



(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) 
1-0022609

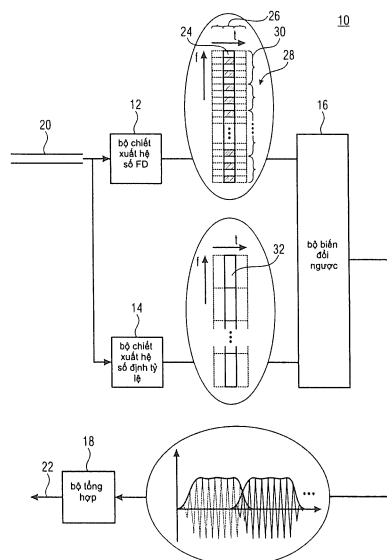
(51)⁷ **G10L 19/022**

(13) **B**

- (21) 1-2016-00610 (22) 15.07.2014
(86) PCT/EP2014/065169 15.07.2014 (87) WO2015/010965 29.01.2015
(30) 13177373.1 22.07.2013 EP
13189334.9 18.10.2013 EP
(45) 25.12.2019 381 (43) 25.04.2016 337
(73) Fraunhofer-Gesellschaft zur Foerderung der angewandten Forschung e.V. (DE)
Hansastrasse 27c, 80686 Muenchen, Germany
(72) DICK, Sascha (DE), HELMRICH, Christian (DE), HOELZER, Andreas (AT)
(74) Công ty Luật TNHH AMBYS Hà Nội (AMBYS HANOI)

(54) **BỘ GIẢI MÃ VÀ BỘ MÃ HÓA ÂM THANH MIỀN TẦN SỐ HỖ TRỢ CHUYỂN ĐỔI ĐỘ DÀI BIẾN ĐỔI VÀ PHƯƠNG PHÁP GIẢI MÃ VÀ MÃ HÓA ÂM THANH MIỀN TẦN SỐ HỖ TRỢ CHUYỂN ĐỔI ĐỘ DÀI BIẾN ĐỔI**

(57) Sáng chế đề cập đến bộ giải mã và bộ mã hóa âm thanh miền tần số hỗ trợ chuyển đổi độ dài biến đổi và phương pháp giải mã và mã hóa âm thanh miền tần số hỗ trợ chuyển đổi độ dài biến đổi có khả năng hỗ trợ thêm độ dài biến đổi nhất định theo cách tương thích ngược, bằng cách sau đây: các hệ số miền tần số của khung tương ứng được truyền theo cách đan xen bất kể việc báo hiệu sự tín hiệu hóa đối với các khung mà trong đó độ dài biến đổi thực sự áp dụng cho nó, và ngoài ra việc chiết xuất hệ số miền tần số và việc chiết xuất hệ số thang tỷ lệ hoạt động độc lập với sự tín hiệu hóa. Nhờ cách này, các bộ lập mã/bộ giải mã âm thanh miền tần số loại cũ, không nhạy đối với sự tín hiệu hóa, tuy nhiên sẽ có khả năng hoạt động không sai hỏng và tái tạo chất lượng hợp lý. Đồng thời, các bộ lập mã/bộ giải mã âm thanh miền tần số có khả năng hỗ trợ độ dài biến đổi sẽ đưa ra chất lượng thậm chí tốt hơn bất kể tính tương thích ngược. Do đó những bất lợi trong việc lập mã do việc lập mã các hệ số miền tần số theo cách rõ ràng đối với các bộ giải mã cũ hơn có liên quan, tương tự đối với tính chất không quan trọng do đan xen.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến việc lập mã âm thanh miền tần số hỗ trợ chuyển đổi độ dài biến đổi.

Tình trạng kỹ thuật của sáng chế

Các hệ thống lập mã tiếng nói/âm thanh miền tần số hiện đại như bộ mã hóa - giải mã Opus/Celt của IETF trong tài liệu tham khảo [1], MPEG-4 (HE-)AAC trong tài liệu tham khảo [2] hoặc, cụ thể là, MPEG-D xHE-AAC (USAC) trong tài liệu tham khảo [3], cung cấp các phương tiện để lập mã các khung âm thanh bằng cách sử dụng hoặc một biến đổi dài - khối dài - hoặc tám biến đổi ngắn nối tiếp nhau - các khối ngắn - phụ thuộc vào điểm dừng theo thời gian của tín hiệu.

