TTS và IGF, các bộ lọc TTS khác nhau có các khuếch đại dự báo khác nhau được áp dụng cho vùng nguồn (như bộ lọc

phẳng) và vùng phổ đích (như bộ lọc tạo hình mà không phải là bản đối chiếu chính xác của bộ lọc phẳng đã nêu) của một ô IGF. Trong trường hợp này, tỷ lệ khuếch đại dự báo của hai bộ lọc TTS được áp dụng không bằng 1 nữa và do đó việc điều chỉnh năng lượng bởi tỷ lệ này phải được áp dụng.

Theo biến thể phương án thay thế, thứ tự xử lý sau IGF và TTS được đảo ngược. Trong bộ giải mã, việc này có nghĩa rằng sự điều chỉnh năng lượng bằng xử lý sau IGF được tính toán sau khi lọc TTS và do đó là bước xử lý cuối cùng trước khi biến đổi tổng hợp. Do đó, không kể đến các khuếch đại bộ lọc TTS khác nhau được áp dụng cho một ô trong quá trình mã hóa, năng lượng cuối cùng luôn luôn được điều chỉnh một cách thích hợp bằng xử lý IGF.

Ở phía bộ giải mã, các hệ số bộ lọc TTS được áp dụng cho phổ đầy đủ một lần nữa, nghĩa là phổ lỗi được mở rộng bởi phổ được tái tạo. Việc áp dụng TTS là cần thiết để tạo thành đường bao theo thời gian của phổ tái tạo để phù hợp đường bao của tín hiệu ban đầu. Do đó, các âm vang trước được thể hiện được giảm. Ngoài ra, sự áp dụng này cũng tạo hình theo thời gian nhiều âm lượng tử hoá trong tín hiệu thấp hơn  $f_{IGFstart}$  như thông thường với TNS truyền thống.

Trong các bộ mã hóa truyền thống, việc vá phổ trên tín hiệu âm thanh (ví dụ, SBR) làm lỗi sự tương quan phổ ở các đường biên vát và do đó làm suy yếu đường bao theo thời gian của tín hiệu âm thanh do sự đưa đến phép tán sắc. Do đó, một lợi ích khác của thực hiện sự điền đầy ô IGF đối với tín hiệu dư là, sau khi áp dụng bộ lọc tạo hình TTS, các đường biên ô có tương quan liền khối, dẫn đến việc tái tạo theo thời gian trung thực hơn của tín hiệu.

Kết quả của tín hiệu được xử lý tương ứng được thể hiện trên Fig.7e. Khi so sánh, phiên bản không lọc (Fig.7c, dải bên phải) tín hiệu lọc TTS thể hiện sự giảm rõ ràng của các âm vang trước và sau không mong muốn (Fig.7e, dải bên phải).

Hơn nữa, như được mô tả, Fig.7a minh họa bộ mã hoá phù hợp với bộ giải mã của Fig.7b hoặc bộ giải mã của Fig.6a. Về cơ bản, thiết bị để mã hoá tín hiệu âm thanh bao gồm 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 được minh họa trong khối 702, là phép biểu diễn phổ giá trị phức hợp. Hơn nữa,

bộ lọc dự báo như 704 để thực hiện việc dự báo theo tần số được cung cấp để tạo ra các giá trị dư phổ, trong đó bộ lọc dự báo 704 được xác định bởi thông tin bộ lọc dự báo được suy ra từ tín hiệu âm thanh và được chuyển tiếp vào bộ phân kênh bit 710, như được minh họa ở 714 trên Fig.7a. Hơn nữa, bộ mã hóa âm thanh như bộ mã hóa âm thanh theo tâm thính học 704 được đề xuất. Bộ mã hóa âm thanh được tạo cấu hình để mã hóa tập hợp thứ nhất của các phần phổ thứ nhất của các giá trị dư phổ để thu được tập hợp thứ nhất được mã hóa của các giá trị phổ thứ nhất. Ngoài ra, bộ mã hóa theo tham số như bộ mã hóa được minh họa ở 706 trên Fig.7a được đề xuất để mã hóa tập hợp thứ hai của các phần phổ thứ hai. Tốt hơn là, tập hợp thứ nhất của các phần phổ thứ nhất được mã hóa với độ phân giải phổ cao hơn so với tập hợp thứ hai của các phần phổ thứ hai.

Cuối cùng, như được minh họa trên Fig.7a, giao diện đầu ra được cung cấp để xuất ra tín hiệu được mã hóa bao gồm tập hợp thứ hai được mã hóa theo tham số của các phần phổ thứ hai, tập hợp thứ nhất được mã hóa của các phần phổ thứ nhất và thông tin bộ lọc được minh họa như “thông tin phụ TTS” ở 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 để tính toán thông tin bộ lọc. Hơn nữa, bộ lọc dự báo được tạo cấu hình để tính toán các giá trị dư phổ bằng cách sử dụng cùng các giá trị phổ của phép biểu diễn phổ được sử dụng để tính toán thông tin bộ lọc.

Tốt hơn là, bộ lọc TTS 704 được tạo cấu hình theo cách tương tự như đã biết đối với các bộ mã hóa âm thanh của kỹ thuật ưu tiên áp dụng công cụ TNS theo tiêu chuẩn AAC.

Sau đây, phương án khác sử dụng việc giải mã kênh đôi được mô tả trong nội dung của các hình vẽ từ Fig.8a đến Fig.8e. Hơn nữa, việc tham khảo được thực hiện cho sự mô tả của các yếu tố tương ứng trong nội dung của các hình vẽ Fig.2a, Fig.2b (mã hóa kênh ghép 228 và giải mã kênh ghép 204).

Fig.8a minh họa bộ giải mã âm thanh để tạo ra tín hiệu kênh đôi được giải mã. Bộ giải mã âm thanh bao gồm bốn bộ giải mã âm thanh 802 để giải mã tín hiệu kênh đôi được mã hóa để thu được tập hợp thứ nhất của các phần phổ thứ nhất và ngoài ra

bộ giải mã theo tham số 804 để cung cấp dữ liệu tham số đối với tập hợp thứ hai của các phần phổ thứ hai và ngoài ra, việc nhận biết kênh đôi nhận biết phép biểu diễn kênh đôi khác nhau thứ nhất hoặc thứ hai đối với các phần phổ thứ hai. Ngoài ra, bộ tái tạo tần số 806 được cung cấp 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 và dữ liệu tham số đối với phần thứ hai và nhận biết kênh đôi đối với phần thứ hai. Fig.8b minh họa các tổ hợp khác nhau đối với các phép biểu diễn kênh đôi trong khoảng nguồn và khoảng đích. Khoảng nguồn có thể là trong phép biểu diễn kênh đôi thứ nhất và khoảng đích cũng có thể là trong phép biểu diễn kênh đôi thứ nhất. Theo cách khác, khoảng nguồn có thể là trong phép biểu diễn kênh đôi thứ nhất và khoảng đích có thể là trong phép biểu diễn kênh đôi thứ hai. Hơn nữa, khoảng nguồn có thể là trong phép biểu diễn kênh đôi thứ hai và khoảng đích có thể là trong phép biểu diễn kênh đôi thứ nhất như được biểu thị trong cột thứ ba của Fig.8b. Cuối cùng, cả khoảng nguồn và khoảng đích có thể là trong phép biểu diễn kênh đôi thứ hai. Theo phương án, phép biểu diễn kênh đôi thứ nhất là phép biểu diễn kênh đôi riêng biệt trong đó kênh đôi của tín hiệu kênh đôi được biểu diễn theo cách riêng rẽ. Khi đó, phép biểu diễn kênh đôi thứ hai là phép biểu diễn ghép trong đó kênh đôi của phép biểu diễn kênh đôi được biểu diễn ghép, nghĩa là trong đó việc xử lý khác hoặc biến đổi phép biểu diễn được đòi hỏi để tính toán lại phép biểu diễn kênh đôi riêng biệt như được đòi hỏi để xuất ra cho các loa tương ứng.

Theo phương án, phép biểu diễn kênh đôi thứ nhất có thể là phép biểu diễn trái/phải (left/right - L/R) và phép biểu diễn kênh đôi thứ hai là phép biểu diễn âm thanh nổi ghép. Tuy nhiên, các phép biểu diễn kênh đôi khác ngoài sự dự báo bên trái/bên phải hoặc M/S hoặc âm thanh nổi có thể được áp dụng và được sử dụng đối với sáng chế.

Fig.8c minh họa lưu đồ đối với các hoạt động được thực hiện bởi bộ giải mã âm thanh của Fig.8a. Ở bước 812, bộ giải mã âm thanh 802 thực hiện việc giải mã khoảng nguồn. Khoảng nguồn có thể bao gồm, đối với Fig.3a, các dải hệ số tỷ lệ từ SCB1 đến SCB3. Hơn nữa, có thể có việc nhận biết kênh đôi đối với mỗi dải hệ số tỷ lệ và dải hệ số tỷ lệ 1 có thể, ví dụ, trong phép biểu diễn thứ nhất (như L/R) và dải hệ số tỷ lệ thứ ba có thể trong phép biểu diễn kênh đôi thứ hai như M/S hoặc trộn giảm/dự báo. Do đó, bước 812 có thể dẫn đến các phép biểu diễn khác nhau đối với các dải khác

nhau. Sau đó, ở bước 814, bộ tái tạo tần số 806 được tạo cấu hình để chọn lọc khoảng nguồn đối với phép biểu diễn tần số. Ở 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 đích. Nếu cả hai phép biểu diễn giống nhau, bộ tái tạo tần số 806 cung cấp phép biểu diễn tần số riêng biệt đối với mỗi kênh của tín hiệu kênh đôi. Tuy nhiên, khi cả hai phép biểu diễn như được phát hiện trong khối 818 không giống nhau, khi đó dòng tín hiệu 824 được chọn 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 để tái tạo khoảng đích. Do đó, bộ giải mã của Fig.8a giúp tái tạo khoảng đích được biểu thị như có sự nhận biết kênh đôi thứ hai bằng cách sử dụng khoảng nguồn trong phép biểu diễn kênh đôi thứ nhất. Về bản chất, sáng chế còn cho phép tái tạo khoảng đích bằng cách sử dụng khoảng nguồn có cùng nhận biết kênh đôi. Và ngoài ra, sáng chế cho phép tái tạo khoảng đích có nhận biết kênh đôi biểu thị phép biểu diễn kênh đôi ghép và sau đó biến đổi phép biểu diễn này thành phép biến đổi kênh riêng biệt được yêu cầu để lưu trữ hoặc truyền tới các loa tương ứng đối với tín hiệu kênh đôi.

Nhấn mạnh rằng kênh đôi của phép biểu diễn kênh đôi có thể là kênh đôi âm thanh nổi như kênh bên trái và kênh bên phải. Tuy nhiên, tín hiệu cũng có thể là tín hiệu đa kênh có, ví dụ, năm kênh và kênh loa trầm phụ hoặc thậm chí có nhiều kênh hơn. Khi đó, việc xử lý kênh đôi theo cặp như được mô tả trong nội dung của các hình vẽ từ Fig.8a đến Fig.8e có thể được thực hiện trong đó các cặp có thể, ví dụ, là kênh bên trái và kênh bên phải, kênh bao quanh bên trái và kênh bao quanh bên phải và kênh trung tâm và kênh LFE (loa trầm phụ). Các cặp khác bất kỳ có thể được sử dụng để biểu diễn, ví dụ sáu kênh đầu vào bởi ba quy trình xử lý kênh đôi.

Fig.8d minh họa biểu đồ khối của bộ giải mã của sáng chế tương ứng với Fig.8a. Khoảng nguồn hoặc bộ giải mã lõi 830 có thể tương ứng với bộ giải mã âm thanh 802. Các khối khác 832, 834, 836, 838, 840, 842 và 846 có thể là các phần của bộ tái tạo tần số 806 của Fig.8a. Cụ thể là, khối 832 là bộ biến đổi phép biểu diễn để biến đổi phép biểu diễn khoảng nguồn trong các dải riêng rẽ sao cho, tại đầu ra của khối 832, một mặt, tập hợp hoàn thiện của khoảng nguồn trong phép biểu diễn thứ nhất và mặt khác, trong phép biểu diễn kênh đôi thứ hai có mặt. Hai phép biểu diễn



khoảng nguồn hoàn thiện này có thể được lưu trữ trong bộ nhớ 834 đối với hai phép biểu diễn của khoảng nguồn.

Khi đó, khối 836 áp dụng việc tạo ra lát tần số bằng cách sử dụng, như trong đầu vào, khoảng nguồn ID và ngoài ra sử dụng như là đầu vào ID kênh đôi cho khoảng đích. Dựa trên ID kênh đôi đối với khoảng đích, bộ tạo lát tần số truy cập bộ nhớ 834 và nhận phép biểu diễn kênh đôi của khoảng nguồn phù hợp với ID kênh đôi đối với khoảng đích đưa vào trong bộ tạo lát tần số ở 835. Do đó, khi ID kênh đôi đối với khoảng đích biểu thị việc xử lý âm thanh nối ghép, thì bộ tạo lát tần số 836 truy cập bộ nhớ 834 để thu được phép biểu diễn âm thanh nối ghép của khoảng nguồn được biểu thị bởi khoảng nguồn ID 833.

Bộ tạo lát tần số 836 thực hiện hoạt động này đối với mỗi khoảng đích và đầu ra của bộ tạo lát tần số sao cho mỗi kênh của phép biểu diễn kênh được nhận biết bởi việc nhận biết kênh đôi có mặt. Khi 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 biết bởi việc nhận biết kênh đôi. Với mục tiêu này, các tham số điều chỉnh đường bao được đòi hỏi và các tham số này được truyền từ bộ mã hoá đến bộ giải mã trong cùng phép biểu diễn kênh đôi như được mô tả. Khi việc nhận biết kênh đôi trong khoảng đích sẽ được xử lý bởi bộ điều chỉnh đường bao có nhận biết kênh đôi biểu thị phép biểu diễn kênh đôi ngoài dữ liệu đường bao đối với khoảng đích này, khi đó cần có bộ biến đổi tham số 840 biến đổi các tham số đường bao thành phép biểu diễn kênh đôi. Ví dụ, khi việc nhận biết kênh đôi đối với một dải biểu thị việc mã hóa âm thanh nối ghép và khi các tham số đối với khoảng đích này được truyền dưới dạng các tham số đường bao L/R, khi đó bộ biến đổi tham số tính toán các tham số đường bao âm thanh nối ghép từ các tham số đường bao L/R như được mô tả sao cho phép biểu diễn tham số thích hợp được sử dụng để điều chỉnh đường bao phổ của khoảng đích.

Theo phương án được ưu tiên khác, các tham số đường bao được truyền như các tham số âm thanh nối ghép khi âm thanh nối ghép được sử dụng trong lát đích.

Khi các tác giả sáng chế giả định rằng việc đưa vào trong bộ điều chỉnh đường bao 838 là tập hợp chứa các khoảng đích có các phép biểu diễn kênh đôi khác nhau,

khi đó đầu ra của bộ điều chỉnh đường bao 838 là tập hợp của các khoảng đích trong các phép biểu diễn kênh đôi khác nhau. Khi khoảng đích có phép biểu diễn ghép như M/S, khi đó khoảng đích này được xử lý bởi bộ biến đổi phép biểu diễn 842 để tính toán phép biểu diễn tách biệt được yêu cầu để lưu trữ hoặc truyền đến các loa. Tuy nhiên, khi khoảng đích đã có phép biểu diễn riêng biệt, dòng tín hiệu 844 được chọn và bộ biến đổi phép biểu diễn 842 là được bỏ qua. Ở đầu ra của khối 842, phép biểu diễn phổ kênh đôi là phép biểu diễn kênh đôi riêng biệt thu được mà sau đó còn có thể được xử lý như được biểu thị bởi khối 846, trong đó việc xử lý tiếp tục này có thể, ví dụ, là sự chuyển đổi tần số/thời gian hoặc xử lý theo yêu cầu khác bất kỳ.

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

Theo phương án, chỉ khoảng khôi phục bắt đầu với tần số bắt đầu IGF 309 của Fig.3a có nhận biết hai kênh đối với các dải khôi phục khác nhau. Theo phương án khác, việc này cũng được áp dụng đối với khoảng tần số thấp hơn tần số bắt đầu IGF 309.

Theo phương án khác, nhận biết dải nguồn và nhận biết dải đích có thể được xác định theo cách thích ứng bằng phép phân tích tương tự. Tuy nhiên, việc xử lý kênh đôi của sáng chế cũng có thể được áp dụng khi có sự ghép cố định của khoảng nguồn vào khoảng đích. Khoảng nguồn có thể được sử dụng để tái tạo, đối với tần số, khoảng đích rộng hơn bằng hoạt động điền đầy lát tần số sóng hài hoặc hoạt động điền đầy lát tần số sao chép bằng cách sử dụng hai hoặc nhiều hoạt động điền đầy lát tần số tương tự với xử lý đối với nhiều bản vá đã biết từ việc xử lý AAC hữu hiệu ở mức cao.

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

tích phổ 866 được cung cấp để thực hiện phép phân tích để xác định, mà các phần phổ sẽ được mã hoá với độ phân giải cao, nghĩa là để phát hiện tập hợp thứ nhất của các phần phổ thứ nhất và còn để phát hiện tập hợp thứ hai của các phần phổ thứ hai.

Hơn 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 sự nhận biết kênh đôi nhận biết phép biểu diễn kênh đôi thứ nhất hoặc phép biểu diễn kênh đôi thứ hai.

Tùy thuộc vào kết quả của bộ phân tích kênh đôi, dải trong phép biểu diễn phổ thứ hai được tham số hoá bằng cách sử dụng phép biểu diễn kênh đôi thứ nhất hoặc phép biểu diễn kênh đôi thứ hai và việc này được thực hiện bởi bộ mã hoá tham số 868. Khoảng tần số lỗi, nghĩa là dải tần số thấp hơn tần số bắt đầu IGF 309 của Fig.3a được mã hoá bởi bộ mã hoá lỗi 870. Kết quả của các khối 868 và 870 được đưa vào giao diện đầu ra 872. Như được biểu thị, bộ phân tích kênh đôi cung cấp sự nhận biết kênh đôi đối với mỗi dải cao hơn tần số bắt đầu IGF hoặc đối với toàn bộ khoảng tần số và việc nhận biết kênh đôi này cũng được chuyển tiếp vào giao diện đầu ra 872 sao cho dữ liệu này cũng được bao gồm trong tín hiệu được mã hoá 873 được xuất ra bởi giao diện đầu ra 872.

Hơn nữa, tốt hơn là bộ mã hoá âm thanh bao gồm bộ biến đổi theo dải 862. Dựa vào sự quyết định của bộ phân tích kênh đôi 862, tín hiệu đầu ra của bộ chuyển đổi phổ thời gian 862 được biến đổi thành phép biểu diễn được biểu thị bởi bộ phân tích kênh đôi và, cụ thể là, bởi ID kênh đôi 835. Do đó, đầu ra của bộ biến đổi theo dải 862 là tập hợp của các dải tần số trong đó mỗi dải tần số có thể là trong phép biểu diễn kênh đôi thứ nhất hoặc phép biểu diễn kênh đôi khác nhau thứ hai. Khi sáng chế được áp dụng trong dải đầy đủ, nghĩa là khi khoảng nguồn và khoảng khôi phục đều được xử lý bởi bộ biến đổi theo dải, bộ phân tích phổ 860 có thể phân tích phép biểu diễn này. Tuy nhiên, theo cách khác, bộ phân tích phổ 860 cũng có thể phân tích đầu ra tín hiệu bởi bộ chuyển đổi phổ thời gian như được biểu thị bởi đường dẫn điều khiển 861. Do đó, bộ phân tích phổ 860 có thể áp dụng phép phân tích âm được ưu tiên ở đầu ra của bộ biến đổi băng thông 862 hoặc đầu ra của bộ chuyển đổi phổ thời gian 860 trước khi được xử lý bởi bộ biến đổi theo dải 862. Hơn nữa, bộ phân tích phổ có thể áp dụng việc nhận biết khoảng nguồn phù hợp tốt nhất đối với khoảng đích nhất định dựa trên

kết quả của bộ biến đổi theo dải 862 hoặc dựa trên kết quả của bộ chuyển đổi thời gian-phổ 860.

Tiếp theo, việc tham khảo được thực hiện đến 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 đã được mô tả trong nội dung của Fig.3a và Fig.3b.

Trạng thái hiện đại của bộ mã hóa âm thanh của kỹ thuật hiện nay áp dụng các kỹ thuật khác nhau để giảm thiểu lượng dữ liệu biểu diễn tín hiệu âm thanh đã cho. Bộ mã hóa âm thanh tương tự USAC [1] áp dụng phép biến đổi thời gian thành tần số tương tự MDCT để có phép biểu diễn phổ của tín hiệu âm thanh đã cho. Các hệ số MDCT này được lượng tử hoá bằng cách khai thác các khía cạnh tâm thính học của hệ thống nghe của con người. Nếu tốc độ bit sẵn có bị giảm, sự lượng tử hoá thô hơn đến số lượng lớn các giá trị phổ bằng 0 mà dẫn đến các giả âm nghe được ở phía bộ giải mã. Để cải thiện chất lượng theo cảm quan, trạng thái của bộ giải mã của kỹ thuật hiện nay điền đầy các phần phổ bằng 0 này với nhiễu âm ngẫu nhiên. Phương pháp IGF thu hoạch các ô từ tín hiệu không bằng 0 còn lại để điền đầy các khoảng trống này trong phổ. Đây là điều cốt yếu đối với chất lượng theo cảm quan của tín hiệu âm thanh được giải mã mà đường bao phổ và sự phân phối năng lượng của các hệ số phổ được bảo toàn. Phương pháp điều chỉnh năng lượng được thể hiện ở đây sử dụng thông tin phụ được truyền để khôi phục đường bao MDCT phổ của tín hiệu âm thanh.

Trong eSBR [15], tín hiệu âm thanh được lấy mẫu giảm ít nhất bởi hệ số hai và phần tần số cao của phổ được đặt bằng 0 đầy đủ ngoài [1, 17]. Phần bị xóa này được thay thế bằng các kỹ thuật tham số, eSBR, ở phía bộ giải mã. eSBR ngụ ý việc sử dụng của biến đổi bổ sung, phép biến đổi QMF mà được sử dụng để thay thế phần tần số cao trống rỗng và để lấy mẫu lại tín hiệu âm thanh [17]. Việc này bổ sung độ phức tạp tính toán và việc tiêu thụ bộ nhớ cho bộ mã hóa âm thanh.

Bộ mã hóa USAC [15] đem lại khả năng điền đầy các lỗ phổ (đường dẫn phổ bằng 0) với nhiễu âm ngẫu nhiên mà có các phía giảm sau đây: nhiễu âm ngẫu nhiên không thể bảo toàn cấu trúc mịn theo thời gian của tín hiệu chuyển tiếp và nhiễu âm ngẫu nhiên không thể bảo toàn cấu trúc sóng hài của tín hiệu âm.

Khu vực trong đó eSBR hoạt động ở phía bộ giải mã bị xoá bỏ một cách hoàn toàn bởi bộ mã hoá [1]. Do đó, eSBR có xu hướng xoá bỏ các đường dẫn tiếng nói trong vùng tần số cao hoặc các cấu trúc sóng hài méo của tín hiệu ban đầu. Do độ phân giải tần số QMF của eSBR rất thấp và chỉ có thể chèn lại các thành phần hình sin trong độ phân giải thô của giàn lọc cơ bản, việc tái tạo các thành phần âm trong eSBR trong khoảng tần số sao chép có độ chính xác rất thấp.

eSBR sử dụng các kỹ thuật để điều chỉnh các năng lượng của các khu vực được vá, sự điều chỉnh đường bao phổ [1]. Kỹ thuật này sử dụng các giá trị năng lượng được truyền trên lưới thời gian tần số QMF để tạo hình lại đường bao phổ. Trạng thái này của kỹ thuật trong kỹ thuật hiện nay không xử lý các phổ bị xoá bỏ một phần và do độ phân giải thời gian cao nó có xu hướng cân lượng tương đối lớn các bit để truyền các giá trị năng lượng thích hợp hoặc để áp dụng việc lượng tử hoá thô cho các giá trị năng lượng.

Phương pháp IGF không cần sự biến đổi bổ sung do phương pháp này sử dụng phép biến đổi MDCT truyền thống mà được tính toán như được mô tả trong [15].

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

- a) áp dụng biến đổi MDCT cửa sổ cho tín hiệu âm thanh đầu vào [16, phần 4.6], tính toán MDST cửa sổ một cách tùy ý, hoặc ước lượng MDST cửa sổ từ MDCT tính toán
- b) áp dụng TNS/TTS đối với các hệ số MDCT [15, phần 7.8]
- c) tính toán năng lượng trung bình đối với mỗi dải hệ số tỷ lệ MDCT cao hơn tần số bắt đầu IGF ( $f_{IGFstart}$ ) đến tần số kết thúc IGF ( $f_{IGFstop}$ )
- d) lượng tử hoá các giá trị năng lượng trung bình

$f_{IGFstart}$  và  $f_{IGFstop}$  là các tham số đã cho của người dùng.

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

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

- a) Khử lượng tử hoá các giá trị MDCT được truyền
- b) Áp dụng việc điền đầy nhiễu âm USAC truyền thông nếu được tín hiệu hóa
- c) Áp dụng điền đầy ô IGF
- d) Khử lượng tử hoá các giá trị năng lượng được truyền
- e) Điều chỉnh theo dải hệ số tỷ lệ đường bao phổ
- f) Áp dụng TNS/TTS nếu được tín hiệu hóa

Cho  $\hat{x} \in \mathbb{R}^N$  là phép biểu diễn phổ giá trị thực được biến đổi MDCT của tín hiệu âm thanh của sở có chiều dài của sở  $2N$ . Việc biến đổi này được mô tả trong [16]. Bộ mã hoá tùy ý áp dụng TNS đối với  $\hat{x}$ .

Trong [16, 4.6.2], việc phân chia  $\hat{x}$  trong các dải hệ số tỷ lệ được mô tả. Các dải hệ số tỷ lệ là tập hợp gồm tập hợp các chỉ số và được xác định ở đây với  $scb$ .

Các giới hạn của mỗi  $scb_k$  với  $k = 0, 1, 2, \dots, max\_sfb$  được xác định bởi dãy  $swb\_offset$  (16, 4.6.2), trong đó  $swb\_offset[k]$  và  $swb\_offset[k+1]-1$  xác định chỉ số thứ nhất và sau cùng đối với đường dẫn hệ số phổ thấp nhất và cao nhất được chứa trong  $scb_k$ . Các tác giả sáng chế biểu thị dải hệ số tỷ lệ:

$$scb_k := \{swb\_offset[k], 1+swb\_offset[k], 2+swb\_offset[k], \dots, swb\_offset[k+1]-1\}$$

Nếu công cụ IGF được sử dụng bởi bộ mã hoá, người dùng xác định tần số bắt đầu IGF và tần số kết thúc IGF. Hai giá trị này được lập bản đồ đến chỉ số dải hệ số tỷ lệ phù hợp tốt nhất  $igfStartSfb$  và  $igfStopSfb$ . Cả hai giá trị được tín hiệu hóa trong dòng bit cho bộ giải mã.

[16] mô tả phép biến đổi khối dài và khối ngắn. Đối với các khối dài, chỉ có một tập hợp các hệ số phổ cùng với một tập hợp các hệ số tỷ lệ được truyền đến bộ giải mã. Đối với các khối ngắn, tám cửa sổ ngắn với tám tập hợp khác nhau của các hệ số phổ được tính toán. Để lưu tốc độ bit, các hệ số tỷ lệ của tám cửa sổ khối ngắn này được phân nhóm bởi bộ mã hoá.

Trong trường hợp IGF, phương pháp được thể hiện ở đây sử dụng các dải hệ số tỷ lệ truyền thống để nhóm các giá trị phổ mà được truyền đến bộ giải mã:

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

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

Để lượng tử hoá

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

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

Giả sử rằng bộ mã hoá quyết định để nhóm các tập hợp hệ số tỷ lệ  $num\_window\_group$ .

Các tác giả sáng chế biểu thị với  $w$ , việc phân chia nhóm này của tập hợp  $\{0, 1, 2, \dots, 7\}$  mà là các chỉ số của tám cửa sổ ngắn.  $w_l$  biểu thị tập hợp phụ thứ  $l$  của  $w$ , trong đó  $l$  có nghĩa là chỉ số của nhóm cửa sổ,  $0 \leq l < num\_window\_group$ .

Để tính toán khối ngắn, người dùng xác định tần số bắt đầu/kết thúc IGF được lập bản đồ đối với các dải hệ số tỷ lệ thích hợp. Tuy nhiên, để đơn giản khi biểu thị đối với các khối ngắn  $k = igfStartSfb, 1 + igfStartSfb, 2 + igfStartSfb, \dots, igfEndSfb$ .

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

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

Để lượng tử hoá

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

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

Các công thức mã hóa được kể đến trên đây hoạt động bằng cách sử dụng chỉ các hệ số MDCT giá trị thực  $\hat{x}$ . Để thu được sự phân bố năng lượng ổn định hơn trong

khoảng IGF, tức là, để giảm các dao động biên độ theo thời gian, phương pháp thay thế có thể được sử dụng để tính toán các giá trị  $\hat{E}_k$ :

Cho  $\hat{x}_r \in \mathbb{R}^N$  là phép biểu diễn phổ giá trị thực được biến đổi MDCT của tín hiệu âm thanh của sổ có chiều dài của sổ  $2N$ , và  $\hat{x}_i \in \mathbb{R}^N$  là phép biểu diễn phổ được biến đổi MDST giá trị thực của cùng phần tín hiệu âm thanh. Phép biểu diễn phổ MDST  $\hat{x}_i$  có thể được tính toán một cách chính xác hoặc được ước lượng từ  $\hat{x}_r$ .  $\hat{c} = (\hat{x}_r, \hat{x}_i) \in \mathbb{C}^N$  biểu thị phép biểu diễn phổ phức hợp của tín hiệu âm thanh của sổ, có  $\hat{x}_r$  là phần thực của nó và  $\hat{x}_i$  là phần ảo của nó. Bộ mã hoá tuỳ ý áp dụng TNS đối với  $\hat{x}_r$  và  $\hat{x}_i$ .

Giờ đây, năng lượng của tín hiệu trong khoảng IGF có thể được đo với

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

Các năng lượng giá trị thực và phức hợp của dải khôi phục, tức là, ô mà sẽ được sử dụng ở phía bộ giải mã trong sự khôi phục khoảng IGF  $\text{scb}_k$ , được tính toán với:

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

trong đó  $\text{tr}_k$  là tập hợp các chỉ số – khoảng lát nguồn ghép, theo sự phụ thuộc của  $\text{scb}_k$ . Trong hai công thức trên đây, thay cho tập hợp chỉ số  $\text{scb}_k$ , tập hợp  $\overline{\text{scb}_k}$  (được xác định sau trong phần mô tả này) có thể được sử dụng để tạo ra  $\text{tr}_k$  để tạo ra các giá trị chính xác hơn  $E_r$  và  $E_r$ .

Tính toán

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

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

với

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



giờ đây phiên bản ổn định hơn của  $E_k$  được tính toán, do việc tính toán của  $E_k$  với các giá trị MDCT chỉ bị suy yếu bởi thực tế rằng các giá trị MDCT không tuân theo định lý của Parseval và do đó chúng không phản chiếu thông tin năng lượng hoàn toàn của các giá trị phổ.  $\tilde{E}_k$  được tính toán như trên đây.

Như được lưu ý trên đây, đối với các khối ngắn, giả sử rằng bộ mã hoá quyết định để nhóm các tập hợp hệ số tỷ lệ  $num\_window\_group$ . Như trên đây,  $w_l$  có nghĩa là tập hợp phụ thứ  $l$  của  $w$ , trong đó  $l$  có nghĩa là chỉ số của nhóm cửa sổ,  $0 \leq l < num\_window\_group$ .

Một lần nữa, phiên bản thay thế được nêu trên đây để tính toán phiên bản ổn định hơn của  $E_{kl}$  có thể được tính toán. Với việc xác định  $\hat{c} := (\hat{c}_r, \hat{c}_i) \in \mathbb{C}^N$ ,  $\hat{c}_r \in \mathbb{R}^N$  là MDCT được biến đổi và  $\hat{c}_i \in \mathbb{R}^N$  là tín hiệu âm thanh cửa sổ được biến đổi MDST có chiều dài  $2N$ , tính toán:

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

Theo cách tương tự tính toán:

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

và tiếp tục với hệ số  $f_{kl}$

$$f_{kl} = \frac{E_{okl}}{E_{tkl}}$$

mà được sử dụng để điều chỉnh  $E_{rkl}$  được tính toán trước:

$$E_{kl} = \sqrt{f_{kl} E_{rkl}}$$

$\tilde{E}_{kl}$  được tính toán như trên đây.

Quy trình không chỉ sử dụng năng lượng của dải khôi phục được suy ra từ dải khôi phục phức hợp hoặc từ các giá trị MDCT, nhưng cũng sử dụng thông tin năng lượng từ khoảng nguồn cung cấp sự khôi phục năng lượng cải thiện hơn.

Cụ thể là, bộ tính toán tham số 1006 được tạo cấu hình để tính toán thông tin năng lượng đối với dải khôi phục bằng cách sử dụng thông tin về năng lượng của dải khôi phục và ngoài ra bằng cách sử dụng thông tin về năng lượng của khoảng nguồn sẽ được sử dụng để khôi phục dải khôi phục.

Hơn nữa, bộ tính toán tham số 1006 được tạo cấu hình để tính toán thông tin năng lượng ( $E_{ok}$ ) đối với dải khôi phục của phổ phức hợp của tín hiệu ban đầu, để tính toán thông tin năng lượng khác ( $E_{rk}$ ) đối với khoảng nguồn của phần giá trị thực của phổ phức hợp của tín hiệu ban đầu sẽ được sử dụng để khôi phục dải khôi phục, và trong đó bộ tính toán tham số được tạo cấu hình để tính toán thông tin năng lượng đối với dải khôi phục bằng cách sử dụng thông tin năng lượng ( $E_{ok}$ ) và thông tin năng lượng khác ( $E_{rk}$ ).

Hơn nữa, bộ tính toán tham số 1006 được tạo cấu hình để xác định thông tin năng lượng thứ nhất ( $E_{ok}$ ) của dải hệ số tỷ lệ sẽ được khôi phục của phổ phức hợp của tín hiệu ban đầu, để xác định thông tin năng lượng thứ hai ( $E_{tk}$ ) trên khoảng nguồn của phổ phức hợp của tín hiệu ban đầu sẽ được sử dụng để khôi phục dải hệ số tỷ lệ sẽ 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 ban đầu sẽ được sử dụng để khôi phục dải hệ số tỷ lệ sẽ được khôi phục, để xác định thông tin trọng số dựa trên mối quan hệ giữa ít nhất hai thông tin năng lượng thứ nhất, thông tin năng lượng thứ hai và thông tin năng lượng thứ ba và để lấy trọng số một trong thông tin năng lượng thứ nhất và thông tin năng lượng thứ ba bằng cách sử dụng thông tin trọng số để thu được thông tin năng lượng trọng số và để sử dụng thông tin năng lượng trọng số như thông tin năng lượng đối với dải khôi phục.

Các ví dụ cho các tính toán là như sau, nhưng nhiều ví dụ khác có thể xuất hiện đối với chuyên gia trong lĩnh vực này khi xét đến nguyên lý chung trên đây:

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 nhận biết thực tế rằng mặc dù chỉ các giá trị MDCT thực được xử lý ở phía bộ giải mã, việc tính toán thực là – do sự chồng lấp và bổ sung – miền thời gian tạo bí danh quy trình huỷ bỏ được tạo ra hoàn toàn bằng cách sử dụng các số phức hợp. Tuy nhiên, cụ thể là, việc xác định 918 của thông tin năng lượng ô của các phần phổ khác 922, 923 của dải khôi phục 920 đối với các giá trị tần số khác với phần phổ thứ nhất 921 có các tần số trong dải khôi phục 920 dựa trên các giá trị MDCT thực. Do đó, thông tin năng lượng được truyền đến bộ giải mã sẽ thường nhỏ hơn so với thông tin năng lượng  $E_{ok}$  trên dải khôi phục của phổ phức hợp của tín hiệu ban đầu. Ví dụ đối với trường hợp C trên đây, điều này có nghĩa rằng hệ số  $f_k$  (thông tin trọng số) sẽ nhỏ hơn 1.

Ở phía bộ giải mã, nếu công cụ IGF được tín hiệu hóa là ON, các giá trị được truyền  $\hat{E}_k$  thu được từ dòng bit và sẽ được lượng tử hoá với

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

với  $k = \text{igfStartSfb}, 1 + \text{igfStartSfb}, 2 + \text{igfStartSfb}, \dots, \text{igfEndSfb}$ .

Bộ giải mã lượng tử hoá các giá trị MDCT được truyền thành  $x \in \mathbb{R}^N$  và tính toán năng lượng còn lại:

$$sE_k := \sum_{i \in \text{scb}_k} x_i^2$$

trong đó,  $k$  là nằm trong khoảng như được xác định trên đây.

Các tác giả sáng chế đã biểu thị  $\overline{scb_k} = \{i | i \in scb_k \wedge x_i = 0\}$ . Tập hợp này chứa tất cả các chỉ số của dải hệ số tỷ lệ và  $scb_k$  mà đã được lượng tử hoá đến 0 bởi bộ mã hoá.

IGF có phương pháp băng con (không được mô tả ở đây) được sử dụng để điền đầy các khoảng trống phổ thu được từ việc lượng tử hoá thô của các giá trị phổ MDCT ở phía bộ mã hoá bằng cách sử dụng các giá trị khác 0 của MDCT được truyền.  $x$  ngoài ra sẽ chứa các giá trị mà thay thế tất cả các giá trị bằng 0 trước. Năng lượng ô được tính toán theo:

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

trong đó  $k$  nằm trong khoảng như được xác định trên đây.

Năng lượng bị mất trong dải khôi phục được tính toán theo:

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

Và hệ số khuếch đại để điều chỉnh thu được theo:

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

Với

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

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

$$x_i := g' x_i$$

với  $i \in \overline{scb_k}$  và  $k$  nằm trong khoảng như được xác định trên đây.

Việc này tạo hình lại đường bao phổ  $x$  để tạo hình đường bao phổ ban đầu  $\hat{x}$ .

Với chuỗi cửa sổ ngắn, tất cả các phép tính toán như được nêu trên đây tương tự trong nguyên lý, nhưng việc tạo nhóm các dải hệ số tỷ lệ được xét đến. Các tác giả sáng chế đã biểu thị dưới dạng  $E_{k,i}$  các giá trị năng lượng tạo nhóm lượng tử hoá thu được từ dòng bit. Tính toán:

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

và

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

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

Tính toán:

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

Và

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

Với

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

Áp dụng

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

với  $i \in \overline{scb_{k,l}}$ .

Đối với các ứng dụng tốc độ bit thấp, có thể tạo nhóm từng cặp các giá trị  $E_k$  là có thể thực hiện được mà không hao tổn quá nhiều độ chính xác. Phương pháp này chỉ được áp dụng với các khối dài:

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

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

Một lần nữa, sau khi lượng tử hoá tất cả các trị số  $E_{k \gg 1}$  được truyền đến bộ giải mã.

Fig.9a minh họa thiết bị để giải mã tín hiệu âm thanh được mã hoá bao gồm phép biểu diễn được mã hoá của tập hợp thứ nhất của các phần phổ thứ nhất và phép biểu diễn được mã hoá của dữ liệu tham số biểu thị các năng lượng phổ cho tập hợp thứ hai của các phần phổ thứ hai. Tập hợp thứ nhất của các phần phổ thứ nhất được biểu thị ở 901a trên Fig.9a, và phép biểu diễn được mã hoá của dữ liệu tham số được biểu thị ở 901b trên Fig.9a. Bộ giải mã âm thanh 900 được cung cấp để giải mã phép biểu diễn được mã hoá 901a của tập hợp thứ nhất của các phần phổ thứ nhất để thu được tập hợp thứ nhất được giải mã của các phần phổ thứ nhất 904 và để giải mã phép biểu diễn được mã hoá của dữ liệu tham số để thu được dữ liệu tham số được giải mã 902 đối với tập hợp thứ hai của các phần phổ thứ hai biểu thị các năng lượng riêng rẽ đối với các dải khôi phục riêng rẽ, trong đó các phần phổ thứ hai được định vị trong các dải khôi phục. Hơn nữa, bộ tái tạo tần số 906 được cung cấp để khôi phục các giá trị phổ của dải khôi phục bao gồm phần phổ thứ hai. Bộ tái tạo tần số 906 sử dụng phần phổ thứ nhất của tập hợp thứ nhất của các phần phổ thứ nhất và thông tin năng lượng riêng rẽ đối với dải khôi phục, trong đó dải khôi phục bao gồm phần phổ thứ nhất và phần phổ thứ hai. Bộ tái tạo tần số 906 bao gồm bộ tính toán 912 để xác định thông tin năng lượng còn lại bao gồm năng lượng tích tụ của phần phổ thứ nhất có các tần số trong dải khôi phục. Hơn 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 khác của dải khôi phục và đối với các giá trị tần số sẽ khác với phần phổ thứ nhất, trong đó các giá trị tần số này có các tần số trong dải khôi phục, trong đó các phần phổ khác sẽ được tạo ra bởi việc tạo ra tần số bằng cách sử dụng phần phổ thứ nhất khác với phần phổ thứ nhất trong dải khôi phục.

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

Việc tham khảo được thực hiện đến Fig.9c minh họa dải khôi phục 920 nhất định. Dải khôi phục bao gồm phần phổ thứ nhất trong dải khôi phục như phần phổ thứ

nhất 306 trên Fig.3a được minh họa theo sơ đồ ở 921. Hơn nữa, phần còn lại của các giá trị phổ trong dải khôi phục 920 sẽ được tạo ra bằng cách sử dụng vùng nguồn, ví dụ, từ dải hệ số tỷ lệ 1, 2, 3 thấp hơn tần số bắt đầu điền đầy khoảng trống thông minh 309 của Fig.3a. Bộ tái tạo tần số 906 được tạo cấu hình để tạo ra các giá trị phổ thô đối với các phần phổ thứ hai 922 và 923. Sau đó, hệ 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 dải tần số 922, 923 để thu được các phần phổ thứ hai được khôi phục và điều chỉnh trong dải khôi phục 920 mà giờ đây có cùng độ phân giải phổ, nghĩa là cùng khoảng cách đường dẫn như phần phổ thứ nhất 921. Quan trọng là cần hiểu rằng phần phổ thứ nhất trong dải khôi phục được minh họa ở 921 trên Fig.9c được giải mã bởi bộ giải mã âm thanh 900 và không bị tác động bởi khối thực hiện điều chỉnh đường bao 916 của Fig.9b. Thay vào đó, phần phổ thứ nhất trong dải khôi phục được biểu thị ở 921 ở phía bên trái được giữ nguyên, do phần phổ thứ nhất này được xuất ra bởi băng thông đầy đủ hoặc bộ giải mã âm thanh tốc độ đầy đủ 900 qua đường dẫn 904.

Sau đây, ví dụ cụ thể với các số thực được mô tả. Năng lượng còn lại được tính toán bởi khối 912 là, ví dụ, năm đơn vị năng lượng và năng lượng này là năng lượng của bốn đường dẫn phổ được biểu thị ví dụ trong phần phổ thứ nhất 921.

Hơn nữa, giá trị năng lượng E3 đối với dải khôi phục tương ứng với dải hệ số tỷ lệ 6 của Fig.3b hoặc Fig.3a bằng 10 đơn vị. Quan trọng là, giá trị năng lượng không chỉ bao gồm năng lượng của các phần phổ 922, 923, mà còn bao gồm năng lượng đầy đủ của dải khôi phục 920 như được tính toán ở phía bộ mã hoá, nghĩa là trước khi thực hiện việc phân tích phổ bằng cách sử dụng, ví dụ, màn chắn âm. Do đó, mười đơn vị năng lượng bao phủ các phần phổ thứ nhất và thứ hai trong dải khôi phục. Khi đó, giả sử rằng năng lượng của dữ liệu khoảng nguồn đối với các khối 922, 923 hoặc đối với dữ liệu khoảng đích thô đối với khối 922, 923 bằng tám đơn vị năng lượng. Do đó, năng lượng mất của năm đơn vị được tính toán.

Dựa trên năng lượng mất được chia bởi năng lượng ô tEk, hệ số khuếch đại 0,79 được tính toán. Khi đó, các đường dẫn phổ thô đối với các phần phổ thứ hai 922, 923 được nhân với hệ số khuếch đại được tính toán. Do đó, chỉ các giá trị phổ đối với các phần phổ thứ hai 922, 923 được điều chỉnh và các đường dẫn phổ đối với phần phổ thứ

nhất 921 không bị tác động bởi việc điều chỉnh đường bao này. Tiếp theo để nhân các giá trị phổ thô với các phần phổ thứ hai 922, 923, dải khôi phục hoàn toàn đã được tính toán gồm có các phần phổ thứ nhất trong dải khôi phục, và gồm có các đường dẫn phổ trong các phần phổ thứ hai 922, 923 trong dải khôi phục 920.

Tốt hơn là, khoảng nguồn để tạo ra dữ liệu phổ thô trong các dải 922, 923, đối với tần số, thấp hơn tần số bắt đầu IGF 309 và dải khôi phục 920 cao hơn tần số bắt đầu IGF 309.

Hơn nữa, tốt hơn là các đường biên dải khôi phục trùng khớp với các đường biên dải hệ số tỷ lệ. Do đó, dải khôi phục có, theo một phương án, kích cỡ của các dải hệ số tỷ lệ tương ứng của bộ giải mã âm thanh lõi hoặc có kích cỡ sao cho, khi việc tạo cặp năng lượng được áp dụng, giá trị năng lượng đối với dải khôi phục cung cấp năng lượng của hai hoặc số nguyên lớn hơn của các dải hệ số tỷ lệ. Do đó, khi giả sử rằng sự tích tụ năng lượng được thực hiện đối với dải hệ số tỷ lệ 4, dải hệ số tỷ lệ 5 và dải hệ số tỷ lệ 6, khi đó đường biên tần số thấp hơn của dải khôi phục 920 bằng đường biên thấp hơn của dải hệ số tỷ lệ 4 và đường biên tần số cao hơn của dải khôi phục 920 trùng khớp với đường biên cao hơn của dải hệ số tỷ lệ 6.

Sau đây, Fig.9d được mô tả để thể hiện các chức năng khác của bộ giải mã của Fig.9a. Bộ giải mã âm thanh 900 nhận các giá trị phổ được giải lượng tử hoá tương ứng với các phần phổ thứ nhất của tập hợp thứ nhất của các phần phổ và ngoài ra, các hệ số tỷ lệ đối với các dải hệ số tỷ lệ như được minh họa trên Fig.3b được cung cấp vào khối tạo tỷ lệ nghịch đảo 940. Khối tạo tỷ lệ nghịch đảo 940 cung cấp tất cả các tập hợp thứ nhất của các phần phổ thứ nhất thấp hơn tần số bắt đầu IGF 309 của Fig.3a và ngoài ra, các phần phổ thứ nhất cao hơn tần số bắt đầu IGF, nghĩa là các phần phổ thứ nhất 304, 305, 306, 307 của Fig.3a mà tất cả được định vị trong dải khôi phục như được minh họa ở 941 trên Fig.9d. Hơn nữa, các phần phổ thứ nhất trong dải nguồn được sử dụng để điền đầy lát tần số trong dải khôi phục được cung cấp cho bộ điều chỉnh/bộ tính toán đường bao 942 và khối này ngoài ra nhận thông tin năng lượng đối với dải khôi phục được cung cấp như thông tin phụ tham số vào tín hiệu âm thanh được mã hoá như được minh họa ở 943 trên Fig.9d. Khi đó, bộ điều chỉnh/bộ tính toán đường bao 942 cung cấp các chức năng của Fig.9b và Fig.9c và cuối cùng xuất ra các



giá trị phổ được điều chỉnh đối với các phần phổ thứ hai trong ô khôi phục. Các giá trị phổ điều chỉnh này 922, 923 đối với các phần phổ thứ hai trong dải khôi phục và các phần phổ thứ nhất 921 trong dải khôi phục được biểu thị rằng đường dẫn 941 trên Fig.9d biểu diễn theo cách ghép phép biểu diễn phổ hoàn toàn của dải khôi phục.

Tiếp theo, việc tham khảo được thực hiện đến các hình vẽ từ Fig.10a đến 10b để giải thích các phương án được ưu tiên của bộ mã hoá âm thanh để mã hoá tín hiệu âm thanh để cung cấp hoặc tạo ra tín hiệu âm thanh được mã hoá. Bộ mã hoá bao gồm bộ chuyển đổi thời gian/phổ 1002 nạp cho bộ phân tích phổ 1004, và bộ phân tích phổ 1004 một mặt được kết nối với bộ tính toán tham số 1006 và mặt khác bộ mã hoá âm thanh 1008. Bộ mã hoá âm thanh 1008 cung cấp phép biểu diễn được mã hoá của tập hợp thứ nhất của các phần phổ thứ nhất và không bao phủ tập hợp thứ hai của các phần phổ thứ hai. Theo cách khác, bộ tính toán tham số 1006 cung cấp thông tin năng lượng cho dải khôi phục bao phủ các phần phổ thứ nhất và thứ hai. Hơn nữa, bộ mã hoá âm thanh 1008 được tạo cấu hình để tạo ra phép biểu diễn được mã hoá thứ nhất của tập hợp thứ nhất của các phần phổ thứ nhất có độ phân giải phổ thứ nhất, trong đó bộ mã hoá âm thanh 1008 cung cấp các hệ số tỷ lệ đối với tất cả các dải của phép biểu diễn phổ được tạo ra bởi khối 1002. Ngoài ra, như được minh họa trên Fig.3b, bộ mã hoá cung cấp thông tin năng lượng ít nhất cho các dải khôi phục được định vị, đối với tần số, cao hơn tần số bắt đầu IGF 309 như được minh họa trên Fig.3a. Do đó, đối với các dải khôi phục ưu tiên trùng khớp với các dải hệ số tỷ lệ hoặc với các nhóm gồm các dải hệ số tỷ lệ, hai giá trị đã cho, nghĩa là hệ số tỷ lệ tương ứng từ bộ mã hoá âm thanh 1008 và, ngoài ra, thông tin năng lượng được xuất ra bởi bộ tính toán tham số 1006.

Tốt hơn là, bộ mã hoá âm thanh có các dải hệ số tỷ lệ với các băng thông tần số khác nhau, nghĩa là với số lượng khác nhau của các giá trị phổ. Do đó, bộ tính toán tham số bao gồm bộ chuẩn hoá 1012 để chuẩn hoá các năng lượng của băng thông khác nhau với băng thông của dải khôi phục đặc hiệu. Với mục tiêu này, bộ chuẩn hoá 1012 nhận, dưới dạng đầu vào, năng lượng trong dải và số lượng của các giá trị phổ trong dải bộ chuẩn hoá 1012 khi đó xuất ra năng lượng được chuẩn hoá cho mỗi dải khôi phục/hệ số tỷ lệ.