Đối với các tín hiệu âm thanh nhất định như tiếng mưa hoặc tiếng vỗ tay ở khán phòng lớn, cả việc lập mã khối ngắn và dài đều không tạo ra chất lượng thỏa đáng ở tốc độ bit thấp. Điều này có thể được giải thích bởi mật độ của các chuyển tiếp nổi bật trong việc ghi lại này; việc lập mã chỉ với các khối dài có thể gây ra tần số và nhiễu theo thời gian nghe được của lỗi lập mã, còn được biết đến là tiếng vang trước, trong khi sự lập mã chỉ với các khối ngắn nhìn chung là không hiệu quả do tăng quá mức dữ liệu, dẫn đến các lỗ hổng quang phổ.

Theo đó, sẽ là thuận lợi khi có được giải pháp lập mã âm thanh miền tần số để hỗ trợ các độ dài biến đổi mà cũng phù hợp đối với các dạng vừa được nêu của các tín hiệu âm thanh. Cụ thể, sẽ là khả thi để thiết lập bộ mã hóa - giải mã âm thanh miền tần số mới, hỗ trợ việc chuyển đổi giữa tập hợp của các độ dài biến đổi mà, ví dụ, bao gồm độ dài biến đổi mong muốn nhất định phù hợp với dạng nhất định của tín hiệu âm thanh. Tuy nhiên, đó không phải là nhiệm vụ dễ dàng để có được bộ mã hóa-giải mã âm thanh miền tần số mới được chấp nhận trong thị trường. Các bộ mã hóa - giải mã phổ biến đã có sẵn và được sử dụng thường xuyên. Theo đó, sẽ là thuận lợi để có thể

có được giải pháp cho phép các bộ mã hóa - giải mã âm thanh miền tần số hiện nay được mở rộng theo cách để hỗ trợ thêm độ dài biến đổi mới, được mong muốn, nhưng tuy nhiên vẫn giữ tính tương thích ngược với các bộ lập mã và các bộ giải mã hiện nay.

Tài liệu tham khảo

- [1] Internet Engineering Task Force (IETF), RFC 6716, "Definition of the Opus Audio Codec," Proposed Standard, Sep. 2012. Available online at <http://tools.ietf.org/html/rfc6716>.
- [2] International Organization for Standardization, ISO/IEC 14496-3:2009, "Information Technology – Coding of audio-visual objects – Part 3: Audio," Geneva, Switzerland, Aug. 2009.
- [3] M. Neuendorf et al., "MPEG Unified Speech and Audio Coding – The ISO/MPEG Standard for High-Efficiency Audio Coding of All Content Types," in Proc. 132nd Convention of the AES, Budapest, Hungary, Apr. 2012. Also to appear in the Journal of the AES, 2013.
- [4] International Organization for Standardization, ISO/IEC 23003-3:2012, "Information Technology – MPEG audio – Part 3: Unified speech and audio coding," Geneva, Jan. 2012.
- [5] J.D.Johnston and A.J.Ferreira, "Sum-Difference stereo Transform Coding", in Proc. IEEE ICASSP-92, Vol. 2, March 1992.
- [6] N.Rettelbach, et al., European Patent EP2304719A1, "Audio Encoder, Audio Decoder, Methods for Encoding and Decoding an Audio Signal, Audio Stream and Computer Program", April 2011.

Bản chất kỹ thuật của sáng chế

Theo đó, mục đích của sáng chế là đề xuất giải pháp như nêu trên để cho phép các bộ mã hóa - giải mã âm thanh miền tần số hiện nay được mở rộng, theo cách tương thích ngược, theo hướng hỗ trợ độ dài biến đổi thêm sao cho chuyển đổi giữa các độ dài biến đổi cũng bao gồm độ dài biến đổi mới này.

Mục đích này của sáng chế đạt được bởi đối tượng trong các điểm yêu cầu bảo hộ độc lập được kèm theo ở đây.

Sáng chế được dựa trên việc phát hiện ra rằng bộ mã hóa - giải mã âm thanh miền tần số có thể được cung cấp với khả năng hỗ trợ thêm độ dài biến đổi nhất định theo cách tương thích ngược, khi các hệ số miền tần số của khung tương ứng được truyền theo cách đan xen không kể việc tín hiệu hóa báo hiệu đối với các khung mà trong đó độ dài biến đổi áp dụng vào một cách thực tế, và ngoài ra khi việc chiết xuất hệ số miền tần số và việc chiết xuất hệ số thang tỷ lệ hoạt động độc lập với sự tín hiệu hóa. Nhờ biện pháp này, các bộ lập mã/bộ giải mã âm thanh miền tần số loại cũ, không nhảy đối với sự tín hiệu hóa, tuy nhiên sẽ có khả năng hoạt động không có lỗi và với việc tái tạo chất lượng hợp lý. Đồng thời, các bộ lập mã/bộ giải mã âm thanh miền tần số chịu trách nhiệm cho việc chuyển đổi đến/từ độ dài biến đổi được hỗ trợ thêm sẽ đạt được chất lượng thậm chí tốt hơn bất kể tính tương thích ngược. Theo các bất lợi về hiệu quả lập mã do việc lập mã của các hệ số miền tần số theo cách rõ ràng đối với các bộ giải mã cũ hơn, có liên quan, tương tự đối với bản chất tương đối yếu do sự đan xen.

Việc thực hiện thuận lợi sáng chế là đối tượng của điểm yêu cầu bảo phụ thuộc.

Mô tả vắn tắt các hình vẽ

Cụ thể là, các phương án được ưu tiên của sáng chế được mô tả dưới đây kết hợp với các hình vẽ, trong đó:

Fig.1 thể hiện biểu đồ khối dưới dạng sơ đồ của bộ giải mã âm thanh miền tần số theo một phương án;

Fig.2 thể hiện sự minh họa dưới dạng sơ đồ chức năng của bộ biến đổi ngược theo Fig.1;

Fig.3 thể hiện sự minh họa dưới dạng sơ đồ của phương án thay thế có thể có của quy trình lọc định hình nhiễu âm miền thời gian (temporal noise shaping-TNS) ngược trên Fig.2 có hướng ngược lại theo phương án;

Fig.4 thể hiện khả năng lựa chọn các cửa sổ khi sử dụng việc chia tách biến đổi đối với cửa sổ kết thúc - bắt đầu dài trong USAC theo phương án; và

Fig.5 thể hiện sơ đồ khối của bộ mã hóa âm thanh miền tần số theo phương án.

Mô tả chi tiết sáng chế

Fig.1 thể hiện bộ giải mã âm thanh miền tần số hỗ trợ việc chuyển đổi độ dài biến đổi theo phương án của sáng chế. Bộ giải mã âm thanh miền tần số thuộc Fig.1 nhìn chung là được biểu thị với việc sử dụng ký hiệu tham chiếu 10 và bao gồm bộ chiết xuất hệ số miền tần số 12, bộ chiết xuất hệ số định tỷ lệ 14, bộ biến đổi ngược 16, và bộ tổng hợp 18. Tại đầu vào của chúng, bộ chiết xuất hệ số miền tần số và bộ chiết xuất hệ số định tỷ lệ 12 và 14 có sự truy cập đến dòng dữ liệu phản hồi 20. Các đầu ra của bộ chiết xuất hệ số miền tần số 12 và bộ chiết xuất hệ số định tỷ lệ 14 được kết nối với các đầu vào tương ứng của bộ biến đổi ngược 16. Đầu ra của bộ biến đổi ngược 16, lần lượt, được kết nối với đầu vào của bộ tổng hợp 18. Bộ tổng hợp này phát ra tín hiệu âm thanh được khôi phục tại đầu ra 22 của bộ mã hóa 10.