Hơn nữa, bộ tính toán tham số 1006a của Fig.10a bao gồm bộ tính toán giá trị năng lượng nhận thông tin điều khiển từ bộ mã hoá lỗi hoặc âm thanh 1008 như được minh họa bởi đường dẫn 1007 trên Fig.10a. Thông tin điều khiển này có thể bao gồm thông tin về các khối dài/ngắn được sử dụng bởi bộ mã hoá âm thanh và/hoặc thông tin tạo nhóm. Do đó, trong khi thông tin về các khối dài/ngắn và thông tin tạo nhóm về các cửa sổ ngắn liên quan đến việc tạo nhóm “thời gian”, ngoài ra thông tin tạo nhóm có thể đề cập đến việc tạo nhóm phổ, nghĩa là việc tạo nhóm của hai dải hệ số tỷ lệ vào trong dải khôi phục đơn lẻ. Do đó, bộ tính toán giá trị năng lượng 1014 xuất ra giá trị năng lượng đơn lẻ đối với mỗi dải được tạo nhóm bao phủ phần phổ thứ nhất và thứ hai khi chỉ có phần phổ được tạo nhóm.

Fig.10d minh họa phương án khác để thực thi việc tạo nhóm phổ. Với mục tiêu này, khối 1016 được tạo cấu hình để tính toán các giá trị năng lượng đối với hai dải liền kề. Khi đó, trong khối 1018, các giá trị năng lượng đối với các dải liền kề được so sánh và, khi các giá trị năng lượng không khác biệt quá nhiều hoặc khác biệt ít hơn so với được xác định bởi, ví dụ, ngưỡng, khi đó giá trị đơn lẻ (chuẩn hoá) cho cả hai dải được tạo ra như được biểu thị trong khối 1020. Như được minh họa bởi đường dẫn 1019, khối 1018 có thể được thông qua. Hơn nữa, việc tạo ra giá trị đơn lẻ cho hai hoặc nhiều dải được thực hiện bởi khối 1020 có thể được điều khiển bởi điều khiển tốc độ bit của bộ mã hoá 1024. Do đó, khi tốc độ bit giảm, điều khiển tốc độ bit được mã hoá 1024 điều khiển khối 1020 để tạo ra giá trị chuẩn hoá đơn lẻ cho hai hoặc nhiều dải cho dù việc so sánh trong khối 1018 sẽ không được cho phép để tạo nhóm các giá trị thông tin năng lượng.

Trong trường hợp bộ mã hoá âm thanh thực hiện việc tạo nhóm hai hoặc nhiều cửa sổ ngắn, việc tạo nhóm này cũng được áp dụng cho thông tin năng lượng. Khi bộ mã hoá lỗi thực hiện việc tạo nhóm hai hoặc nhiều khối ngắn, khi đó đối với hai hoặc nhiều khối này, chỉ tập hợp đơn lẻ của các hệ số tỷ lệ được tính toán và truyền. Ở phía bộ giải mã, bộ giải mã âm thanh khi đó áp dụng cùng tập hợp của các hệ số tỷ lệ cho cả hai cửa sổ được tạo nhóm.

Đối với việc tính toán thông tin năng lượng, các giá trị phổ trong dải khôi phục được tích tụ qua hai hoặc nhiều cửa sổ ngắn. Nói cách khác, việc này có nghĩa là các

giá trị phổ trong dải khôi phục nhất định đối với khối ngắn và đối với khối ngắn tiếp theo được tích tụ cùng với nhau và chỉ thông tin năng lượng đơn lẻ được truyền cho dải khôi phục bao phủ hai khối ngắn. Khi đó, ở phía bộ giải mã, việc điều chỉnh đường bao được mô tả đối với các hình vẽ từ Fig.9a đến Fig.9d không được thực hiện một cách riêng rẽ cho mỗi khối ngắn mà được thực hiện cùng nhau đối với tập hợp các cửa sổ ngắn được nhóm.

Khi đó, sự chuẩn hoá tương ứng được áp dụng một lần nữa sao cho cho dù việc tạo nhóm bất kỳ theo tần số hoặc tạo nhóm theo thời gian được thực hiện, việc chuẩn hoá dễ dàng cho phép rằng, đối với việc tính toán thông tin giá trị năng lượng ở phía bộ giải mã, một mặt chỉ giá trị thông tin năng lượng và số lượng các đường dẫn phổ trong dải khôi phục hoặc trong tập hợp của các dải khôi phục được nhóm phải được biết đến đến.

Trong các sơ đồ BWE theo tình trạng kỹ thuật, sự khôi phục của vùng phổ HF cao hơn tần số chéo đã cho thường được dựa trên phép vá phổ. Thông thường, vùng HF gồm có nhiều bản vá liền kề và mỗi trong số các bản vá này được bắt nguồn từ các vùng thông dải (BP) của phổ LF thấp hơn tần số chéo đã cho. Trong phép biểu diễn giàn lọc của tín hiệu, các hệ thống này sao chép tập hợp các hệ số dải phụ liền kề trong số các phổ LF vào trong vùng đích. Các đường biên của các tập hợp được chọn thường phụ thuộc vào hệ thống và không phụ thuộc vào tín hiệu. Đối với một vài nội dung tín hiệu, việc chọn lọc và tính điện này có thể dẫn đến âm sắc và sắc thái không dễ chịu của tín hiệu được khôi phục.

Các phương pháp khác truyền tín hiệu LF vào HF qua sự điều biến băng con đơn lẻ (Single Side Band - SSB) thích ứng. Các phương pháp này có độ phức tạp tính toán cao so với [1] do các phương pháp này hoạt động ở tốc độ lấy mẫu cao ở các mẫu miền thời gian. Ngoài ra, phép vá có thể không ổn định, đặc biệt đối với các tín hiệu không âm (ví dụ, tiếng nói vô thanh) và nhờ đó phép vá thích ứng tín hiệu theo tình trạng kỹ thuật có thể đưa đến các suy yếu vào trong tín hiệu.

Phương pháp của sáng chế được gọi là điền đầy khoảng trống thông minh (IGF) và, trong cấu hình được ưu tiên của nó, phương pháp này được áp dụng trong hệ thống BWE dựa trên biến đổi thời gian-tần số, tương tự ví dụ biến đổi cosin rời rạc cải biên

(MDCT). Tuy nhiên, phần mô tả của sáng chế thường có thể áp dụng, ví dụ theo cách tương tự trong hệ thống dựa trên giàn lọc gương cầu phương (Quadrature Mirror Filterbank - QMF).

Ưu điểm của cấu hình IGF dựa trên MDCT là sự tích hợp liền khối vào trong các bộ mã hóa âm thanh dựa trên MDCT, ví dụ mã hóa âm thanh tiên tiến (Advanced Audio Coding - AAC) MPEG. Sự chia sẻ cùng biến đổi đối với việc mã hóa âm thanh dạng sóng và đối với BWE làm giảm tổng thể độ phức tạp tính toán đối với bộ mã hóa-giải mã âm thanh một cách đáng kể.

Hơn nữa, sáng chế đề xuất giải pháp cho các vấn đề ổn định cổ hữu được tìm thấy trong các sơ đồ vá thích ứng theo tình trạng kỹ thuật.

Hệ thống được đề xuất được dựa trên quan sát rằng đối với một vài tín hiệu, việc chọn lọc và không định hướng có thể dẫn đến các thay đổi âm sắc và các sắc thái tín hiệu. Nếu tín hiệu mà là âm trong vùng nguồn phổ (spectral source region - SSR) nhưng tương tự nhiều âm trong vùng đích phổ (spectral target region - STR), việc vá STR tương tự nhiều âm bởi SSR âm có thể dẫn đến âm sắc không tự nhiên. Âm sắc của tín hiệu cũng có thể thay đổi do cấu trúc âm của tín hiệu có thể bị sắp hàng lỗi hoặc thậm chí bị phá hủy bởi quy trình vá.

Hệ thống IGF được đề xuất thực hiện việc lựa chọn ô thông minh bằng cách sử dụng tương quan chéo như phép đo tính tương tự giữa SSR cụ thể và STR đặc hiệu. Sự tương quan chéo của hai tín hiệu cung cấp phép đo về tính tương tự của các tín hiệu này và cả độ trễ của tương quan tối đa và tín hiệu của nó. Do đó, phương pháp về sự tương quan dựa trên lựa chọn ô cũng có thể được sử dụng để điều chỉnh một cách chính xác độ lệch phổ của phổ sao chép để trở thành gần nhất có thể cấu trúc phổ ban đầu.

Sự đóng góp cơ bản của hệ thống được đề xuất là sự chọn lọc của phép đo tính tương tự thích hợp và cả các kỹ thuật để ổn định hóa quy trình lựa chọn ô. Kỹ thuật được đề xuất cung cấp sự cân bằng tối ưu giữa sự thích ứng tín hiệu tức thời và đồng thời, độ ổn định theo thời gian. Việc cung cấp độ ổn định theo thời gian là đặc biệt quan trọng đối với các tín hiệu mà có tính tương tự ít hơn của SSR và STR và do đó thể hiện các giá trị tương quan chéo thấp hoặc nếu các phép đo tính tương tự được

dùng mà không rõ ràng. Trong các trường hợp này, việc ổn định hoá ngăn ngừa đặc tính giả ngẫu nhiên của việc lựa chọn ô thích ứng.

Ví dụ, nhóm các tín hiệu mà thường đặt ra các vấn đề đối với BWE theo tình trạng kỹ thuật được đặc trưng bởi sự tập trung khác biệt của năng lượng vào các vùng phổ tùy ý, như được thể hiện trên Fig.12a (bên trái). Mặc dù có các phương pháp sẵn có để điều chỉnh đường bao phổ và âm của phổ khôi phục trong vùng đích, đối với một vài 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 được thể hiện trên Fig.12a, biên độ của phổ trong vùng đích của tín hiệu ban đầu cao hơn tần số chéo  $f_{\text{cross}}^*$  (Fig.12a, bên trái) giảm gần như một cách tuyến tính. Ngược lại, trong phổ khôi phục (Fig.12a, bên phải), tập hợp khác biệt của các đáy và các đỉnh có mặt mà được cảm nhận như giả âm sắc thái hóa âm sắc.

Bước quan trọng của phương pháp mới là để xác định các ô trong số chúng tính tương tự tiếp theo dựa trên việc lựa chọn có thể được thực hiện. Trước tiên, các đường biên ô của cả vùng nguồn và vùng đích phải được xác định phù hợp với nhau. Do đó, vùng đích giữa tần số bắt đầu IGF của bộ mã hóa lõi  $f_{IGFstart}$  và tần số sẵn có cao nhất  $f_{IGFstop}$  được chia thành số nguyên tùy ý  $nTar$  của các ô, mỗi trong số các ô này có kích cỡ định trước riêng rẽ. Khi đó, đối với mỗi lát đích  $tar[idx\_tar]$ , tập hợp của các lát nguồn được định cỡ bằng nhau  $src[idx\_src]$  được tạo ra. Bằng cách này, mức tự do cơ bản của hệ thống IGF được xác định. Tổng số các lát nguồn  $nSrc$  được xác định theo băng thông của vùng nguồn,

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

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

Để tăng hơn nữa mức tự do để chọn lọc và điều chỉnh, các lát nguồn có thể được xác định để chồng lấp với nhau bởi hệ số chồng lấp nằm trong khoảng từ 0 đến 1, trong đó 0 có nghĩa là không chồng lấp và 1 có nghĩa là 100% chồng lấp. Trường hợp 100% chồng lấp ngụ ý rằng chỉ có một hoặc không có lát nguồn nào là sẵn có.

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

Đối với lát đích, sự tương quan chéo được tính toán với các lát nguồn khác nhau ở các độ trễ tới các ngăn *xcorr\_maxLag*. Đối với lát đích đã cho *idx\_tar* và lát nguồn *idx\_src*, *xcorr\_val[idx\_tar][idx\_src]* đưa ra giá trị tối đa của tương quan chéo tuyệt đối giữa các ô, trong khi *xcorr\_lag[idx\_tar][idx\_src]* đưa ra độ trễ mà ở đó sự tối đa này xảy ra và *xcorr\_sign[idx\_tar][idx\_src]* đưa ra dấu hiệu tương quan chéo tại *xcorr\_lag[idx\_tar][idx\_src]*.

Tham số *xcorr\_lag* được sử dụng để điều khiển tính chính xác của độ phù hợp giữa các lát nguồn và đích. Tham số này dẫn đến các giả âm giảm và hỗ trợ tốt hơn để bảo toàn âm sắc và sắc điệu của tín hiệu.

Trong một vài kịch bản, có thể xảy ra rằng kích cỡ của lát đích đặc hiệu lớn hơn so với kích cỡ của các lát nguồn sẵn có. Trong trường hợp này, lát nguồn sẵn có được lặp lại thường xuyên như cần thiết để điền đầy lát đích đặc hiệu một cách hoàn toàn. Cũng có thể thực hiện sự tương quan chéo giữa lát đích lớn và lát nguồn nhỏ hơn để có vị trí tốt nhất của lát nguồn trong lát đích dưới dạng độ trễ tương quan chéo *xcorr\_lag* và tín hiệu *xcorr\_sign*.

Sự tương quan chéo của các ô phổ thô và tín hiệu ban đầu có thể không phải là phép đo tính tương tự thích hợp nhất được áp dụng cho các phổ âm thanh với cấu trúc thành tổ mạnh. Việc làm trắng phổ loại bỏ thông tin đường bao thô và do đó nhấn mạnh cấu trúc mịn phổ, mà quan tâm trước hết để đánh giá tính tương tự ô. Việc làm trắng cũng trợ giúp cho việc tạo hình đường bao dễ dàng của STR ở bộ giải mã đối với các vùng được xử lý bởi IGF. Do đó, một cách tùy ý, ô và tín hiệu nguồn được làm trắng trước khi tính toán sự tương quan chéo.

Trong các cấu hình khác, chỉ ô được làm trắng bằng cách sử dụng quy trình định trước. Độ trễ “làm trắng” được truyền biểu thị cho bộ giải mã rằng cùng quy trình làm trắng định trước sẽ được áp dụng cho ô trong IGF.

Để làm trắng tín hiệu, trước tiên sự ước lượng đường bao phổ được tính toán. Sau đó, phổ MDCT được chia theo đường bao phổ. Việc ước lượng đường bao phổ có thể được ước lượng trên phổ MDCT, các năng lượng phổ MDCT, phổ công suất phức hợp dựa trên MDCT hoặc các ước lượng phổ công suất. Tín hiệu mà trên đó đường bao được ước lượng sẽ được gọi là tín hiệu cơ sở từ phần này trở về sau.

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

Nếu tín hiệu cơ sở nằm trong miền năng lượng, phổ MDCT phải được chia theo 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 sở với biến đổi cosin rời rạc (DCT), chỉ giữ lại các hệ số DCT thấp hơn (thiết lập cao nhất đến 0) và sau đó tính toán DCT nghịch đảo
- tính toán đường bao phổ của tập hợp các hệ số dự đoán tuyến tính (Linear Prediction Coefficients - LPC) được tính toán trên khung âm thanh miền thời gian
- lọc tín hiệu cơ sở bằng bộ lọc thông thấp

Tốt hơn là, phương pháp cuối cùng được chọn. Đối với các ứng dụng mà yêu cầu độ phức tạp tính toán thấp, một vài đơn giản hoá có thể được thực hiện để làm trắng phổ MDCT: Trước tiên, đường bao được tính toán bằng phương pháp số trung bình động. Điều này chỉ cần có hai chu kỳ xử lý cho mỗi ngăn MDCT. Sau đó, để ngăn ngừa sự tính toán phân chia và căn bậc hai, đường bao phổ bằng khoảng  $2^n$ , trong đó  $n$  là loga số nguyên của đường bao. Trong miền này, thao tác lấy căn bậc hai đơn giản trở thành thao tác chuyển dịch và hơn nữa, sự phân chia theo đường bao có thể được thực hiện theo thao tác chuyển dịch khác.

Sau khi tính toán sự tương quan của mỗi lát nguồn với mỗi lát đích, đối với tất cả các lát đích  $nTar$ , lát nguồn có độ tương quan cao nhất được chọn để thay thế lát này. Để phù hợp cấu trúc phổ ban đầu tốt nhất, độ trễ tương quan được sử dụng để điều biến phổ được sao chép theo số nguyên của các ngăn biến đổi. Trong trường hợp

các độ trễ lẻ, lát còn được điều biến qua việc nhân theo dãy theo thời gian thay thế  $-1/1$  để bù cho phép biểu diễn tần số – nghịch đảo của mỗi dải khác trong MDCT.

Fig.12c thể hiện ví dụ về sự tương quan giữa lát nguồn và lát đích. Trong ví dụ này, độ trễ tương quan bằng 5, do đó lát nguồn phải được điều biến theo 5 ngăn tới các ngăn tần số cao hơn trong giai đoạn sao chép của thuật toán BWE. Ngoài ra, tín hiệu của lát phải được lật do giá trị tương quan tối đa là âm và sự điều biến bổ sung như được mô tả trên đây tính cho độ trễ lẻ.

Do đó, tổng lượng thông tin phụ để truyền từ bộ mã hoá đến bộ giải mã có thể gồm có các dữ liệu sau đây:

- $tileNum[nTar]$ : chỉ số lát nguồn được chọn cho mỗi lát đích
- $tileSign[nTar]$ : tín hiệu của lát đích
- $tileMod[nTar]$ : độ trễ tương quan cho mỗi lát đích

Việc chia nhỏ và ổn định hoá lát là bước quan trọng trong IGF. Nhu cầu và các ưu điểm của nó được giải thích bằng một ví dụ, giả sử tín hiệu âm thanh âm tĩnh như ví dụ một nốt nhạc cao độ ổn định. Logic cho biết rằng các giả âm ít nhất được đưa đến nếu, đối với vùng đích đã cho, các lát nguồn thường được chọn từ cùng các khung chéo vùng nguồn. Cho dù tín hiệu được giả định là tĩnh, điều kiện này sẽ không được giữ tốt trong mỗi khung do phép đo tính tương tự (ví dụ, sự tương quan) của vùng nguồn tương tự cân bằng khác có thể chiếm ưu thế kết quả tính tương tự (ví dụ, tương quan chéo). Việc này dẫn đến  $tileNum[nTar]$  giữa các khung liên kế để làm dao động giữa hai hoặc ba chọn lọc rất tương tự. Việc này có thể là nguồn của nhiễu âm âm nhạc khó chịu như giả âm.

Để loại bỏ loại các giả âm này, tập hợp các lát nguồn sẽ được chia nhỏ sao cho các thành viên còn lại của tập hợp nguồn không tương tự một cách tối đa. Việc này đạt được qua tập hợp các lát nguồn

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

như sau. Đối với lát nguồn bất kỳ  $s_i$ , các tác giả sáng chế tương quan lát nguồn với tất cả các lát nguồn khác, phát hiện sự tương quan tốt nhất giữa  $s_i$  và  $s_j$  và lưu trữ lát này



trong ma trận  $S_x$ . Ở đây,  $S_x[i][j]$  chứa giá trị tương quan chéo tuyệt đối tối đa giữa  $s_i$  và  $s_j$ . Bổ sung ma trận  $S_x$  dọc theo các cột, thu được tổng tương quan chéo của lát nguồn  $s_i$  với tất cả các lát nguồn khác  $T$ .

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

Ở đây,  $T$  biểu diễn phép đo về sự tương tự của nguồn với các lát nguồn khác. Nếu, đối với lát nguồn bất kỳ  $i$ ,

$$T > \text{ngưỡng}$$

lát nguồn  $i$  có thể được thả từ tập hợp các nguồn tiềm năng do lát nguồn này có tương quan ở mức cao với các nguồn khác. Lát có sự tương quan thấp nhất từ tập hợp các lát mà thoả mãn điều kiện trong phương trình 1 được chọn là lát đại diện cho tập hợp phụ này. Bằng cách này, các tác giả sáng chế đảm bảo rằng các lát nguồn không tương tự tối đa với nhau.

Phương pháp chia nhỏ lát cũng bao gồm bộ nhớ tập hợp lát chia nhỏ được sử dụng trong khung trước. Các lát mà hoạt động trong khung trước được giữ lại trong khung tiếp theo nếu các ứng viên thay thế để chia nhỏ tồn tại.

Coi các lát  $s_3$ ,  $s_4$  và  $s_5$  hoạt động trong số các lát  $\{s_1, s_2, \dots, s_5\}$  trong khung  $k$ , khi đó trong khung  $k+1$  ngay cả khi các lát  $s_1$ ,  $s_3$  và  $s_2$  chắc chắn sẽ được chia nhỏ với  $s_3$  có tương quan tối đa với lát khác,  $s_3$  được giữ lại do lát này là lát nguồn hữu ích trong khung trước và do đó giữ lại lát này trong tập hợp các lát nguồn có lợi để thi hành tính liên tục theo thời gian trong lựa chọn lát. Phương pháp này ưu tiên được áp dụng nếu sự tương quan chéo giữa nguồn  $i$  và mục tiêu  $j$ , được biểu diễn như  $T_x[i][j]$  là ở mức cao.

Phương pháp bổ sung để ổn định hoá lát là giữ lại thứ tự lát từ khung trước  $k-1$  nếu không lát nguồn nào trong khung hiện thời  $k$  tương quan tốt với các lát đích. Việc này có thể xảy ra nếu sự tương quan chéo giữa nguồn  $i$  và mục tiêu  $j$ , được biểu diễn như  $T_x[i][j]$  là rất thấp đối với tất cả  $i, j$

Ví dụ, nếu

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

ngưỡng thử giờ đây sẽ được sử dụng, khi đó

$$\text{tileNum}[\text{nTar}]_k = \text{tileNum}[\text{nTar}]_{k-1}$$

đối với tất cả nTar của khung k này.

Hai kỹ thuật trên đây giảm mạnh các giả âm mà xuất hiện do sự thay đổi nhanh chóng số lượng ô tập hợp qua các khung. Ưu điểm bổ sung khác của việc chia nhỏ và ổn định hoá lát này ở chỗ không cần thông tin phụ nào phải được gửi đến bộ giải mã và cũng không cần sự thay đổi của kết cấu bộ giải mã. Việc chia nhỏ lát được đề xuất này là cách tạo nhả để giảm nhiễu âm âm nhạc tiềm năng như các giả âm hoặc nhiễu âm quá mức trong các vùng phổ ô.

Fig.11a minh họa bộ giải mã âm thanh để giải mã tín hiệu âm thanh được mã hoá. 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 các phần phổ thứ nhất, phép biểu diễn được giải mã có độ phân giải phổ thứ nhất.

Hơn 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 các phần phổ thứ hai có độ phân giải phổ thứ hai thấp hơn so với độ phân giải phổ thứ nhất. Hơn nữa, bộ tái tạo tần số 1106 được cung cấp mà nhận, dưới dạng đầu vào thứ nhất 1101, các phần phổ thứ nhất được giải mã và như là đầu vào thứ hai ở 1103 thông tin tham số bao gồm, đối với mỗi lát tần số mục tiêu hoặc dải khôi phục mục tiêu, thông tin khoảng nguồn. Sau đó, bộ tái tạo tần số 1106 áp dụng việc tái tạo tần số bằng cách sử dụng các giá trị phổ từ khoảng nguồn được nhận biết bởi thông tin so khớp để tạo ra dữ liệu phổ đối với khoảng đích. Khi đó, các phần phổ thứ nhất 1101 và đầu vào của bộ tái tạo tần số 1107 đều được đưa vào trong 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 cũng có thể được thực thi như bộ giải mã âm thanh khác bất kỳ như bộ giải mã miền thời gian hoặc âm thanh tham số.

Như được biểu thị ở Fig.11b, bộ tái tạo tần số 1106 có thể bao gồm các chức năng của khối 1120 minh họa bộ chọn lọc khoảng nguồn - bộ điều biến ô cho các độ

trễ lẹ, bộ lọc làm trắng 1122, khi ký hiệu làm trắng 1123 được cung cấp và ngoài ra, đường bao phổ với các chức năng điều chỉnh thực thi được minh họa trong khối 1128 bằng cách sử dụng dữ liệu phổ thô được tạo ra bởi khối 1120 hoặc khối 1122 hoặc sự hợp tác của cả hai khối. Dù sao, bộ tái tạo tần số 1106 có thể bao gồm chuyển đổi 1124 phản ứng với ký hiệu làm trắng nhận được 1123. Khi ký hiệu làm trắng được thiết lập, đầu ra của bộ chọn lọc khoảng nguồn/bộ điều biến ô cho độ trễ lẹ được đưa vào trong bộ lọc làm trắng 1122. Tuy nhiên, khi đó ký hiệu làm trắng 1123 không được thiết lập cho dải khôi phục nhất định, khi đó đường dẫn thông qua 1126 được hoạt hoá sao cho đầu vào của khối 1120 được cung cấp cho khối điều chỉnh đường bao phổ 1128 mà không làm trắng bất kỳ.

Có thể có nhiều hơn một mức làm trắng (1123) được tín hiệu hóa trong dòng bit và các mức này có thể được tín hiệu hóa cho mỗi ô. Trong trường hợp có ba mức tín hiệu hóa cho mỗi ô, chúng sẽ được mã hóa theo 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 cho các bộ lọc làm trắng khác nhau (1122) mà có thể khác nhau theo cách mà đường bao được tính toán (như được mô tả trên đây).

Bộ giải mã - bộ tái tạo tần số phụ có thể được điều khiển bởi khoảng nguồn ID 1121 khi chỉ sơ đồ lựa chọn ô phổ thô được áp dụng. Tuy nhiên, khi sơ đồ lựa chọn ô phổ mịn – giai điệu được áp dụng, khi đó, ngoài ra, độ trễ khoảng nguồn 1119 được cung cấp. Hơn nữa, nếu việc tính toán sự tương quan tạo ra kết quả âm, khi đó ngoài ra, tín hiệu tương quan cũng có thể được áp dụng cho khối 1120 sao cho mỗi đường dẫn phổ dữ liệu trang được nhân với “-1” để tính toán tín hiệu âm.