Bộ chiết xuất hệ số miền tần số 12 được tạo cấu hình để chiết xuất các hệ số miền tần số 24 của các khung 26 của tín hiệu âm thanh từ dòng dữ liệu 20. Các hệ số miền tần số 24 có thể là các hệ số biến đổi cosin rời rạc được cải biên (modified discrete cosine transform-MDCT) hoặc có thể thuộc vào một số biến đổi khác như biến đổi được nối chồng khác. Theo cách được mô tả thêm dưới đây, các hệ số miền tần số 24 thuộc về khung 26 nhất định mô tả phổ của tín hiệu âm thanh trong khung 26 tương ứng trong sự phân giải quang phổ theo thời gian biến thiên. Các khung 26 thể hiện các phần theo thời gian mà trong đó tín hiệu âm thanh được chia nhỏ một cách nối tiếp theo thời gian. Đặt cùng nhau tất cả các hệ số miền tần số 24 của tất cả các khung, cùng thể hiện ảnh phổ 28 của tín hiệu âm thanh. Các khung 26 có thể, ví dụ, có độ dài bằng nhau. Do loại nội dung âm thanh của tín hiệu âm thanh thay đổi theo thời gian, có thể bất lợi khi mô tả phổ đối với từng khung 26 với sự phân giải quang phổ theo thời gian liên tục bằng cách sử dụng, ví dụ, các biến đổi có độ dài biến đổi bất biến mà mở rộng, ví dụ, độ dài theo thời gian của từng khung 26, tức là bao gồm các giá trị mẫu trong khung 26 này của tín hiệu âm thanh cũng như các mẫu miền thời gian đi trước và theo sau khung tương ứng. Các giả âm tiếng vang trước có thể, ví dụ, do việc truyền phổ hao hụt của khung tương ứng ở dạng các hệ số miền tần số 24. Do đó, theo cách được nêu thêm dưới đây, các hệ số miền tần số 24 của khung 26 tương ứng

mô tả phổ của tín hiệu âm thanh trong khung 26 này trong sự phân giải quang phổ theo thời gian chuyển đổi được bằng việc chuyển đổi giữa các độ dài biến đổi khác nhau. Trong phạm vi bộ chiết xuất hệ số miền tần số 12 có liên quan, tuy nhiên, trường hợp sau là rõ ràng cho bộ chiết xuất tương tự. Bộ chiết xuất hệ số miền tần số 12 hoạt động độc lập với sự tín hiệu hóa bất kỳ, báo hiệu về việc chuyển đổi vừa được đề cập giữa các độ phân giải quang phổ theo thời gian khác nhau đối với các khung 26.

Bộ chiết xuất hệ số miền tần số 12 có thể sử dụng sự lập mã entropi để chiết xuất các hệ số miền tần số 24 từ dòng dữ liệu 20. Ví dụ, bộ chiết xuất hệ số miền tần số có thể sử dụng sự giải mã entropi dựa trên phạm vi, như sự giải mã số học phạm vi biến thiên, để chiết xuất các hệ số miền tần số 24 từ dòng dữ liệu 20 với chỉ định, cho từng hệ số trong số các hệ số miền tần số 24, cùng phạm vi bất kể việc tín hiệu hóa được đề cập trên đây tạo tín hiệu phân giải quang phổ theo thời gian của khung 26 mà hệ số miền tần số tương ứng thuộc về. Theo cách khác, và là ví dụ thứ hai, bộ chiết xuất 12 có thể sử dụng sự giải mã Huffman và xác định tập hợp các từ mã Huffman bất kể việc tín hiệu hóa này đặc trưng hóa sự phân giải của khung 26.

Các khả năng khác nhau tồn tại theo cách các hệ số miền tần số 24 mô tả ảnh phổ 28. Ví dụ, các hệ số miền tần số 24 có thể chỉ thể hiện một số số dư dự báo. Ví dụ, các hệ số miền tần số có thể thể hiện số dư của phép dự báo mà, ít nhất một phần, thu được bằng phép dự báo âm thanh nổi từ tín hiệu âm thanh khác thể hiện kênh âm thanh tương ứng hoặc trộn giảm trong số tín hiệu âm thanh đa kênh mà trong đó tín hiệu ảnh phổ 28 thuộc về. Theo cách khác, hoặc ngoài số dư dự báo, các hệ số miền tần số 24 có thể thể hiện tín hiệu tổng (giữa) hoặc hiệu (bên) theo mẫu âm thanh nổi M/S (giữa/bên-mid/side) trong tài liệu tham khảo [5]. Hơn nữa, các hệ số miền tần số 24 có thể phải trải qua sự định hình nhiễu âm theo thời gian.