Do đó, sáng chế như được minh họa trên Fig.11a, Fig.11b đảm bảo rằng chất lượng âm thanh tối ưu thu được do thực tế rằng khoảng nguồn phù hợp tốt nhất cho

khoảng đích hoặc mục tiêu nhất định được tính toán ở phía bộ mã hoá và được áp dụng ở phía bộ giải mã.

Fig.11c là bộ mã hoá âm thanh nhất định để mã hoá tín hiệu âm thanh bao gồm bộ chuyển đổi thời gian-phổ 1130, bộ phân tích phổ được kết nối một cách nối tiếp 1132 và ngoài ra, bộ tính toán tham số 1134 và bộ mã hóa lỗi 1136. Bộ mã hóa lỗi 1136 xuất ra các khoảng nguồn được mã hoá và bộ tính toán tham số 1134 xuất ra thông tin so khớp cho các khoảng đích.

Các khoảng nguồn được mã hoá được truyền đến bộ giải mã cùng với thông tin khớp cho các khoảng đích sao cho bộ giải mã được minh họa trên Fig.11a ở vị trí để thực hiện việc tái tạo tần số.

Bộ tính toán tham số 1134 được tạo cấu hình để tính toán tính tương tự giữa các phần phổ thứ nhất và các phần phổ thứ hai và để xác định, dựa trên các tính tương tự được tính toán, đối với phần phổ thứ hai phù hợp phần phổ thứ nhất phù hợp với phần phổ thứ hai. Tốt hơn là, các kết quả phù hợp đối với các khoảng nguồn và các khoảng đích khác nhau như được minh họa trên các hình vẽ Fig.12a, Fig.12b để xác định cặp phù hợp được chọn bao gồm phần phổ thứ hai, và bộ tính toán tham số được tạo cấu hình để cung cấp thông tin so khớp này nhận biết cặp phù hợp vào trong tín hiệu âm thanh được mã hoá. Tốt hơn là, bộ tính toán tham số 1134 được tạo cấu hình để sử dụng các vùng đích định trước trong tập hợp thứ hai của các phần phổ thứ hai hoặc các vùng nguồn định trước trong tập hợp thứ nhất của các phần phổ thứ nhất như được minh họa, ví dụ trên Fig.12b. Tốt hơn là, các vùng đích định trước không bị chồng lấp hoặc các vùng nguồn định trước là chồng lấp. Khi các vùng nguồn định trước là tập hợp con của tập hợp thứ nhất của các phần phổ thứ nhất thấp hơn tần số bắt đầu điền đầy khoảng trống 309 của Fig.3a, và tốt hơn là, vùng đích định trước bao phủ vùng phổ thấp hơn trùng khớp, với đường biên tần số thấp hơn của vùng này với tần số bắt đầu điền đầy khoảng trống sao cho các khoảng đích bất kỳ được định vị cao hơn tần số bắt đầu điền đầy khoảng trống và các khoảng nguồn được định vị thấp hơn tần số bắt đầu điền đầy khoảng trống.

Như được mô tả, độ chi tiết mịn thu được bằng cách so sánh vùng đích với vùng nguồn mà không có độ trễ bất kỳ đối với vùng nguồn và cùng vùng nguồn, nhưng có

độ trễ nhất định. Các độ trễ này được áp dụng trong bộ tính toán tương quan chéo 1140 của Fig.11d và việc lựa chọn cặp phù hợp cuối cùng được thực hiện bởi bộ lựa chọn ô 1144.

Hơn nữa, tốt hơn là thực hiện việc làm trắng các khoảng nguồn và/hoặc mục tiêu đích được minh họa ở khối 1142. Khi đó, khối 1142 này cung cấp ký hiệu làm trắng cho dòng bit mà được sử dụng để điều khiển chuyển đổi phía bộ giải mã 1123 của Fig.11b. Hơn nữa, nếu bộ tính toán tương quan chéo 1140 cung cấp kết quả âm, khi đó kết quả âm này cũng được tín hiệu hóa đến bộ giải mã. Do đó, theo phương án được ưu tiên, bộ lựa chọn ô xuất ra khoảng nguồn ID cho khoảng đích, độ trễ, tín hiệu và khối 1142 còn cung cấp ký hiệu làm trắng.

Hơn nữa, bộ tính toán tham số 1134 được tạo cấu hình để thực hiện chia nhỏ lát nguồn 1146 bằng cách giảm số lượng các khoảng nguồn tiềm năng trong đó bản vá nguồn được thả từ tập hợp của các lát nguồn tiềm năng dựa trên ngưỡng tính tương tự. Do đó, khi hai lát nguồn là tương tự hơn hoặc bằng ngưỡng tính tương tự, khi đó một trong hai lát nguồn này được bị bỏ ra khỏi tập hợp nguồn tiềm năng và lát nguồn bị loại bỏ không được sử dụng tiếp để xử lý thêm và cụ thể là, không thể được chọn lọc bởi bộ lựa chọn ô 1144 hoặc không được sử dụng để tính toán sự tương quan chéo giữa các khoảng nguồn và các khoảng đích khác nhau như được thực hiện trong khối 1140.

Các phương án khác nhau đã được mô tả đối với các hình vẽ khác nhau. Các hình vẽ từ Fig.1a đến Fig.5c đề cập đến sơ đồ bộ mã hoá/bộ giải mã bằng thông tốc độ đầy đủ hoặc đầy đủ. Các hình vẽ từ Fig.6a đến Fig.7e đề cập đến sơ đồ bộ mã hoá/bộ giải mã với xử lý TNS hoặc TTS. Các hình vẽ từ Fig.8a đến Fig.8e đề cập đến sơ đồ bộ mã hoá/bộ giải mã với xử lý kênh đôi đặc hiệu. Các hình vẽ từ Fig.9a đến Fig.10d đề cập đến tính toán và áp dụng thông tin năng lượng đặc hiệu và các hình vẽ từ Fig.11a đến Fig.12c đề cập đến cách đặc hiệu để lựa chọn ô.

Tất cả các khía cạnh khác nhau này có thể sử dụng độc lập với nhau theo sáng chế, nhưng ngoài ra, cũng có thể được áp dụng cùng với nhau như được minh họa một cách cơ bản trên các hình vẽ Fig.2a và Fig.2b. Tuy nhiên, việc xử lý kênh đôi đặc hiệu có thể được áp dụng cho sơ đồ bộ mã hoá/bộ giải mã được minh họa trên Fig.13 và

điều tương tự cũng đúng đối với việc xử lý TNS/TTS, việc tính toán và áp dụng thông tin năng lượng đường bao trong dải khôi phục hoặc nhận biết khoảng nguồn thích ứng và áp dụng tương ứng ở phía bộ giải mã. Theo cách khác, khía cạnh tốc độ đầy đủ có thể được áp dụng với hoặc không với xử lý TNS/TTS, với hoặc không với xử lý kênh đôi, với hoặc không với nhận biết khoảng nguồn thích ứng hoặc với các kiểu của các tính toán năng lượng khác đối với phép biểu diễn đường bao phổ. Do đó, rõ ràng rằng các dấu hiệu của một trong các khía cạnh riêng rẽ này cũng có thể được áp dụng trong các khía cạnh khác.

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

Tùy thuộc vào các yêu cầu triển khai nhất định, các phương án của sáng chế có thể được triển khai trong phần cứng hoặc phần mềm. Việc triển khai có thể được thực hiện bằng cách sử dụng phương tiện lưu trữ không tạm thời như phương tiện lưu trữ số, ví dụ đĩa mềm, ổ cứng (Hard Disk Drive - HDD), DVD, Blu-Ray, CD, ROM, PROM, và EPROM, EEPROM hoặc bộ nhớ FLASH, có các tín hiệu điều khiển có thể đọc được bằng điện được lưu trữ trên đó, mà kết hợp (hoặc có thể kết hợp) với hệ thống máy tính khả trình sao cho phương pháp tương ứng được thực hiện. Do đó, phương tiện lưu trữ 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 đọc được nhờ điện tử, chúng có khả năng kết hợp với hệ thống máy tính lập trình được, sao cho một trong các phương pháp được mô tả trong mô ở đây được thực hiện.

Thông thường, các phương án theo sáng chế có thể được triển khai như 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 ví dụ có thể được lưu trữ trên vật mang đọc được bằng máy.

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

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

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

Do đó, phương án khác về phương pháp của sáng chế là dòng dữ liệu hoặc chuỗi tín hiệu thể hiện chương trình máy tính để thực hiện một trong các phương pháp được mô tả ở đây. Dòng dữ liệu hoặc chuỗi tín hiệu có thể, ví dụ được tạo cấu hình để được truyền qua kết nối truyền thông dữ liệu, ví dụ thông qua Internet.

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

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



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

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

## YÊU CẦU BẢO HỘ

## 1. Thiết bị giải mã tín hiệu được mã hoá, bao gồm:

bộ giải mã âm thanh (1102) để giải mã phép biểu diễn được mã hoá của tập hợp thứ nhất của các phần phổ thứ nhất để thu được tập hợp thứ nhất được giải mã của các phần phổ thứ nhất (1101);

bộ giải mã theo tham số (1104) để giải mã phép biểu diễn theo tham số được mã hoá của tập hợp thứ hai của các phần phổ thứ hai để thu được phép biểu diễn được giải mã của phép biểu diễn theo tham số (1103), trong đó phép biểu diễn được giải mã của phép biểu diễn theo tham số (1103) bao gồm, đối với mỗi lát tần số đích, việc nhận biết vùng nguồn (1121) làm thông tin so khớp; và

bộ tái tạo tần số (1106) để tái tạo lát tần số đích sử dụng vùng nguồn từ tập hợp thứ nhất của các phần phổ thứ nhất (1101) được nhận biết bởi thông tin so khớp,

trong đó bộ tái tạo tần số (1106) bao gồm bộ lọc làm trắng có thể điều khiển (1122, 1123),

trong đó phép biểu diễn được giải mã của phép biểu diễn theo tham số (1103) bao gồm thông tin làm trắng (1123),

trong đó bộ tái tạo tần số (1106) được tạo cấu hình để áp dụng bộ lọc làm trắng (1122) cho vùng nguồn được chọn lọc theo thông tin so khớp trước khi thực hiện việc điều chỉnh đường bao phổ (1128), khi thông tin làm trắng cho vùng nguồn biểu thị rằng vùng nguồn sẽ được làm trắng, và

trong đó việc áp dụng bộ lọc làm trắng được thực hiện sao cho thông tin đường bao thô được loại bỏ và cấu trúc mịn phổ được làm nổi bật hoặc trong đó việc áp dụng bộ lọc làm trắng bao gồm việc tính toán ước tính đường bao phổ của vùng nguồn và việc chia phổ của vùng nguồn bởi đường bao phổ được biểu thị bởi ước lượng đường bao phổ.

## 2. Thiết bị theo điểm 1, trong đó bộ giải mã âm thanh (1102) là bộ giải mã âm thanh miền phổ, và

trong đó, thiết bị còn bao gồm bộ chuyển đổi phổ-thời gian (1108) để chuyển đổi phép biểu diễn phổ của các phân phổ thứ nhất và các phân phổ thứ hai được khôi phục thành phép biểu diễn thời gian.

3. Thiết bị theo điểm 1, trong đó thông tin làm trắng bao gồm lát hoặc nhóm các lát, thông tin mức làm trắng biểu thị mức làm trắng được áp dụng cho lát tần số nguồn khi tái tạo lát tần số đích, và

trong đó bộ tái tạo tần số (1106) được tạo cấu hình để áp dụng bộ lọc làm trắng được chọn lọc từ nhóm các bộ lọc làm trắng khác nhau đáp ứng thông tin làm trắng.

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

trong đó bộ tái tạo tần số (1106) bao gồm bộ điều chỉnh vùng nguồn,

trong đó phép biểu diễn được giải mã của phép biểu diễn theo tham số (1103) bao gồm, ngoài việc nhận biết vùng nguồn (1121), thông tin ký hiệu, và

trong đó bộ điều chỉnh vùng nguồn được tạo cấu hình để áp dụng hoạt động để thu được chuyển dịch pha của các giá trị phổ vùng nguồn phù hợp với thông tin tín hiệu.

5. Thiết bị theo một trong số các điểm nêu trên,

trong đó bộ tái tạo tần số (1106) bao gồm bộ điều biến ô (1120),

trong đó phép biểu diễn được giải mã của phép biểu diễn theo tham số (1103) bao gồm độ trễ tương quan ngoài việc nhận biết vùng nguồn (1121), và

trong đó bộ điều biến ô được tạo cấu hình để áp dụng việc điều biến ô phù hợp với độ trễ tương quan được kết hợp với nhận biết vùng nguồn (1121).

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

trong đó bộ tái tạo tần số (1106) bao gồm bộ điều biến ô (1120),

trong đó phép biểu diễn được giải mã của phép biểu diễn theo tham số (1103) bao gồm độ trễ tương quan ngoài việc nhận biết vùng nguồn (1121), và

trong đó bộ điều biến ô (1120) được tạo cấu hình để áp dụng việc điều biến ô sử dụng chuỗi theo thời gian luân phiên bằng  $-1/1$  khi độ trễ tương quan là số lẻ.

7. Thiết bị mã hoá tín hiệu âm thanh, bao gồm:

bộ chuyển đổi thời gian-phổ (1130) để chuyển đổi tín hiệu âm thanh thành phép biểu diễn phổ;

bộ phân tích phổ để phân tích phép biểu diễn phổ để xác định tập hợp thứ nhất của các phần phổ thứ nhất được mã hoá với độ phân giải phổ thứ nhất, và tập hợp thứ hai của các phần phổ thứ hai được mã hoá với độ phân giải phổ thứ hai, trong đó độ phân giải phổ thứ hai thấp hơn độ phân giải phổ thứ nhất;

bộ tính toán tham số (1134) để tính toán các điểm giống nhau giữa các vùng nguồn và các vùng đích được xác định trước, vùng nguồn bao gồm các phần phổ và vùng đích bao gồm các phần phổ thứ hai,

trong đó bộ tính toán tham số (1134) được tạo cấu hình để so sánh các kết quả phù hợp cho các cặp khác nhau của phần phổ thứ nhất và phần phổ thứ hai để xác định cặp phù hợp được chọn lọc và để cung cấp thông tin so khớp nhận biết cặp phù hợp; và

bộ mã hóa lỗi (1136) để mã hoá tập hợp thứ nhất của các phần phổ thứ nhất, trong đó tập hợp thứ nhất của các phần phổ thứ nhất bao gồm các vùng nguồn được xác định trước và các phần phổ khác với các vùng nguồn được xác định trước,

trong đó bộ tính toán tham số được tạo cấu hình để sao chép lát nguồn nhiều lần thành lát đích cho tới khi lát đích được điền đầy với nhiều bản sao chép của lát nguồn, khi lát đích lớn hơn lát nguồn, hoặc trong đó bộ tính toán tham số (1134) được tạo cấu hình để làm trắng theo phổ (1142) phần phổ thứ nhất hoặc thứ hai của các cặp trước khi thực hiện xử lý tương quan để thu được nhận biết phù hợp, trong đó việc làm trắng theo phổ được thực hiện sao cho thông tin đường bao thô được loại bỏ và cấu trúc mịn phổ được làm nổi bật hoặc trong đó việc làm trắng theo phổ (1142) lần lượt bao gồm tính toán ước lượng đường bao phổ của phần phổ thứ nhất hoặc thứ hai và chia phổ của phần phổ thứ nhất hoặc thứ hai, bởi đường bao phổ được biểu thị bằng ước lượng đường bao phổ.

8. Thiết bị theo điểm 7, trong đó bộ tính toán tham số (1134) được tạo cấu hình để sử dụng các vùng đích được xác định trước trong tập hợp thứ hai của các phần phổ thứ

hai hoặc các vùng nguồn được xác định trước trong tập hợp thứ nhất của các phần phổ thứ nhất.

9. Thiết bị theo điểm 7, trong đó bộ tính toán tham số (1134) được tạo cấu hình sao cho các vùng đích được xác định trước là không chồng lấp, hoặc các vùng nguồn được xác định trước là chồng lấp, hoặc trong đó các vùng nguồn được xác định trước là tập hợp con của tập hợp thứ nhất của các phần phổ thứ nhất thấp hơn tần số bắt đầu điền đầy khoảng trống (309), hoặc trong đó vùng đích được xác định trước bao phủ vùng phổ thấp nhất trùng với tần số bắt đầu điền đầy khoảng trống (309).

10. Thiết bị theo một trong số các điểm 7, 8 hoặc 9,

trong đó bộ tính toán tham số (1134) được tạo cấu hình để so sánh các cặp của vùng đích và vùng nguồn với cặp của vùng đích và vẫn vùng nguồn ấy, trong đó vẫn vùng nguồn ấy được chuyển dịch bởi dấu hiệu tương quan để cung cấp thông tin về dấu hiệu tương quan của cặp được lựa chọn làm thông tin so khớp bổ sung.

11. Thiết bị theo điểm bất kỳ trong số các điểm từ 7 đến 10,

trong đó, bộ tính toán tham số (1134) được tạo cấu hình để thực hiện việc xử lý tương quan để thu được kết quả phù hợp cho cặp phần phổ thứ nhất và phần phổ thứ hai, kết quả phù hợp có ký hiệu âm, và

trong đó, bộ tính toán tham số (1134) được tạo cấu hình để cung cấp thông tin cho ký hiệu âm làm thông tin so khớp bổ sung.

12. Thiết bị theo điểm bất kỳ trong số các điểm từ 7 đến 11,

trong đó bộ tính toán tham số (1134) được tạo cấu hình để xác định số nguyên của các lát đích và để xác định nhiều lát nguồn được định kích cỡ bằng nhau cho mỗi lát đích.

13. Thiết bị theo một trong số các điểm từ 7 đến 12,

trong đó bộ tính toán tham số (1134) được tạo cấu hình để tính toán đường bao phổ của phần phổ thứ nhất hoặc phần phổ thứ hai sử dụng ít nhất một trong số các quy trình sau đây:

biến đổi phổ với phép biến đổi cosin rời rạc (discrete cosine transform - DCT), giữ lại các hệ số DCT tần số thấp bằng cách đặt các hệ số DCT cao hơn thành không và tính toán DCT nghịch đảo;

tính toán đường bao phổ của tập hợp các hệ số dự báo tuyến tính được tính toán trên khung âm thanh miền thời gian; hoặc

lọc phổ (lũy thừa) biến đổi cosin rời rạc cải biên với bộ lọc thông thấp.

14. Thiết bị theo một trong số các điểm từ 7 đến 13 còn bao gồm hoạt động chia nhỏ lát nguồn (1146) và bộ nhớ (1148) để lưu trữ thông tin lát nguồn của khung sớm hơn có trước khung hiện thời.

15. Thiết bị theo một trong số các điểm từ 7 đến 14,

trong đó hoạt động chia nhỏ nguồn (1146) bao gồm việc phân tích nhiều lát nguồn đối với tính tương tự của chúng và loại bỏ lát nguồn có tính tương tự với lát nguồn khác lớn hơn ngưỡng được xác định trước từ tập hợp các ô tiềm năng được sử dụng để tính toán tương quan chéo (1140).

16. Thiết bị theo một trong số các điểm từ 11 đến 15,

trong đó bộ tính toán tham số được tạo cấu hình để giữ lại tập hợp thông tin so khớp cho mỗi vùng đích từ khung trước, khi không có vùng nguồn nào trong khung hiện thời tương quan với các vùng đích của khung hiện thời tốt hơn đối với khung trước so với ngưỡng được xác định trước.

17. Phương pháp giải mã tín hiệu được mã hoá bao gồm các bước:

giải mã (1102) phép biểu diễn được mã hoá của tập hợp thứ nhất của các phần phổ thứ nhất để thu được tập hợp thứ nhất được giải mã của các phần phổ thứ nhất (1101);

giải mã (1104) phép biểu diễn theo tham số được mã hoá của tập hợp thứ hai của các phần phổ thứ hai để thu được phép biểu diễn được giải mã của phép biểu diễn theo tham số (1103), trong đó phép biểu diễn được giải mã của phép biểu diễn theo tham số (1103) bao gồm, đối với mỗi lát tần số đích, việc nhận biết vùng nguồn làm thông tin so khớp; và

tái tạo (1106) lát tần số đích sử dụng vùng nguồn từ tập hợp thứ nhất của các phần phổ thứ nhất (1101) được nhận biết bởi thông tin so khớp,

trong đó việc tái tạo (1106) bao gồm việc sử dụng bộ lọc làm trắng có thể điều khiển (1122, 1123),

trong đó phép biểu diễn được giải mã của phép biểu diễn theo tham số (1103) bao gồm thông tin làm trắng (1123),

trong đó việc tái tạo (1106) bao gồm việc áp dụng bộ lọc làm trắng (1122) cho vùng nguồn được chọn lọc theo thông tin so khớp trước khi thực hiện việc điều chỉnh đường bao phổ (1128), khi thông tin làm trắng cho vùng nguồn biểu thị rằng vùng nguồn được làm trắng, và

trong đó việc áp dụng bộ lọc làm trắng được thực hiện sao cho thông tin đường bao thô được loại bỏ và cấu trúc mịn phổ được làm nổi bật hoặc trong đó việc áp dụng bộ lọc làm trắng bao gồm việc tính toán ước lượng đường bao phổ và việc chia phổ của vùng nguồn bởi đường bao phổ được biểu thị bởi ước lượng đường bao phổ.

18. Phương pháp mã hoá tín hiệu âm thanh bao gồm các bước:

chuyển đổi (1130) tín hiệu âm thanh thành phép biểu diễn phổ;

phân tích (1132) phép biểu diễn phổ để xác định tập hợp thứ nhất của các phần phổ thứ nhất được mã hoá với độ phân giải phổ thứ nhất, và tập hợp thứ hai của các phần phổ thứ hai được mã hoá với độ phân giải phổ thứ hai, trong đó độ phân giải phổ thứ hai thấp hơn độ phân giải phổ thứ nhất;

tính toán (1134) các điểm giống nhau giữa các vùng nguồn được xác định trước và các vùng đích, vùng nguồn bao gồm các phần phổ và vùng đích bao gồm các phần phổ thứ hai,

trong đó việc tính toán (1134) bao gồm so sánh các kết quả phù hợp cho các cặp khác nhau của phần phổ thứ nhất và phần phổ thứ hai để xác định cặp phù hợp được lựa chọn và để cung cấp thông tin so khớp nhận biết cặp phù hợp; và

mã hoá (1136) tập hợp thứ nhất của các phần phổ thứ nhất, trong đó tập hợp thứ nhất của các phần phổ thứ nhất bao gồm các vùng nguồn được xác định trước và các phần phổ khác với các vùng nguồn được xác định trước,

trong đó việc tính toán (1134) bao gồm việc sao chép lát nguồn nhiều lần thành lát đích cho tới khi lát đích được điền đầy với nhiều bản sao chép của lát nguồn, khi lát đích lớn hơn lát nguồn, hoặc trong đó việc tính toán (1134) bao gồm việc làm trắng theo phổ (1142) phần phổ thứ nhất hoặc thứ hai của các cặp trước khi thực hiện xử lý tương quan để thu được nhận biết phù hợp, trong đó việc làm trắng theo phổ được thực hiện sao cho thông tin đường bao thô được loại bỏ và phổ cấu trúc mịn được làm nổi bật hoặc trong đó việc làm trắng theo phổ (1142) lần lượt bao gồm việc tính toán ước lượng đường bao phổ của phần phổ thứ nhất hoặc thứ hai và việc chia phổ của phần phổ thứ nhất hoặc thứ hai, bởi đường bao phổ được biểu thị bởi ước lượng đường bao phổ.

19. Vật ghi đọc được bằng máy tính chứa chương trình máy tính để thực hiện phương pháp theo điểm 17, khi được chạy trên máy tính hoặc bộ xử lý.

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



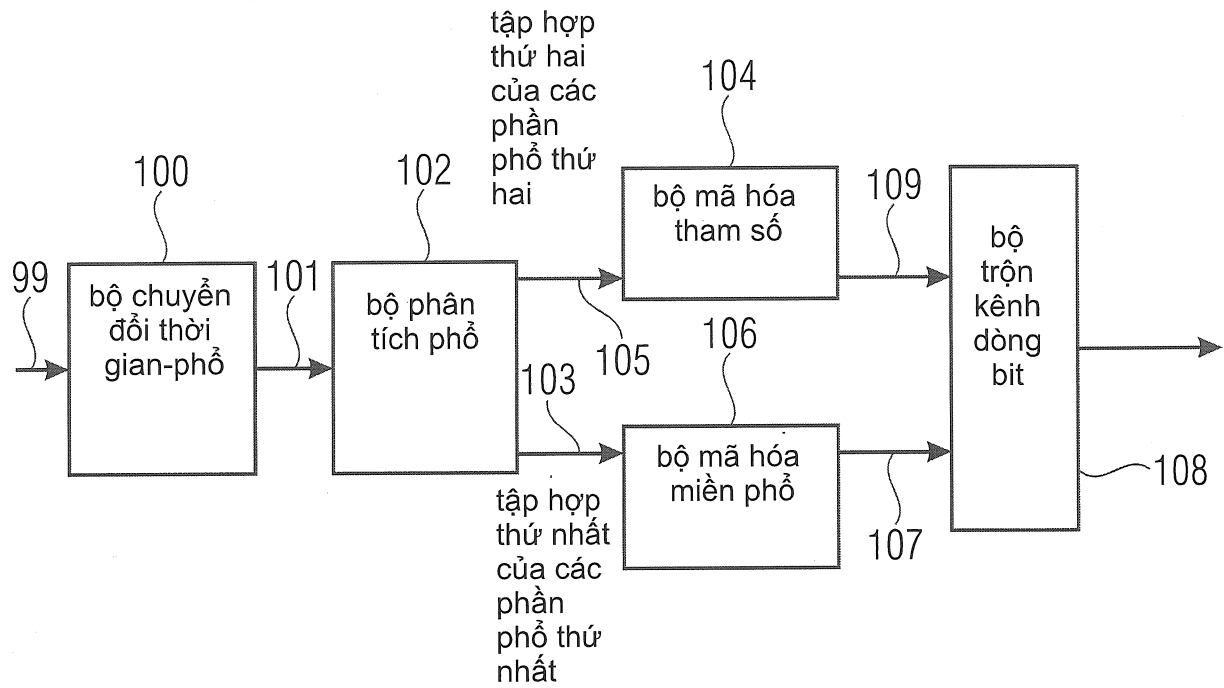


FIG 1A

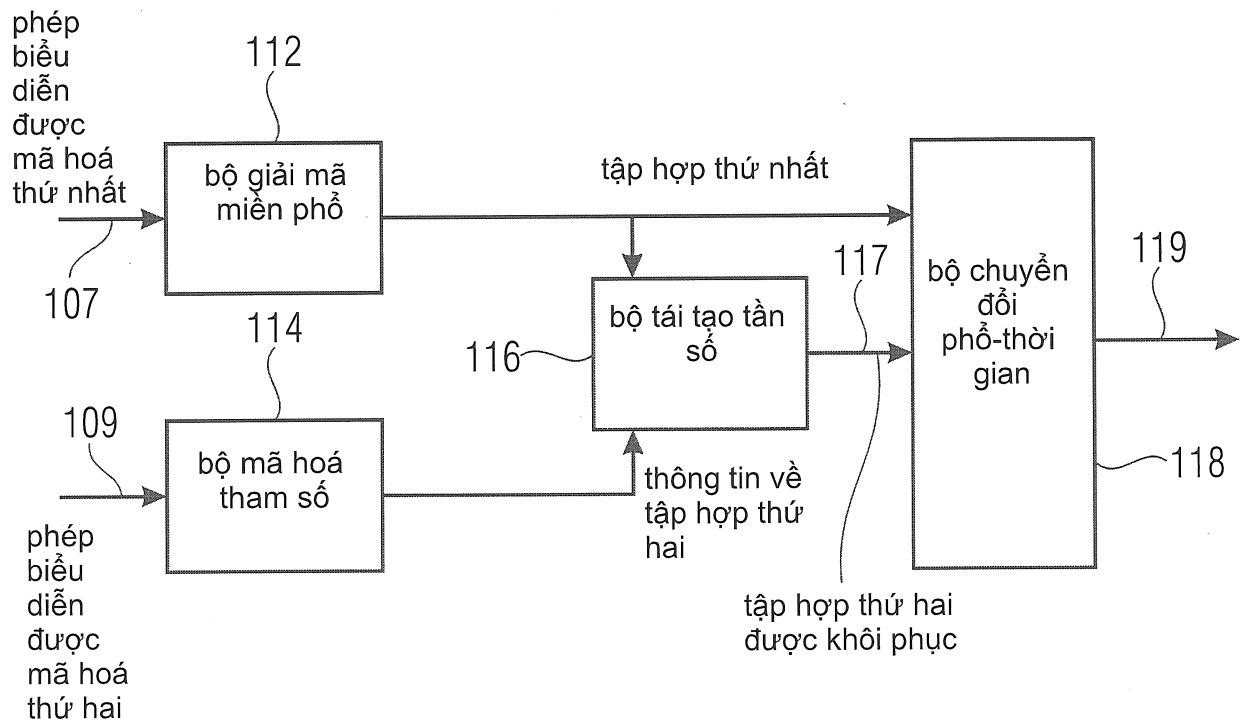


FIG 1B

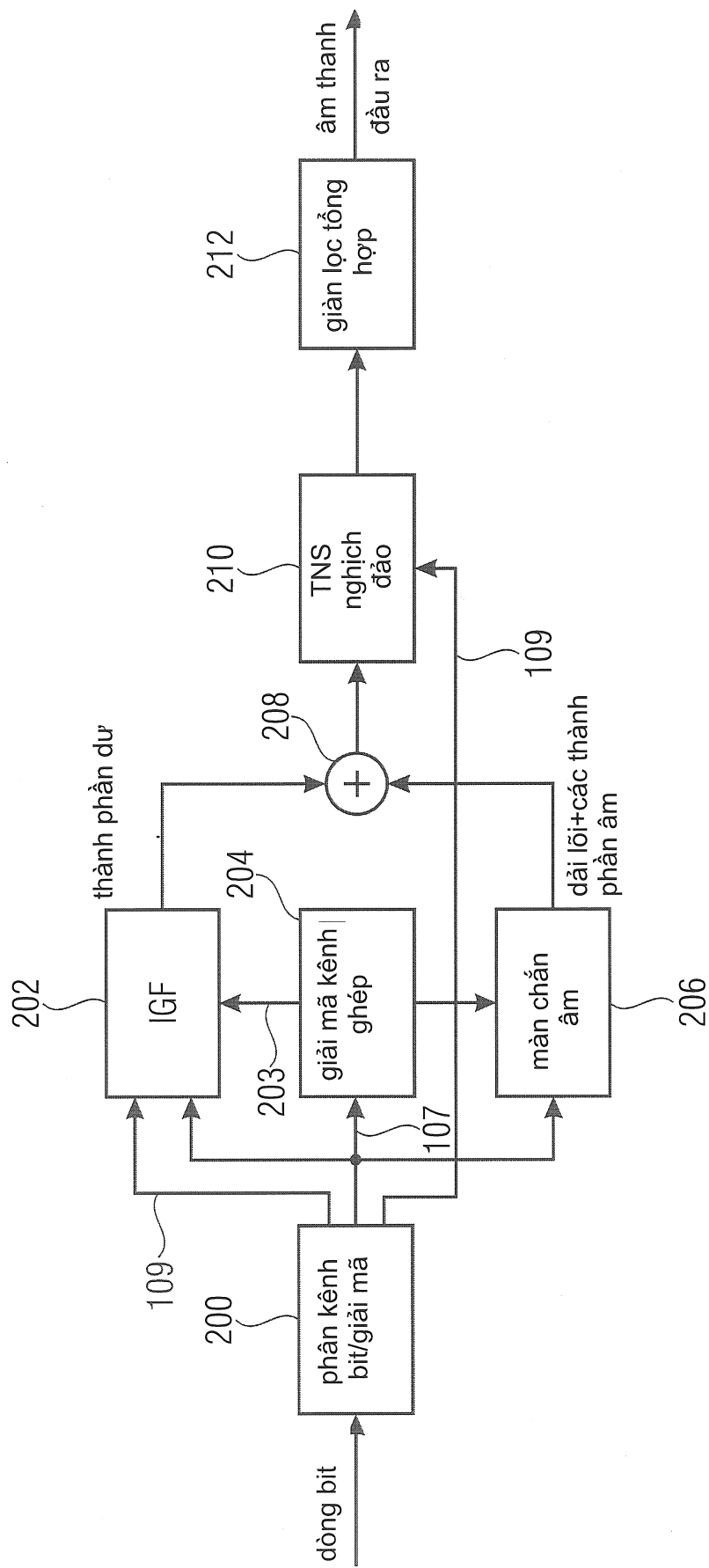


FIG 2A

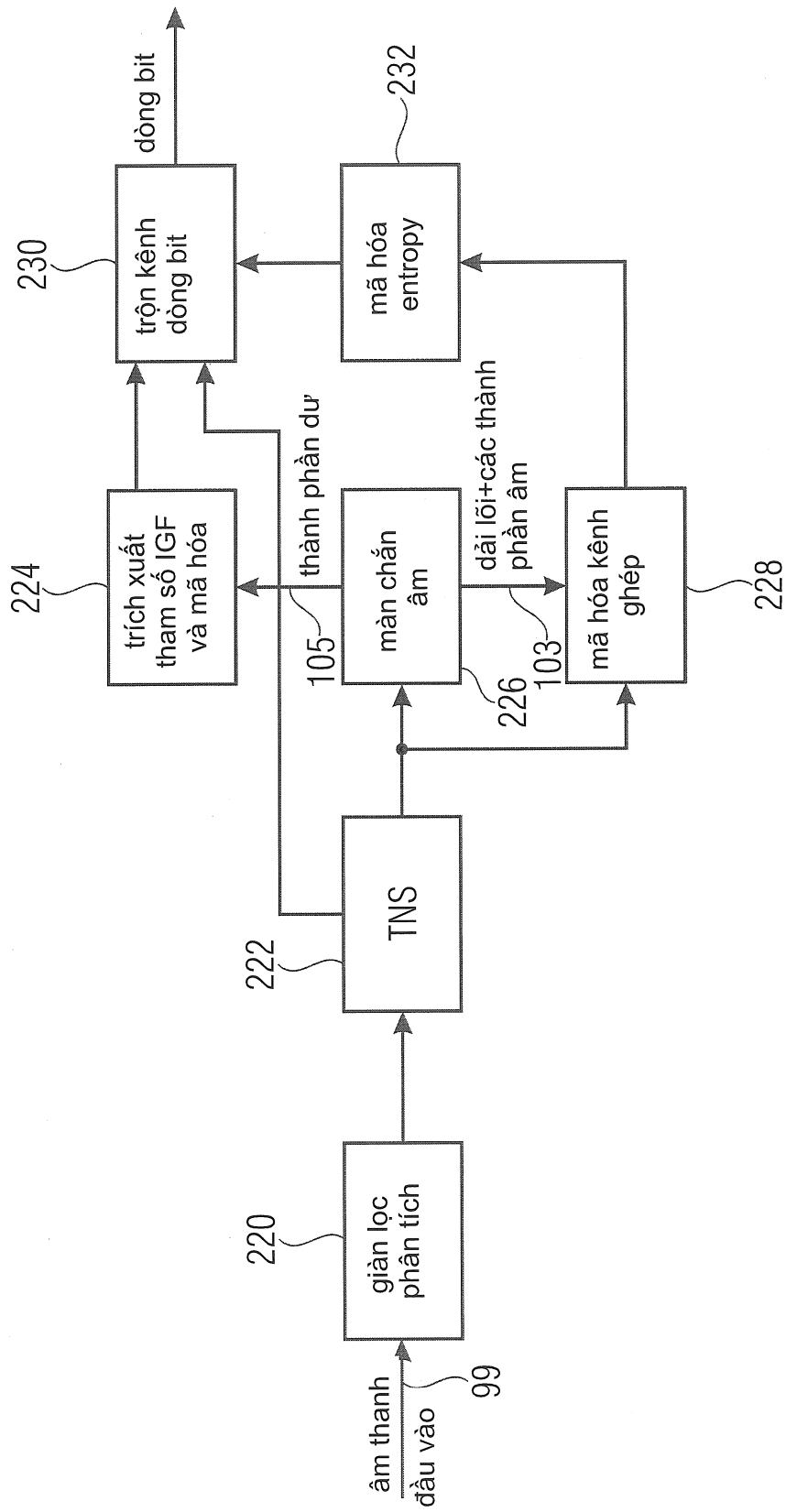


FIG 2B

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

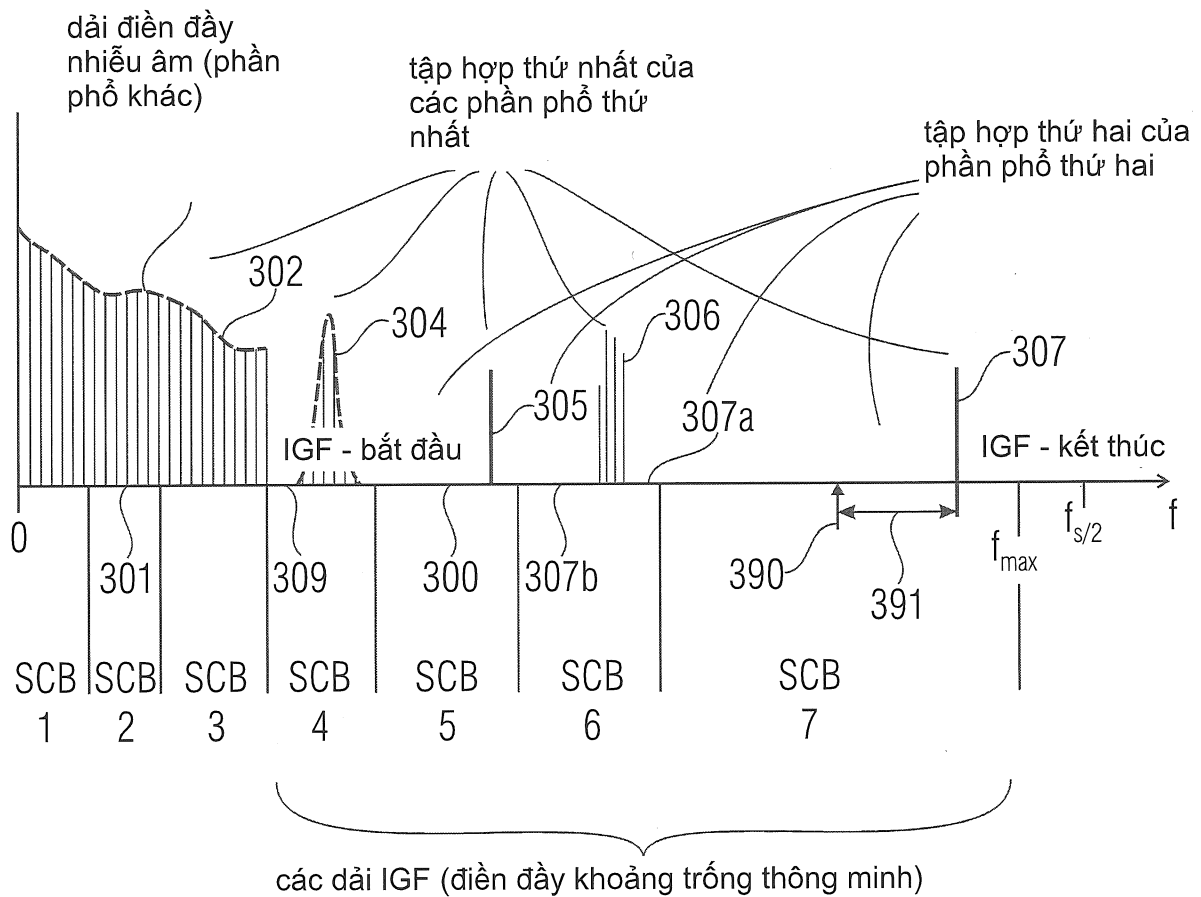


FIG 3A

|      |                 |      |                |                |                |                |
|------|-----------------|------|----------------|----------------|----------------|----------------|
| SCB1 | SCB2            | SCB3 | SCB4           | SCB5           | SCB6           | SCB7           |
| SF1  | SF2             | SF3  | SF4            | SF5            | SF6            | SF7            |
|      |                 |      | E <sub>1</sub> | E <sub>2</sub> | E <sub>3</sub> | E <sub>4</sub> |
|      | NF <sub>2</sub> |      |                |                |                |                |

308 310 312

FIG 3B

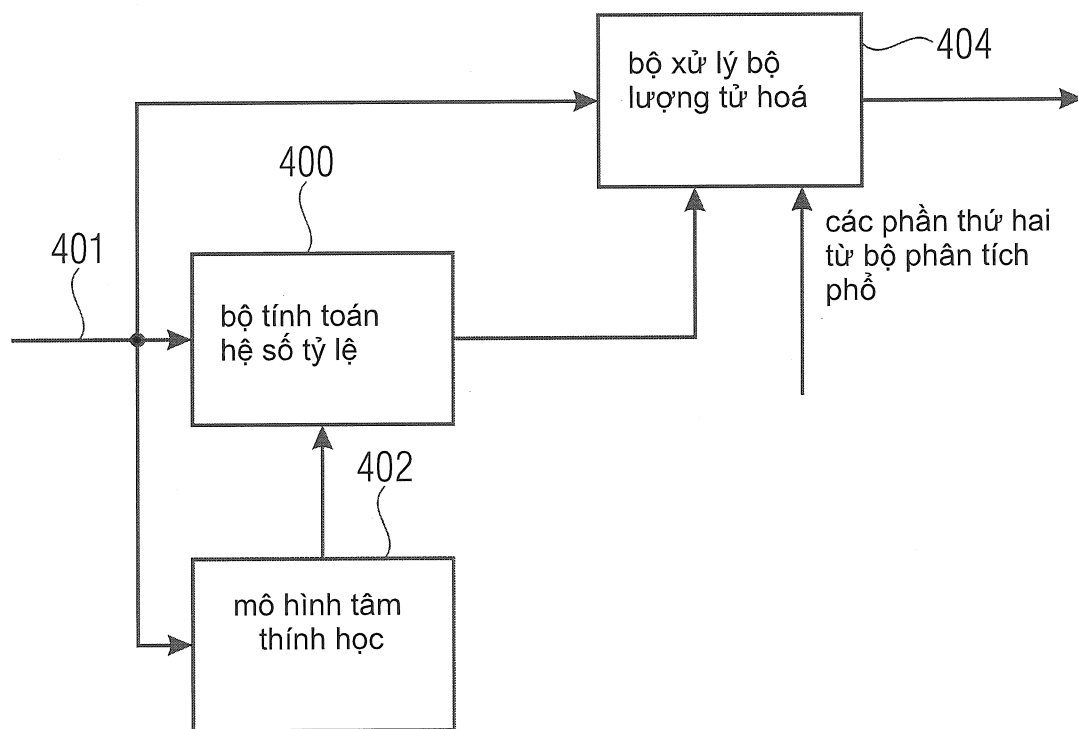


FIG 4A

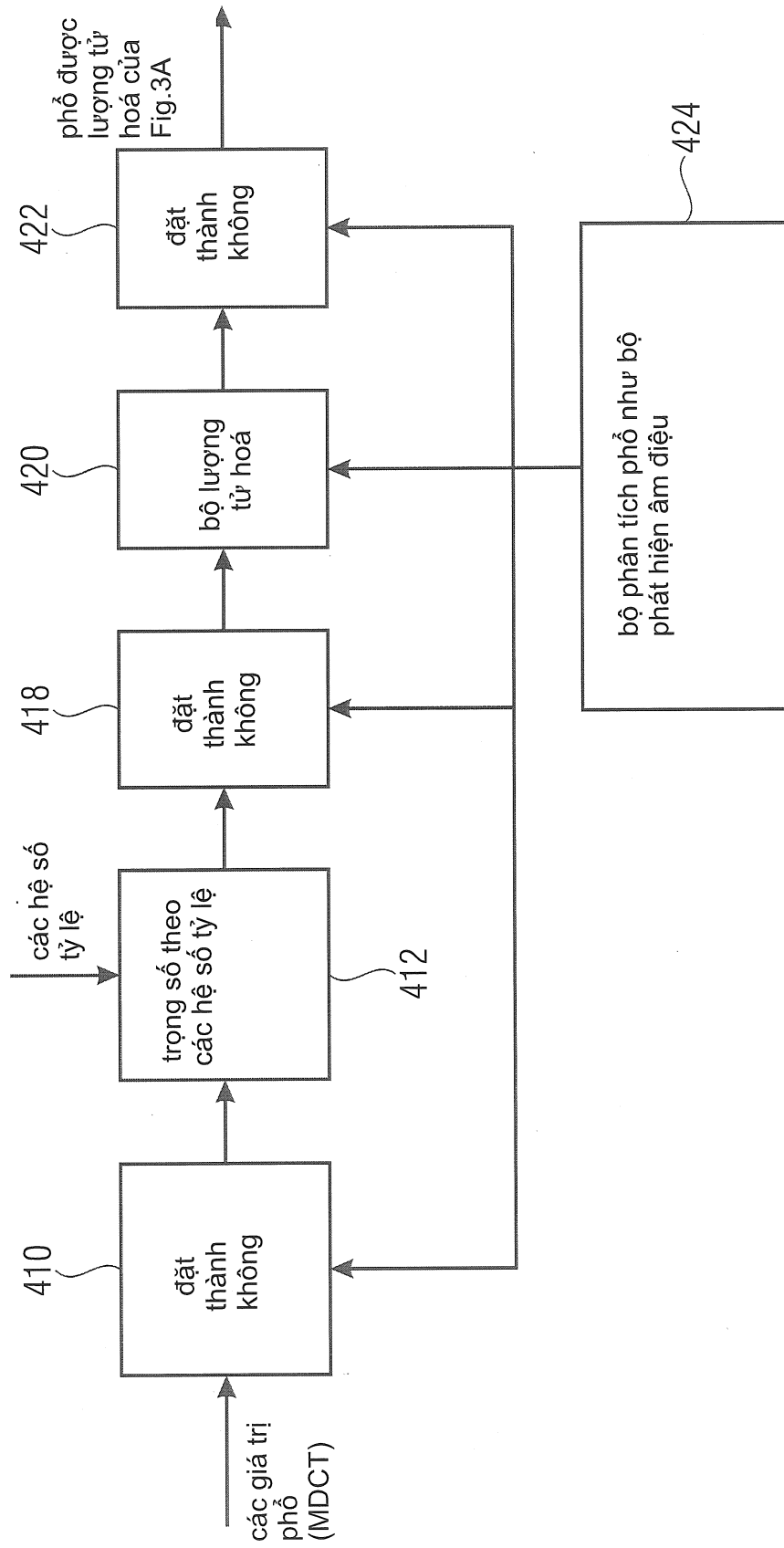


FIG 4B

(BỘ XỬ LÝ LƯỢNG TỬ HÓA)

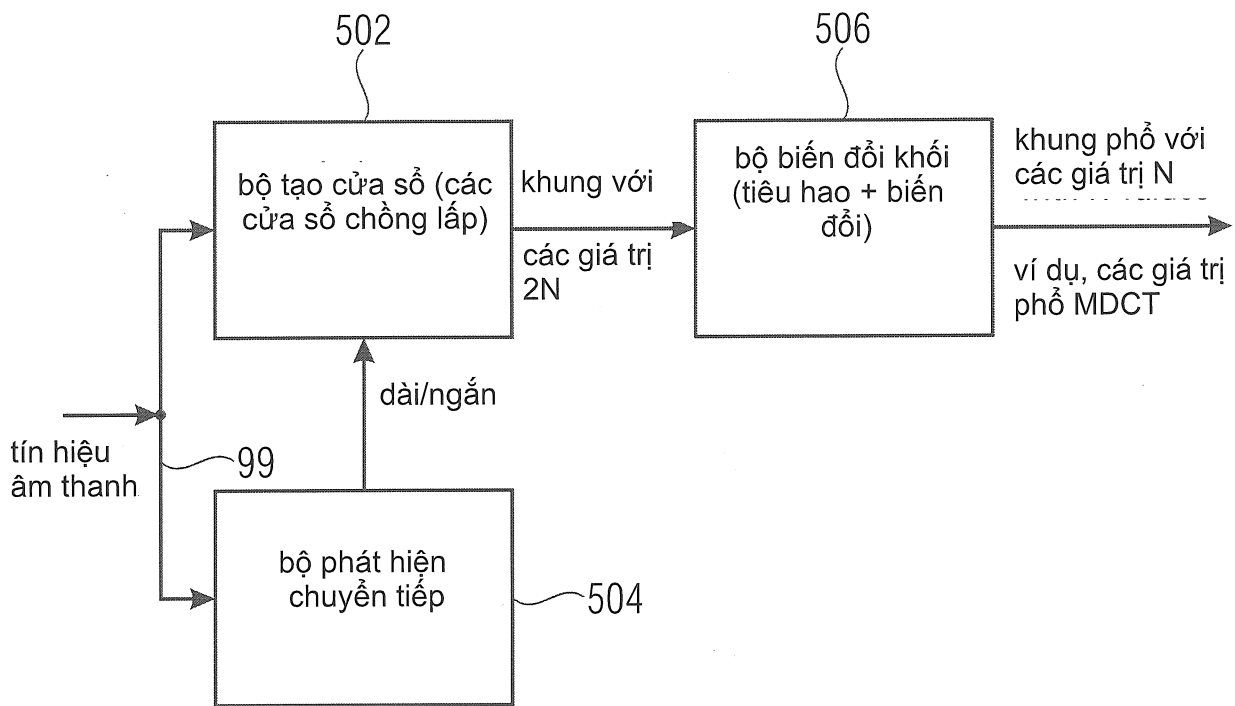


FIG 5A

(CÁC PHẦN PHỔ KHÁC)