Theo đó, các hệ số miền tần số 12 được lượng tử hóa và để giữ lỗi lượng tử hóa thấp hơn ngưỡng phát hiện (hoặc che giấu) tâm thính học, ví dụ, kích cỡ bước lượng tử hóa được biến thiên theo quang phổ theo cách được kiểm soát nhờ các hệ số định tỷ lệ tương ứng kết hợp với các hệ số miền tần số 24. Bộ chiết xuất hệ số định tỷ lệ 14 chịu trách nhiệm để chiết xuất các hệ số định tỷ lệ từ dòng dữ liệu 20.

Tóm lại, chi tiết hơn một chút về việc chuyển đổi giữa các độ phân giải quang phổ theo thời gian khác nhau từ khung này đến khung khác, chú ý tới phần sau đây. Như sẽ được mô tả chi tiết hơn dưới đây, việc chuyển đổi giữa các độ phân giải quang phổ theo thời gian khác nhau sẽ biểu thị rằng, trong khung 26 nhất định, tất cả hệ số miền tần số 24 thuộc về một biến đổi, hoặc là các hệ số miền tần số 24 của khung 26 tương ứng thực tế thuộc về các biến đổi khác nhau như, ví dụ, hai biến đổi, mà độ dài biến đổi của chúng bằng một nửa độ dài biến đổi của một biến đổi vừa được đề cập. Phương án được mô tả sau đây về các hình vẽ một mặt giả định việc chuyển đổi giữa một biến đổi và mặt khác 2 biến đổi, nhưng trong thực tế, việc chuyển đổi giữa một biến đổi và nhiều hơn hai biến đổi sẽ, về nguyên tắc, cũng khả thi như với các phương án được đưa ra dưới đây dễ dàng có thể chuyển sang các phương án thay thế này.

Fig.1 minh họa, bằng cách sử dụng đường gạch bóng, trường hợp làm mẫu là khung hiện thời là dạng được thể hiện bởi hai biến đổi ngắn, một trong đó được bắt nguồn bằng cách sử dụng nửa sau của khung hiện thời 26, và biến đổi còn lại thu được bằng cách biến đổi nửa đầu tiên của khung hiện thời 26 của tín hiệu âm thanh. Do độ dài biến đổi được rút ngắn, độ phân giải quang phổ mà trong đó các hệ số miền tần số 24 mô tả phổ của khung 26 được làm giảm, cụ thể là giảm một nửa trong trường hợp sử dụng hai biến đổi ngắn, trong khi độ phân giải theo thời gian được làm tăng, cụ thể là gấp đôi trong trường hợp này. Trên Fig.1, ví dụ, các hệ số miền tần số 24 được thể hiện bằng các nét gạch bóng sẽ thuộc về biến đổi dẫn trước, trong khi các hệ số không được gạch bóng sẽ thuộc về biến đổi theo sau. Do đó, các hệ số miền tần số cùng định vị theo quang phổ 24, mô tả thành phần quang phổ tương tự của tín hiệu âm thanh bên trong khung 26, nhưng ở các khoảng cách thời gian khác nhau không đáng kể, cụ thể là ở hai cửa sổ biến đổi liên tiếp của khung phân tách biến đổi.

Trong dòng dữ liệu 20, các hệ số miền tần số 24 được truyền theo cách đan xen sao cho các hệ số miền tần số tương ứng về quang phổ của hai biến đổi khác nhau nối tiếp ngay theo nhau. Thậm chí nói cách khác, các hệ số miền tần số 24 của khung biến đổi phân tách, tức là khung 26 mà trong đó việc phân tách biến đổi được báo hiệu trong dòng dữ liệu 20, được truyền dẫn sao cho nếu các hệ số miền tần số 24 như nhận được từ bộ chiết xuất hệ số miền tần số 12 sẽ được sắp xếp một cách nối tiếp theo cách