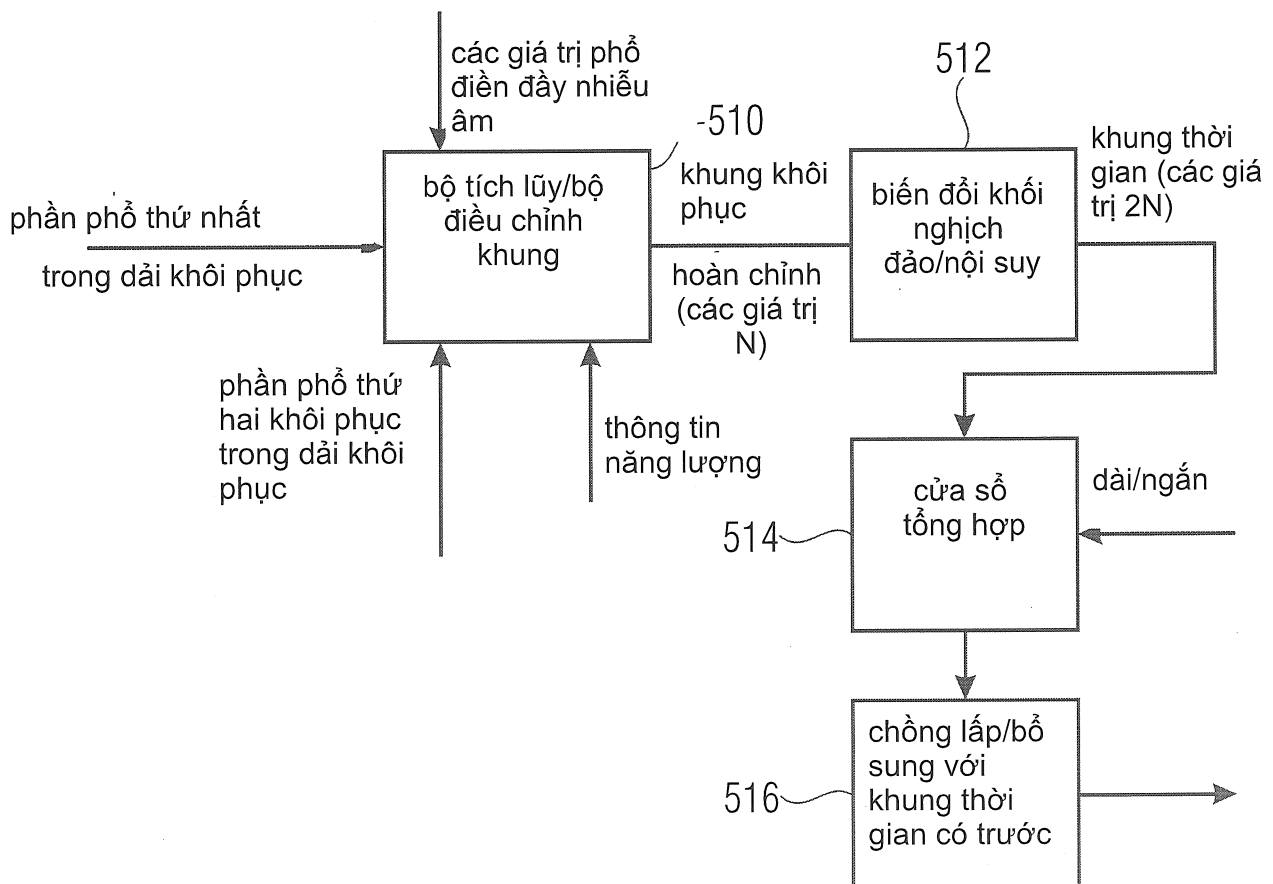


FIG 5B



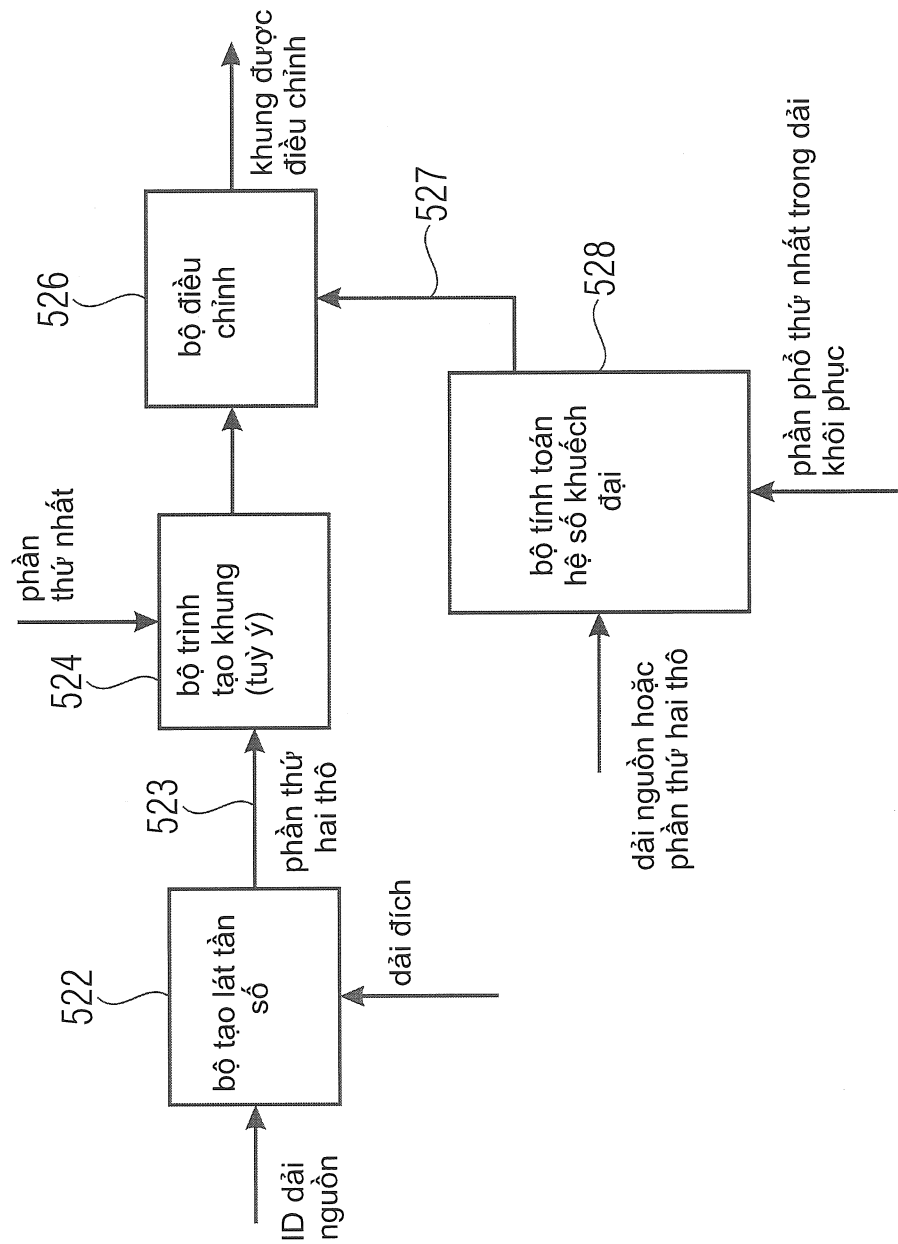


FIG 5C

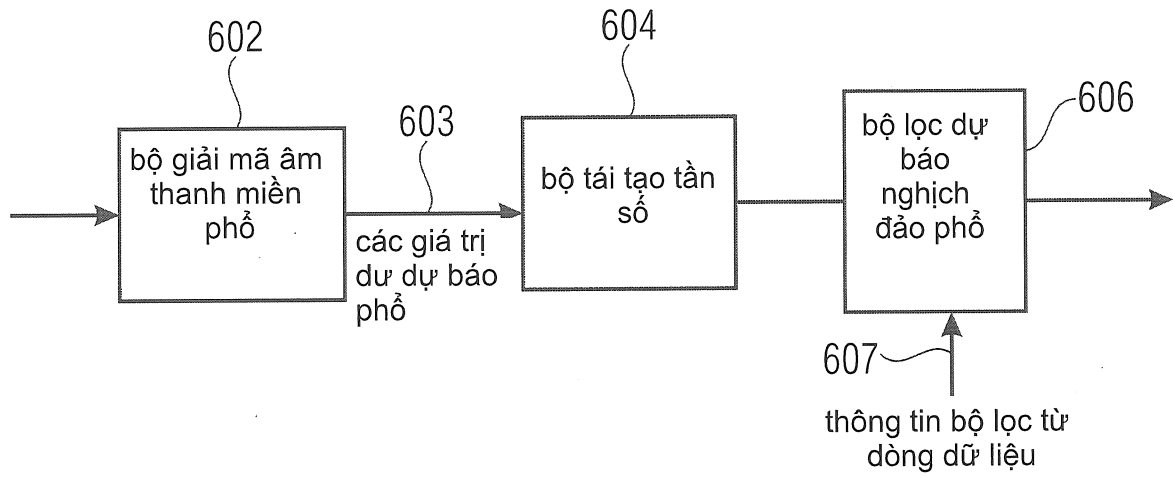


FIG 6A

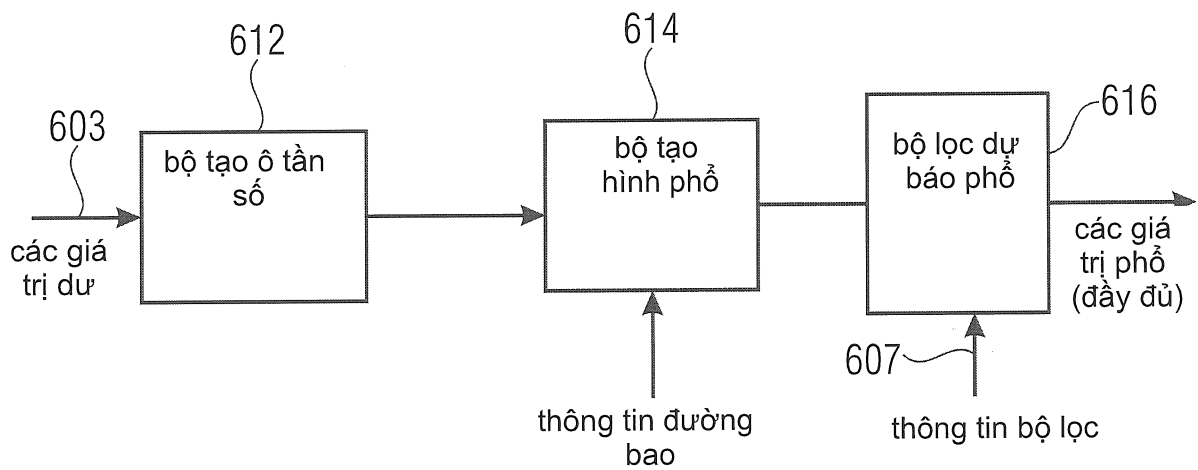


FIG 6B

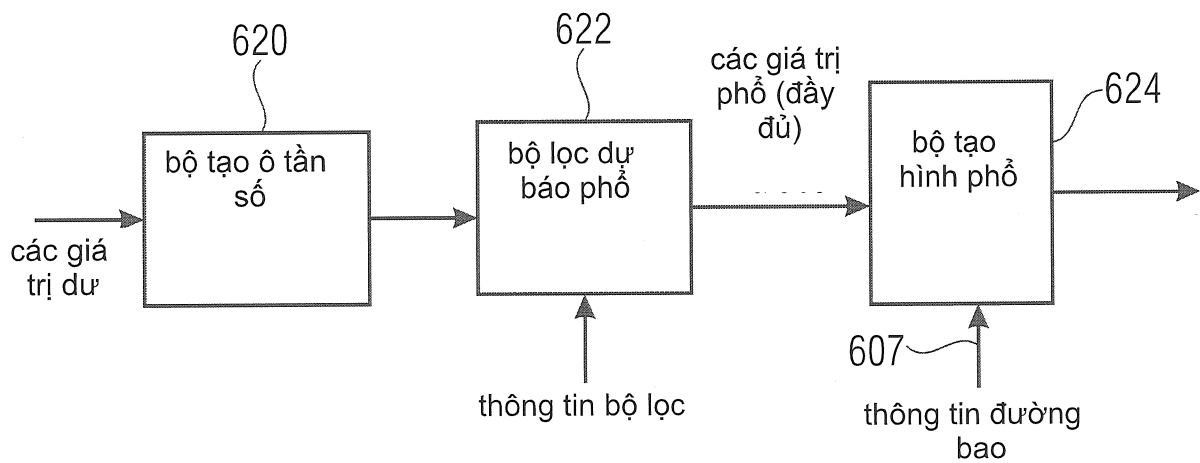


FIG 6C

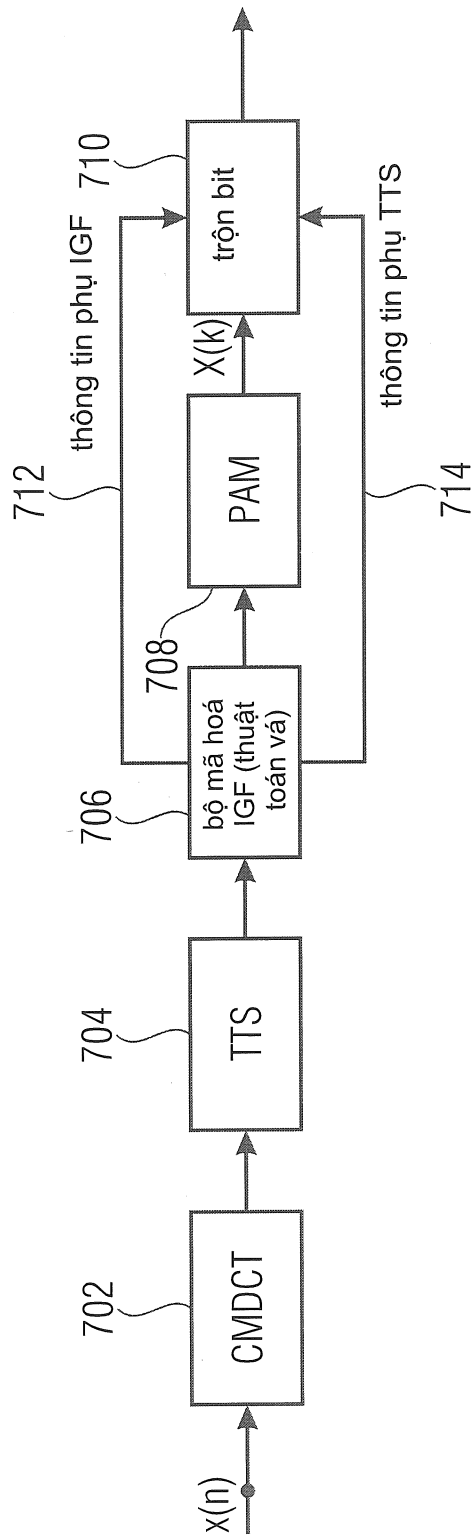


FIG 7A  
(BỘ MÃ HÓA)

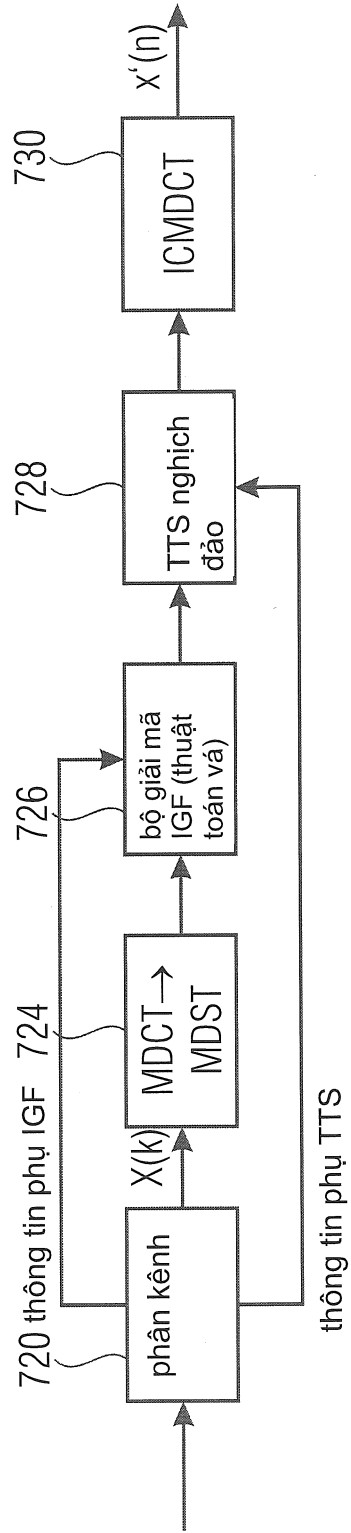


FIG 7B  
(BỘ GIẢI MÃ)

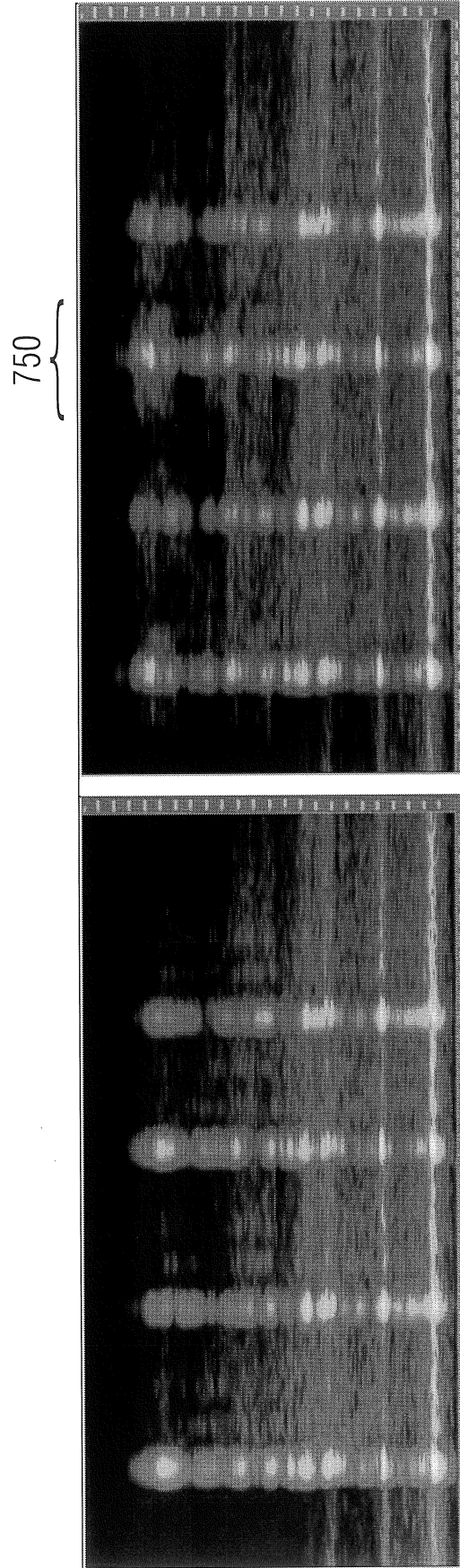
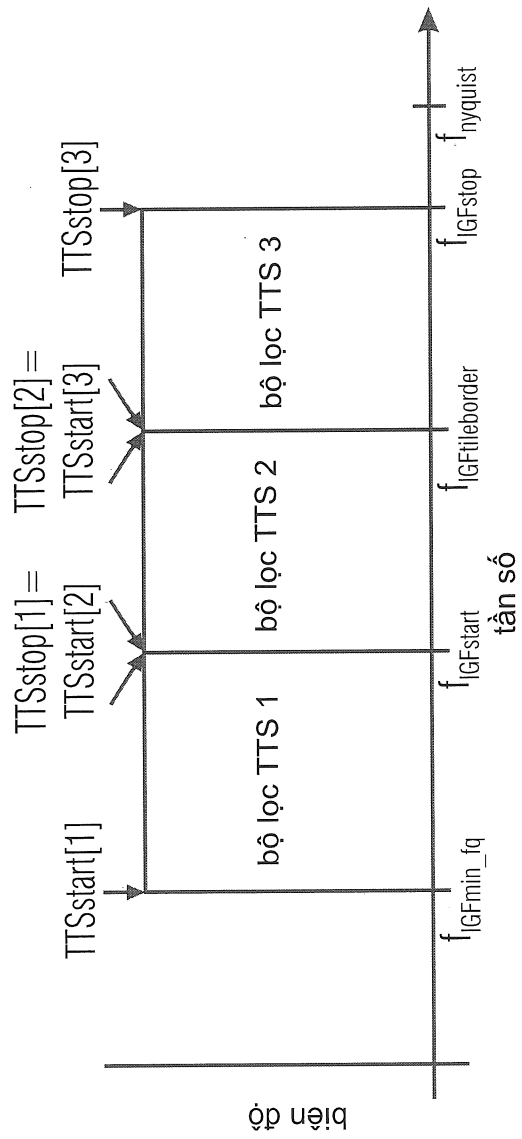
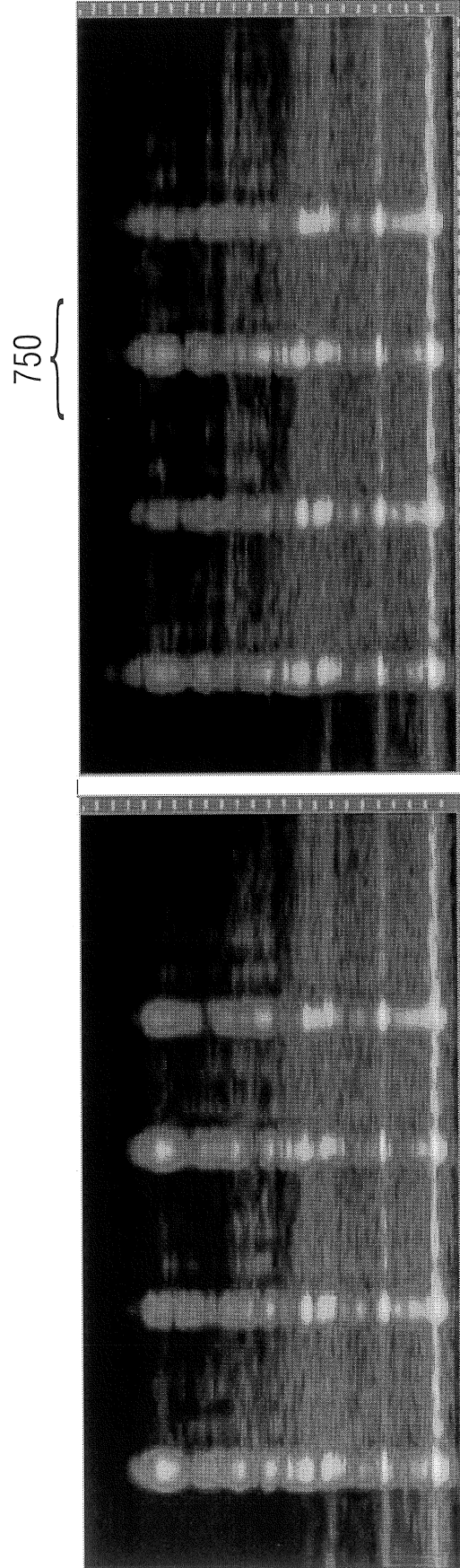


FIG 7C



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

FIG 7D



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

FIG 7E

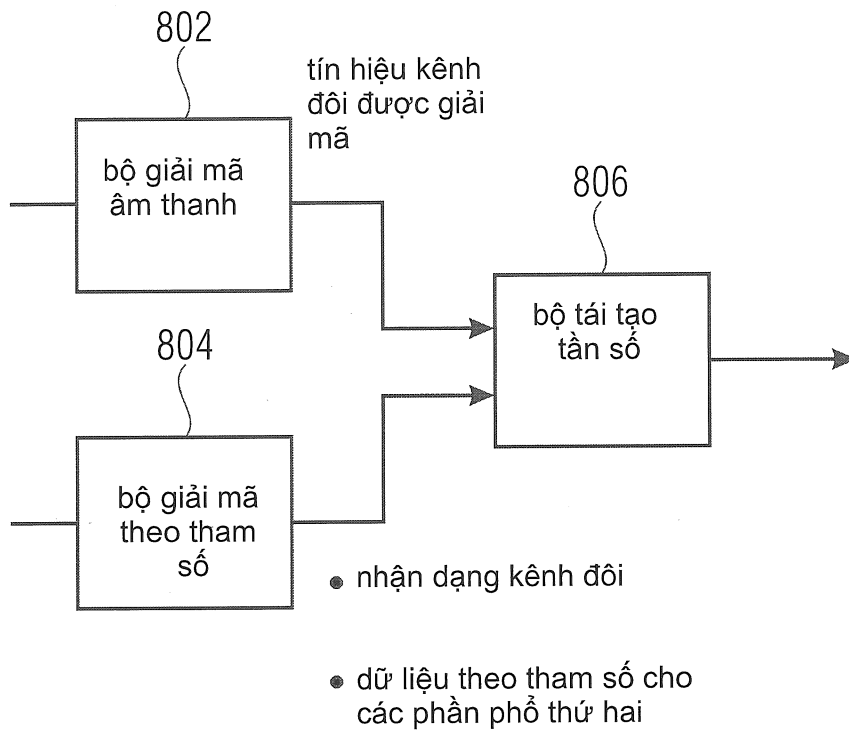


FIG 8A

|              |   |   |   |   |
|--------------|---|---|---|---|
| khoảng nguồn | 1 | 1 | 2 | 2 |
| khoảng đích  | 1 | 2 | 1 | 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)

FIG 8B

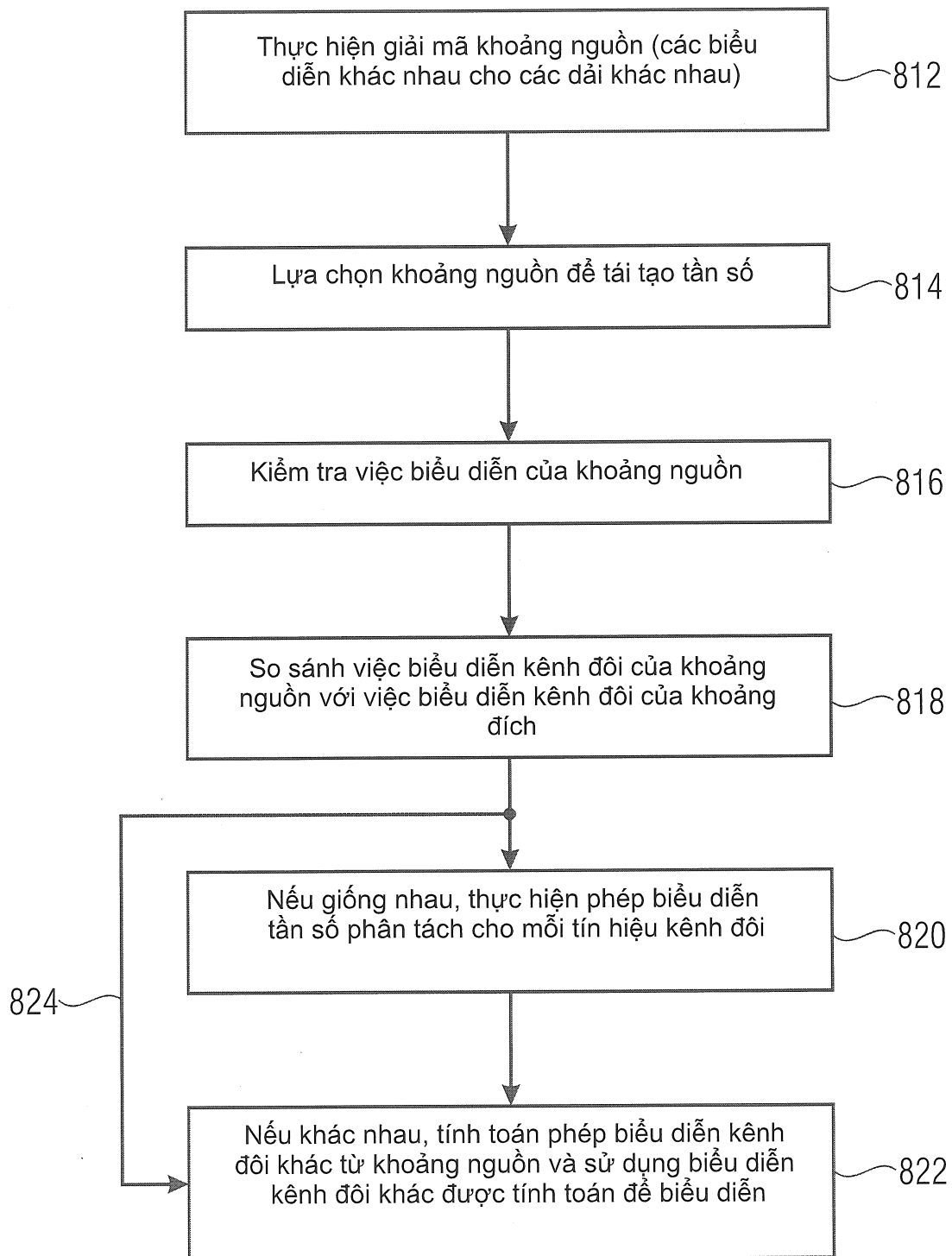


FIG 8C



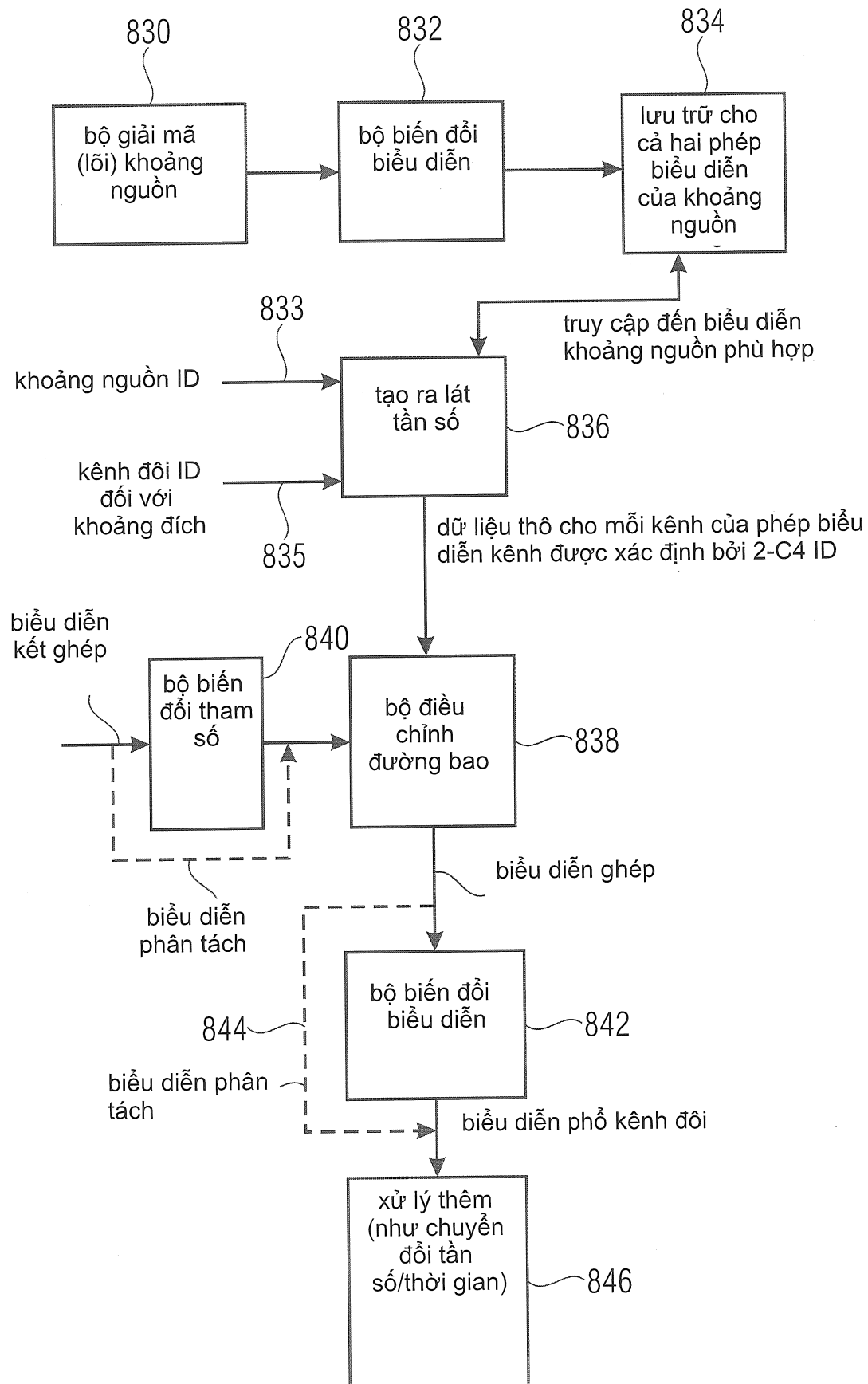


FIG 8D

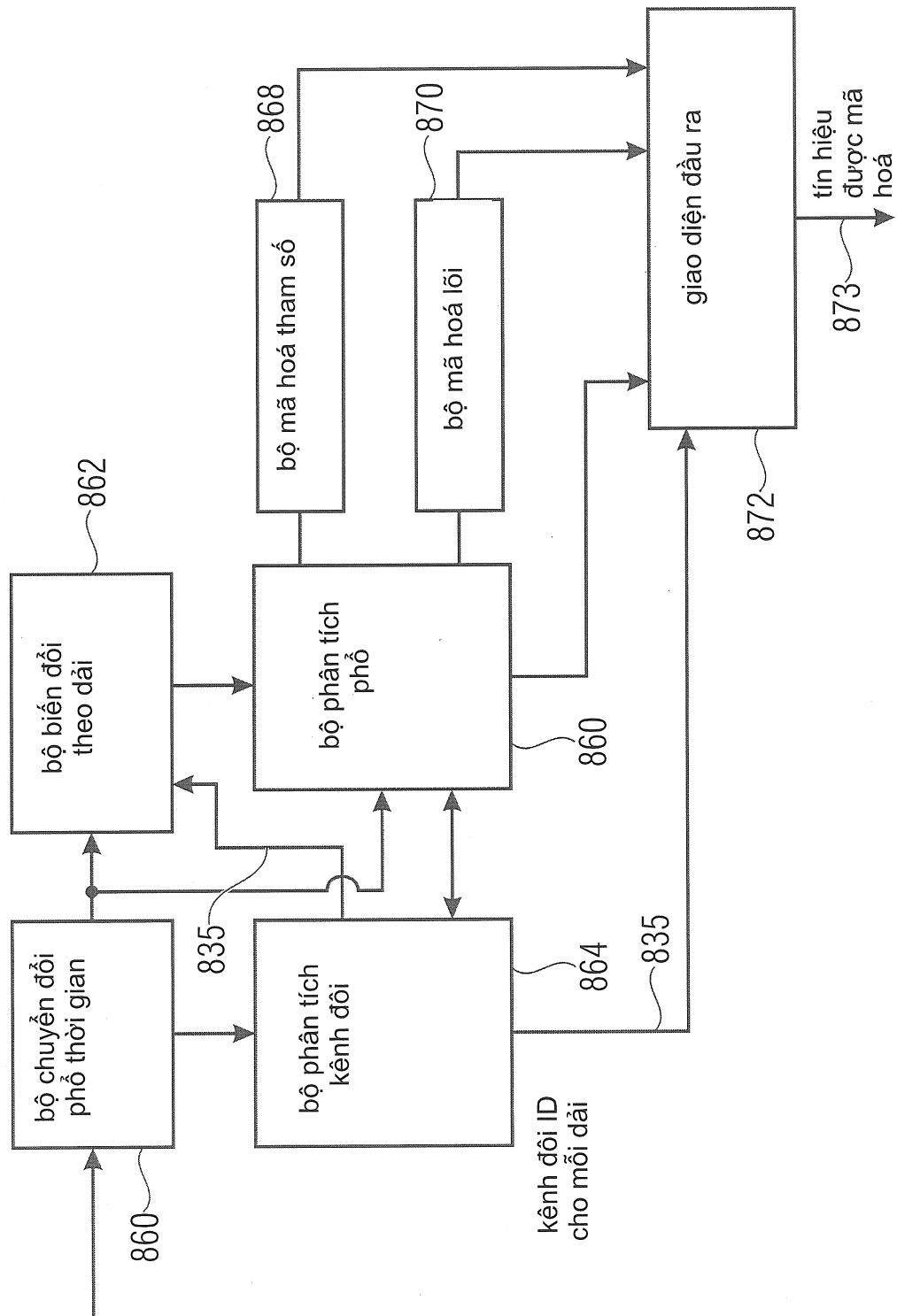


FIG 8E

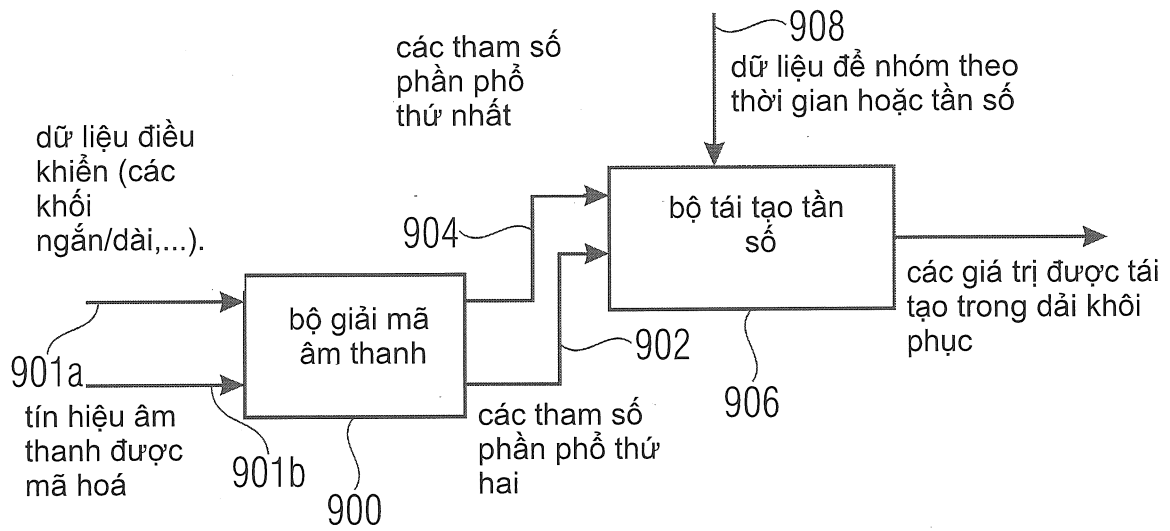


FIG 9A

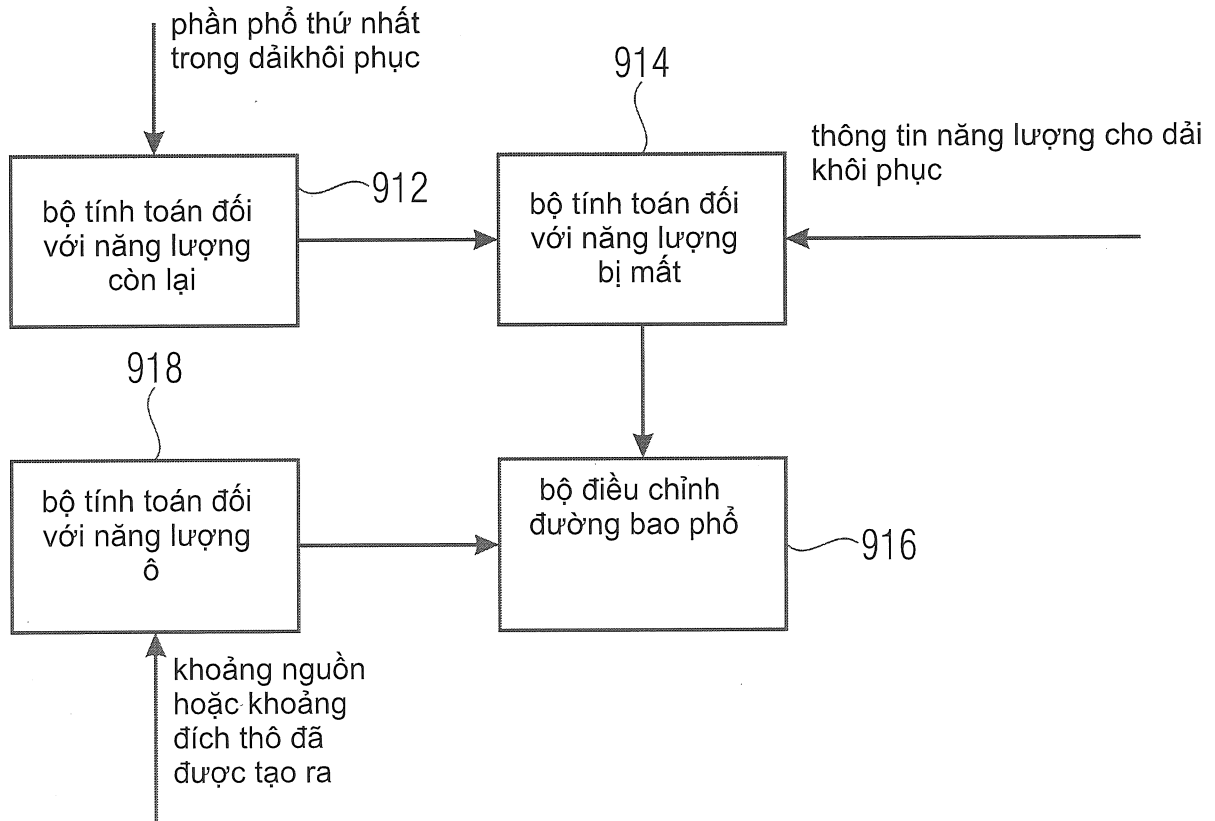
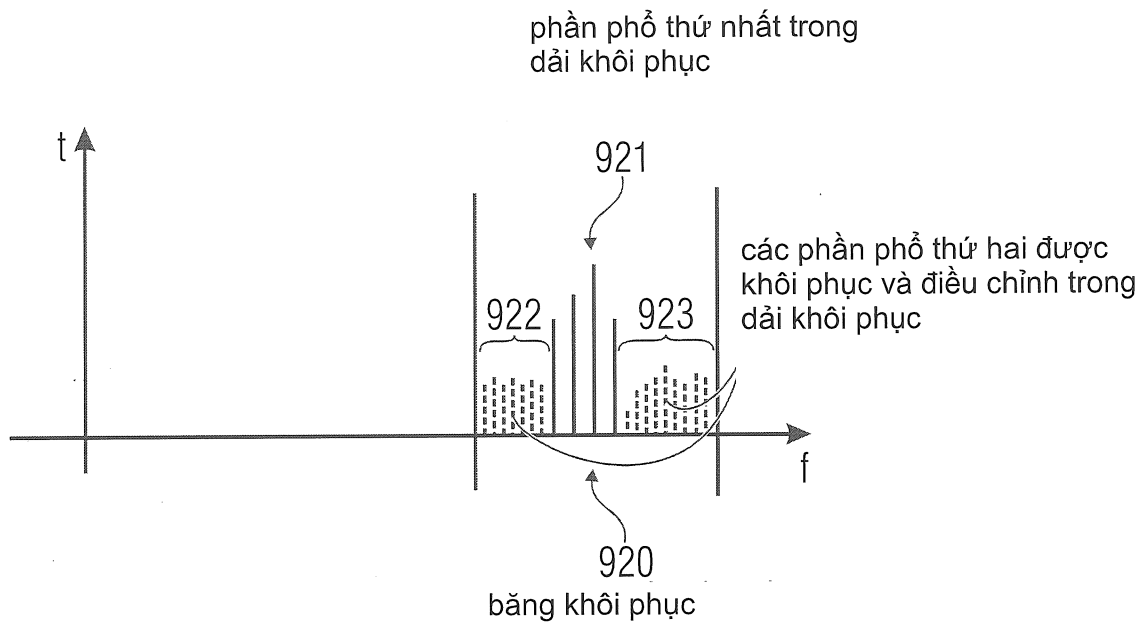


FIG 9B



ví dụ

- năng lượng còn lại:

5 đơn vị

- giá trị năng lượng đối với dải khôi phục

10 đơn vị (bao phủ các phần phổ thứ nhất và thứ hai trong dải khôi phục)

- năng lượng của dữ liệu khoảng nguồn hoặc dữ liệu khoảng đích thô

8 đơn vị

- năng lượng bị mất:

5 đơn vị

- hệ số khuếch đại:

$$g := \sqrt{\frac{mE_k}{pE_k}} = 0,79$$

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

→ phần phổ thứ nhất không bị tác động bởi điều chỉnh đường bao

FIG 9C

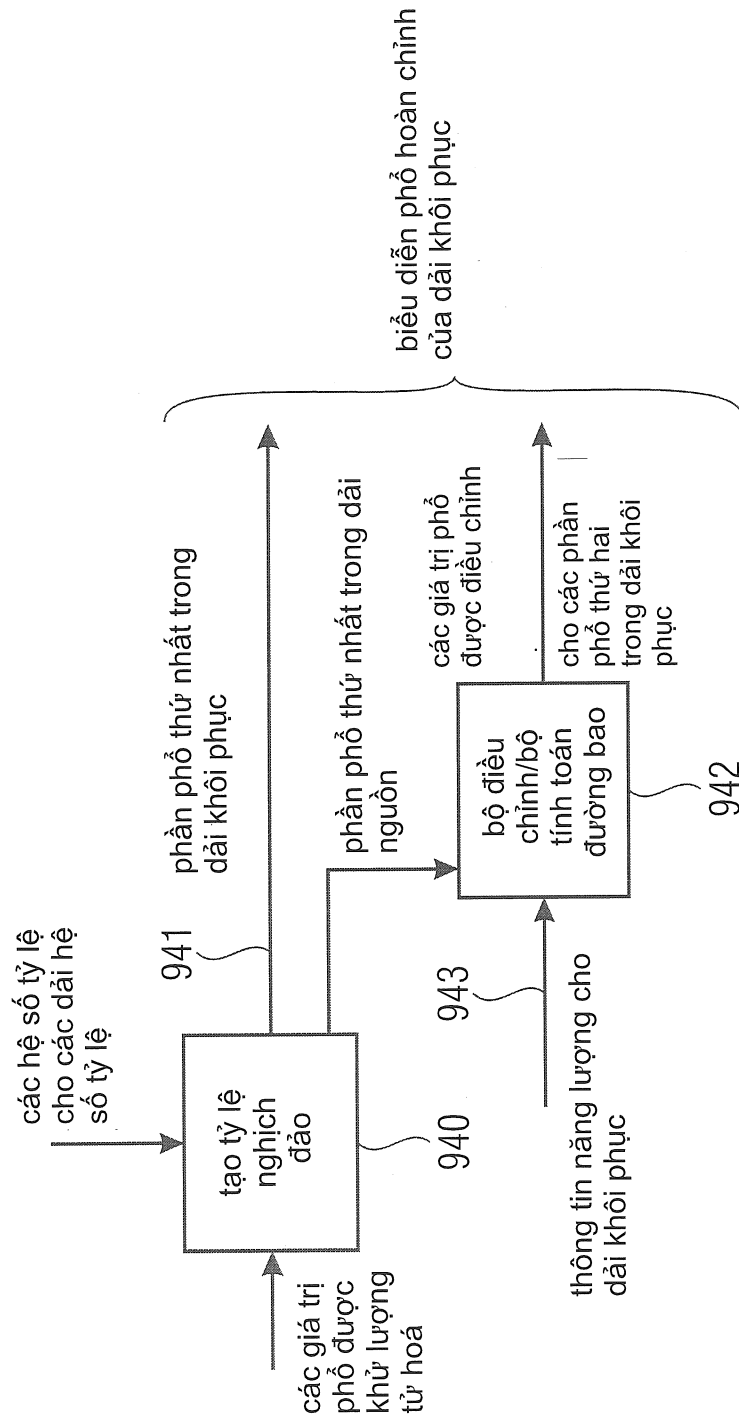


FIG 9D

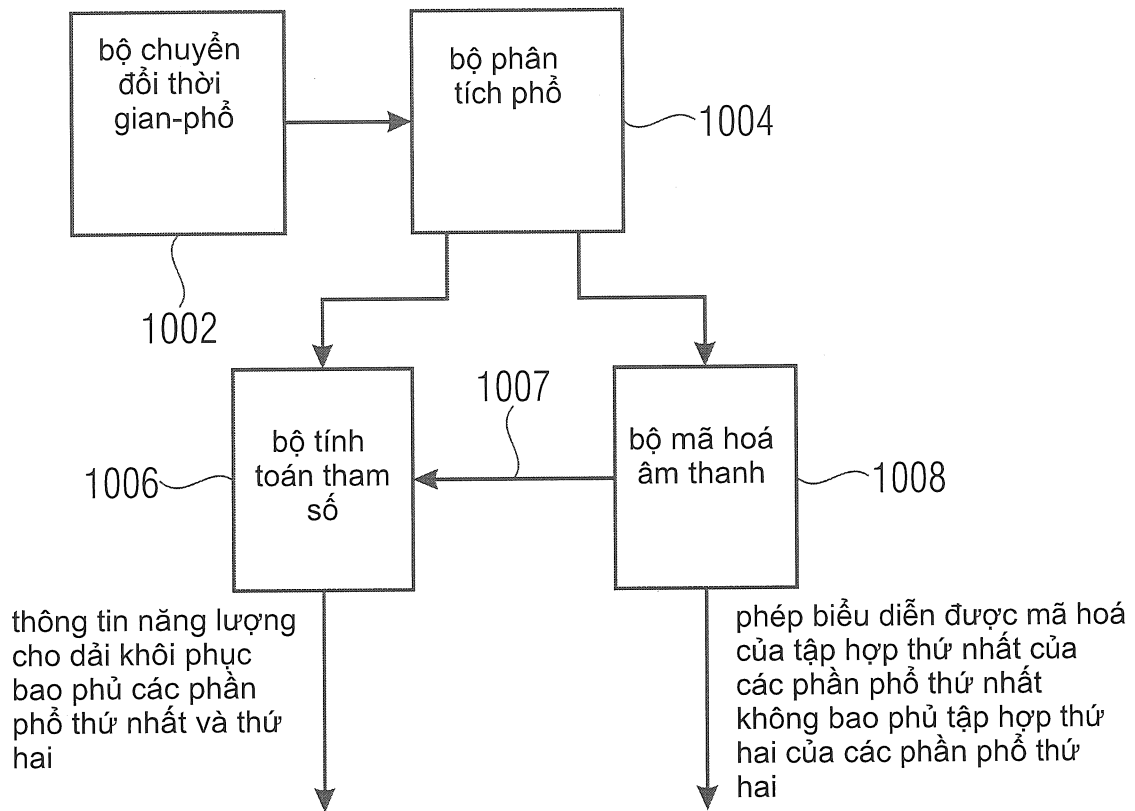


FIG 10A

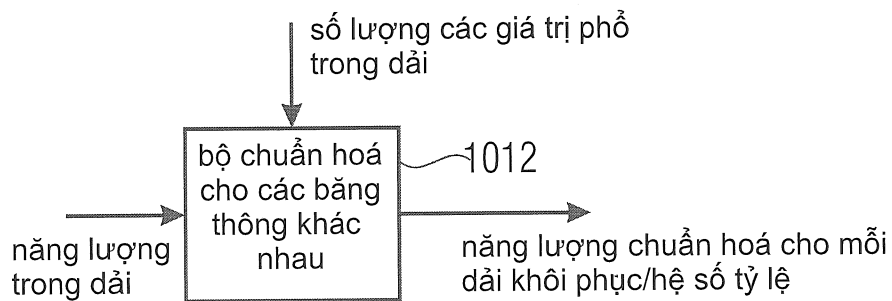


FIG 10B

thông tin điều khiển từ bộ mã hoá lỗi

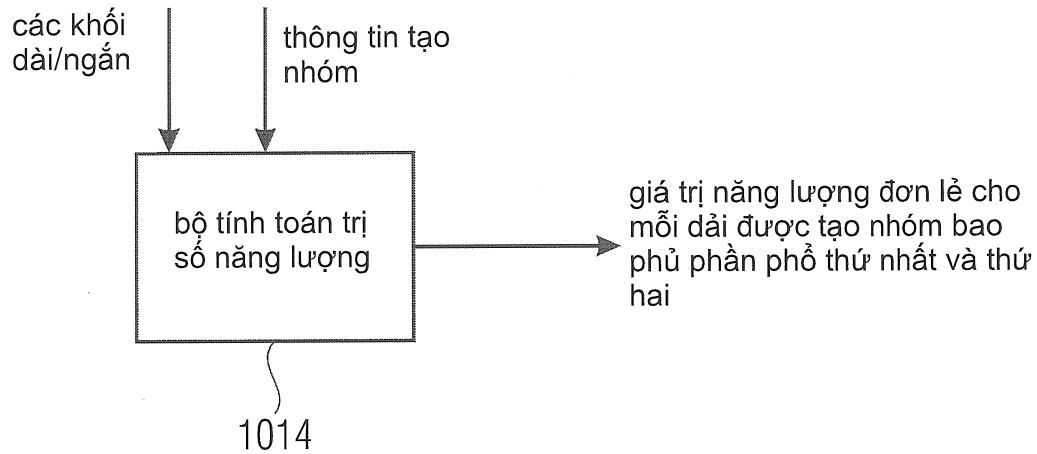


FIG 10C

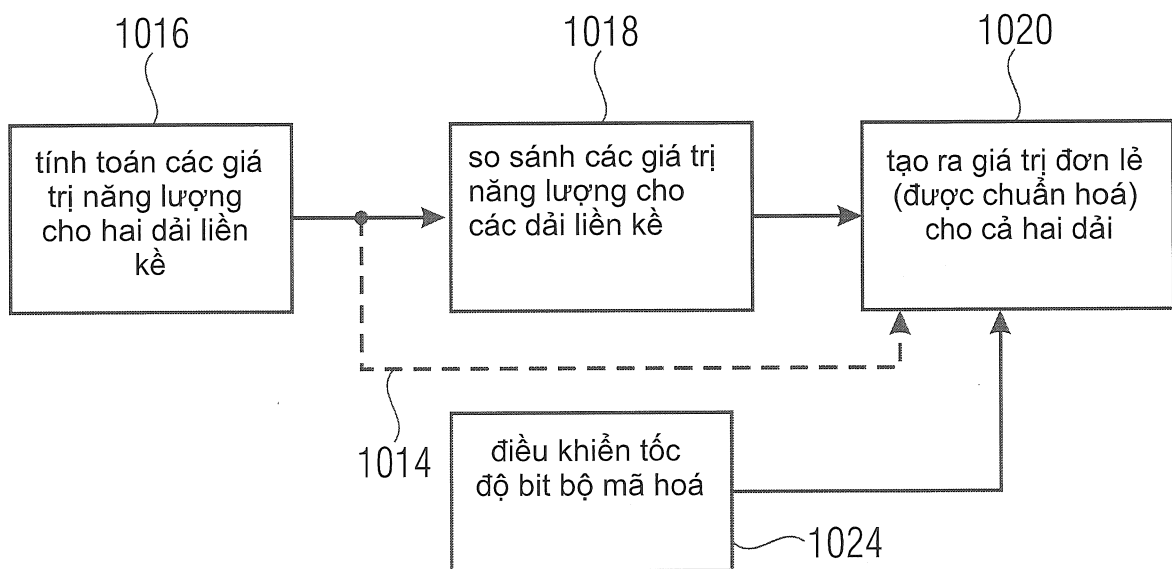


FIG 10D

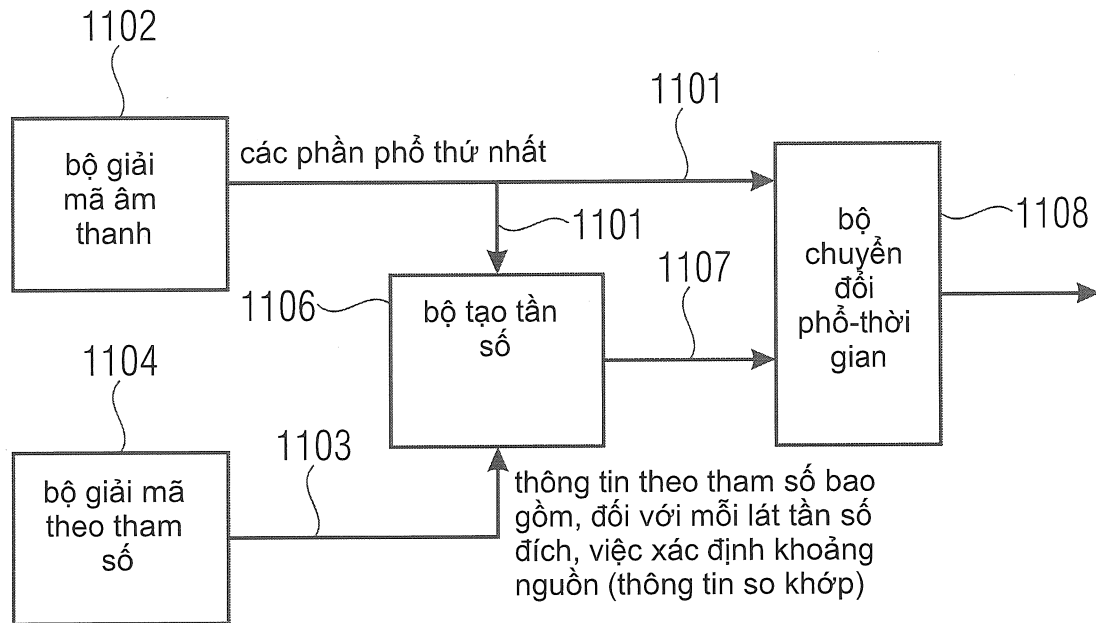


FIG 11A

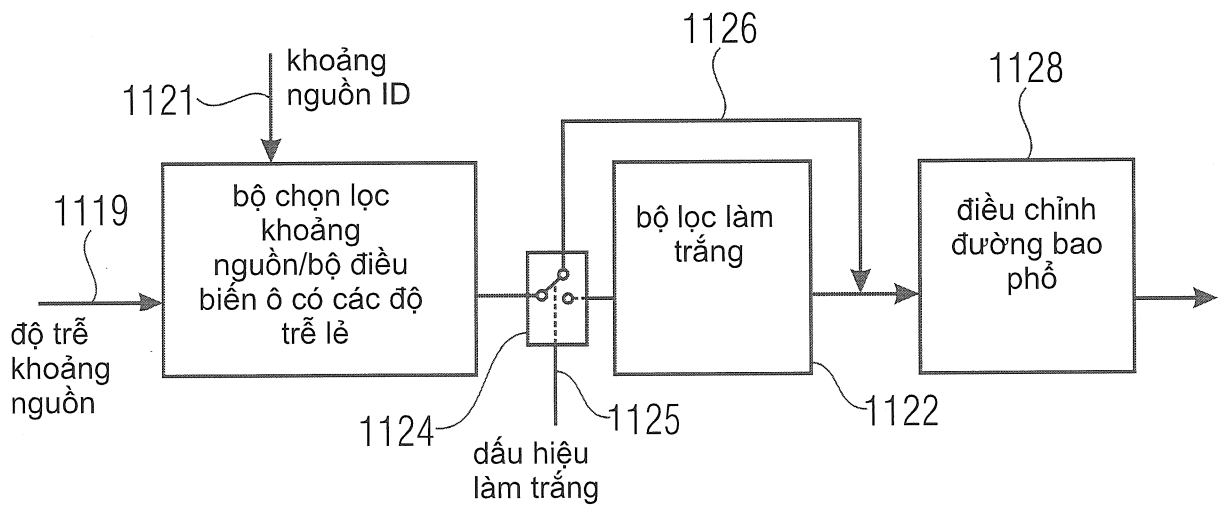


FIG 11B



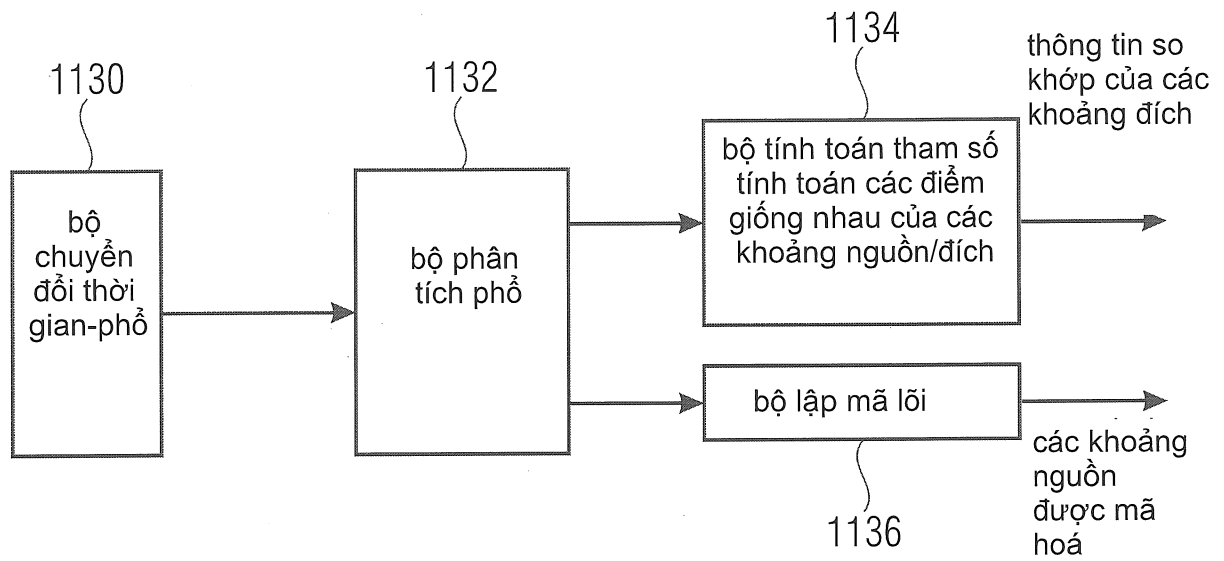


FIG 11C

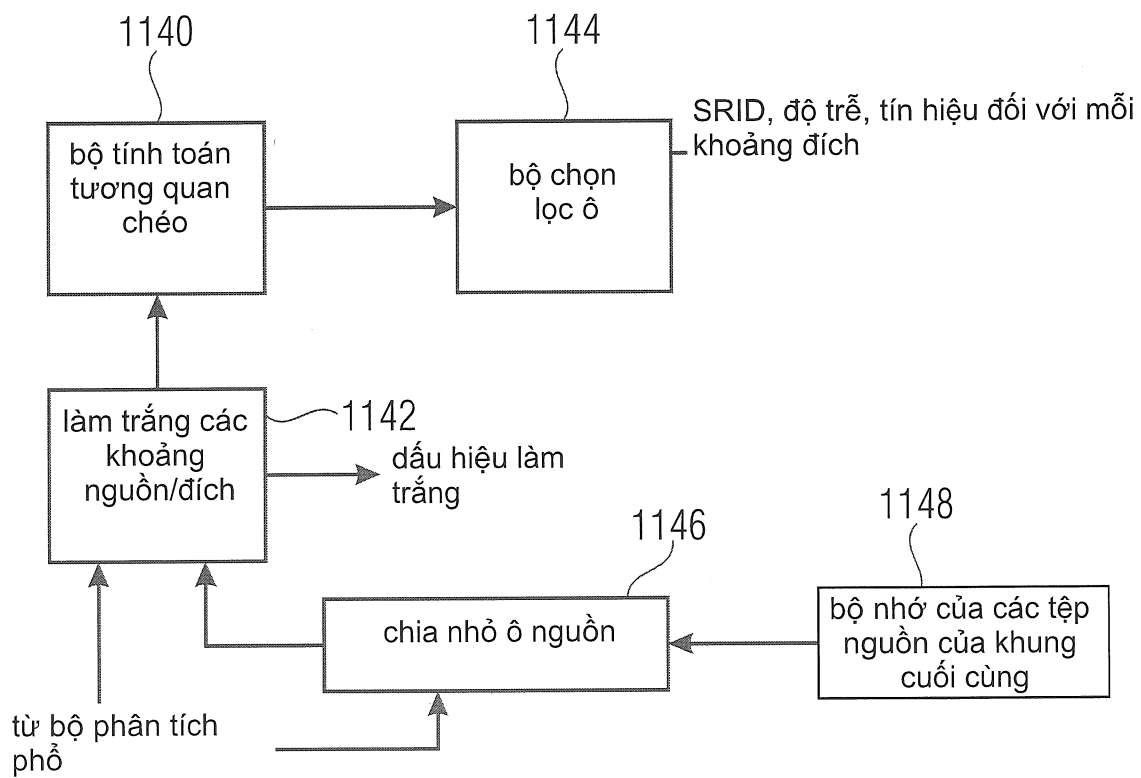


FIG 11D

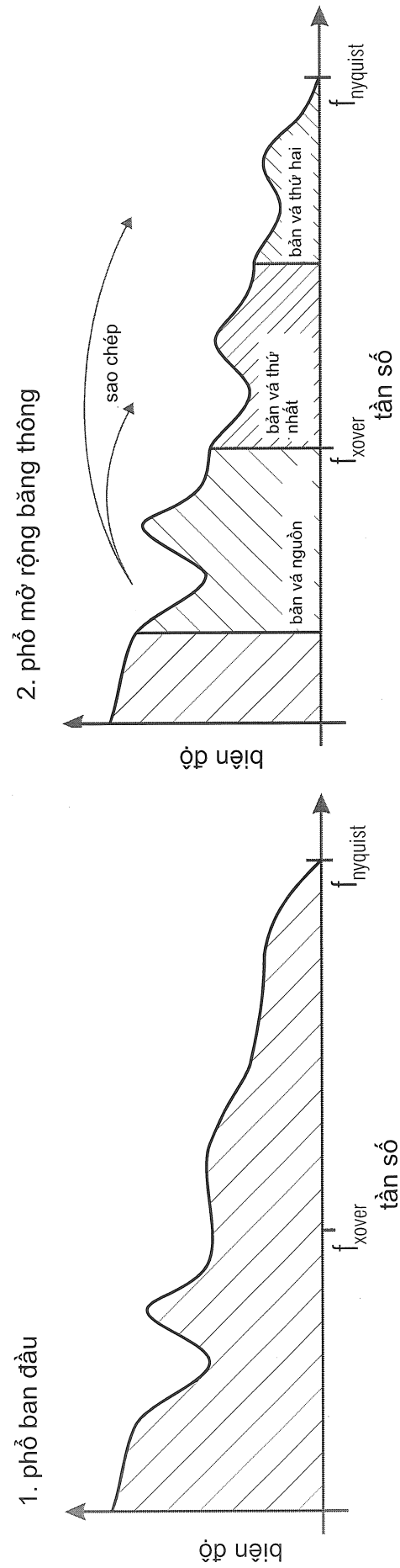


FIG 12A

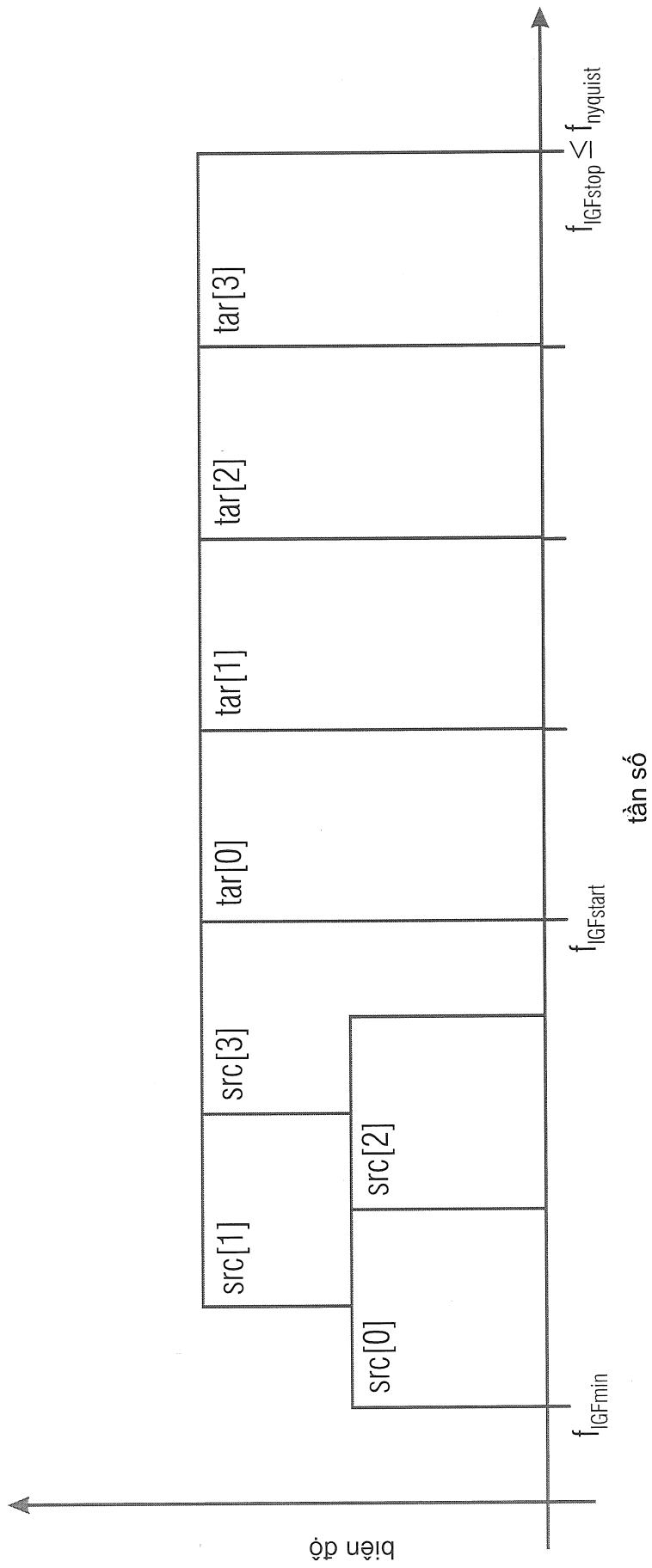


FIG 12B

ví dụ độ tương quan của hai tín hiệu

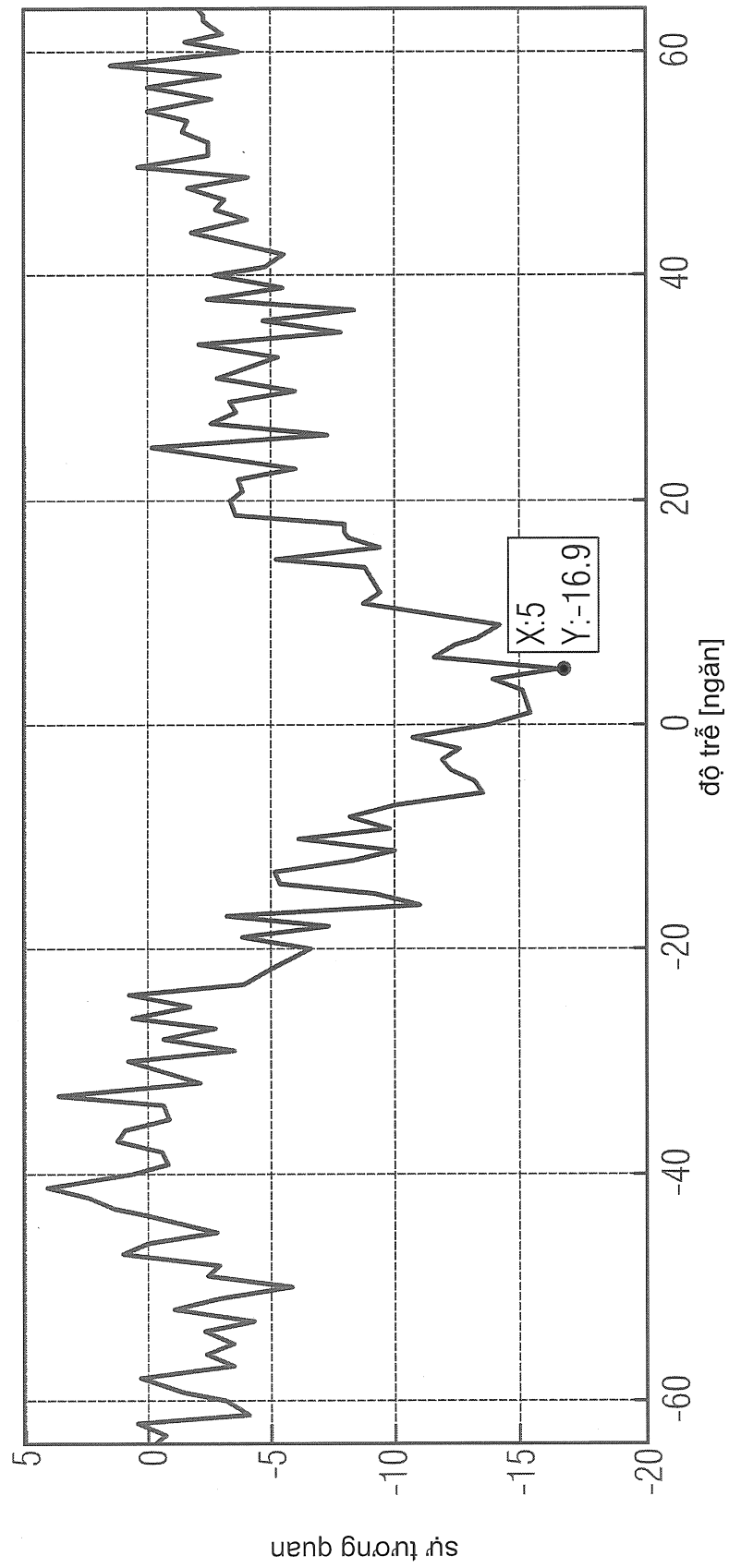


FIG 12C

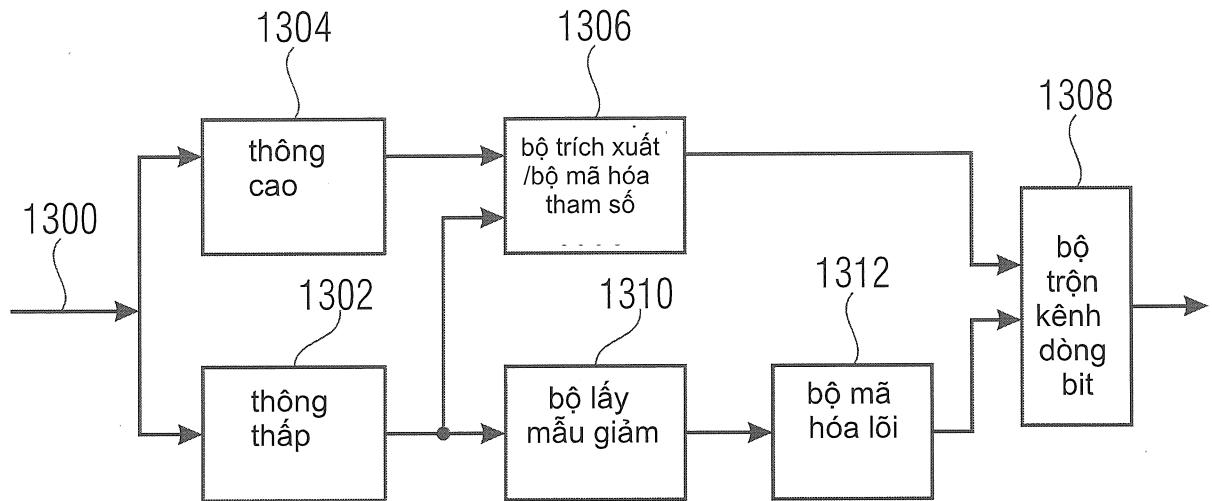


FIG 13A  
(TÍNH TRẠNG KỸ THUẬT)

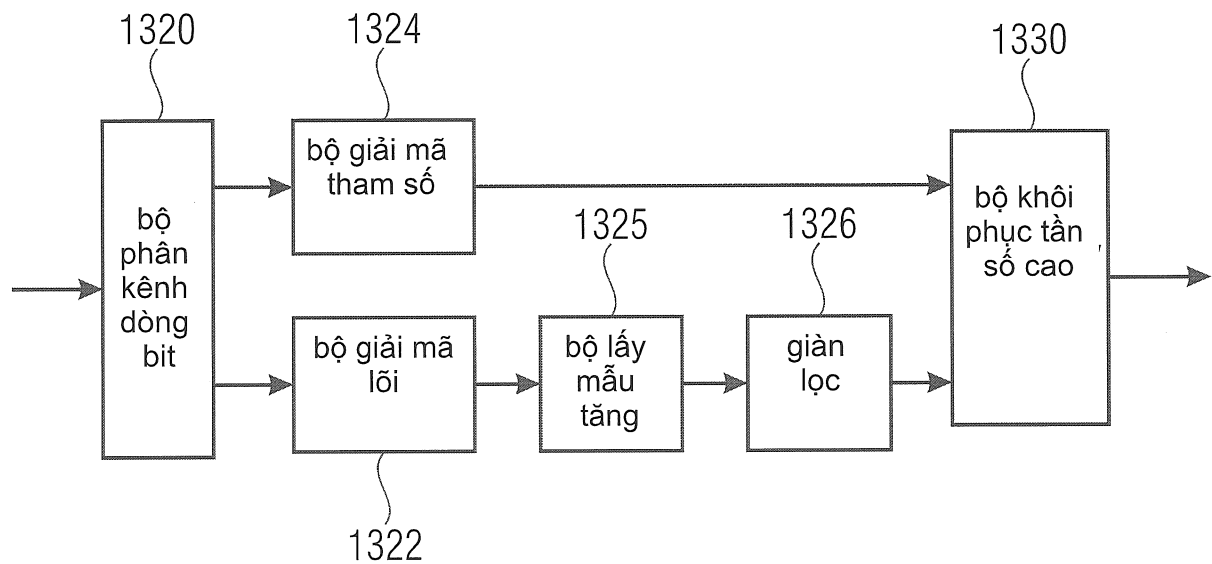


FIG 13B  
(TÍNH TRẠNG KỸ THUẬT)