như thể chúng là các hệ số miền tần số của biến đổi dài, sau đó chúng được sắp xếp trong chuỗi này theo cách đan xen sao cho các hệ số miền tần số cùng định vị về quang phổ 24 ở ngay bên cạnh nhau và các cặp hệ số miền tần số cùng định vị về quang phổ 24 này được sắp xếp theo thứ tự quang phổ/tần số. Thứ vị là, khi được sắp xếp theo cách này, chuỗi gồm các hệ số miền tần số đan xen 24 trông tương tự với chuỗi gồm các hệ số miền tần số 24 thu được bằng một biến đổi dài. Một lần nữa, trong phạm vi bộ chiết xuất hệ số miền tần số 12 được đề cập, việc chuyển đổi giữa các độ dài biến đổi hoặc các độ phân giải quang phổ theo thời gian trong các đơn vị của các khung 26 là rõ ràng cho bộ chiết xuất tương tự, và do đó, việc lựa chọn phạm vi để lập mã entropi các hệ số miền tần số 24 theo cách có khả năng tương thích với phạm vi dẫn đến phạm vi tương tự được chọn - bất kể khung hiện thời thực tế là khung biến đổi dài hoặc khung hiện thời của dạng biến đổi phân tách không có bộ chiết xuất 12 biết đến. Ví dụ, bộ chiết xuất hệ số miền tần số 12 có thể lựa chọn phạm vi sẽ sử dụng đối với hệ số miền tần số nhất định dựa trên các hệ số miền tần số đã được lập mã/giải mã trong vùng lân cận theo quang phổ-thời gian với vùng lân cận theo quang phổ-thời gian này được xác định trong trạng thái đan xen được mô tả trên Fig.1. Điều này có hệ quả sau đây. Hãy hình dung, hệ số miền tần số được lập mã/giải mã hiện thời 24 là phần của biến đổi đầu tiên được biểu thị bằng cách sử dụng đường gạch bóng trên Fig.1. Hệ số miền tần số quang phổ liền kề ngay bên cạnh khi đó thực tế sẽ là hệ số miền tần số 24 của cùng biến đổi dẫn đầu (tức là, hệ số được gạch bóng trên Fig.1). Tuy nhiên, bộ chiết xuất hệ số miền tần số 12 sử dụng để lựa chọn phạm vi, hệ số miền tần số 24 thuộc về biến đổi theo sau, cụ thể là biến đổi lân cận về quang phổ (theo phân giải quang phổ được làm giảm của biến đổi được rút ngắn), giả sử rằng biến đổi sau sẽ lân cận quang phổ sát cạnh của một biến đổi dài của hệ số miền tần số hiện thời 24. Tương tự, trong việc lựa chọn phạm vi đối với hệ số miền tần số 24 của biến đổi theo sau, bộ chiết xuất hệ số miền tần số 12 sẽ sử dụng như là lân cận quang phổ ngay sát hệ số miền tần số 24 thuộc về biến đổi dẫn trước, và thực tế là cùng định vị về quang phổ với hệ số đó. Cụ thể là, thứ tự giải mã được xác định trong số các hệ số 24 của khung hiện thời 26 có thể dẫn trước, ví dụ, từ tần số thấp nhất đến tần số cao nhất. Các quan sát tương tự có hiệu lực trong trường hợp bộ chiết xuất hệ số miền tần số 12 được tạo cấu hình để giải mã entropi các hệ số miền tần số 24 của khung hiện

thời 26 trong các nhóm/bộ dữ liệu của các hệ số miền tần số 24 nối tiếp liên kề khi được tạo sắp xếp không khử đan xen. Thay vì sử dụng bộ dữ liệu của các hệ số miền tần số liên kề về quang phổ 24 chỉ thuộc về cùng biến đổi ngắn, bộ chiết xuất hệ số miền tần số 12 sẽ lựa chọn phạm vi cho bộ dữ liệu nhất định của hỗn hợp các hệ số miền tần số 24 thuộc về các biến đổi ngắn khác nhau trên cơ sở của bộ dữ liệu lân cận về quang phổ của hỗn hợp các hệ số miền tần số 24 thuộc về các biến đổi khác nhau.

Do thực tế là, như được chỉ ra trên đây, trong trạng thái đan xen, phổ tạo thành như thu được bằng 2 biến đổi ngắn trông tương tự với quang phổ thu được bằng một biến đổi dài, bất lợi lập mã entropi dẫn đến từ thao tác bất khả thi của bộ chiết xuất hệ số miền tần số 12 đối với chuyển đổi độ dài biến đổi là thấp.

Sự mô tả của bộ giải mã 10 được tiếp tục với bộ chiết hệ số định tỷ lệ 14 mà, như được đề cập trên đây, chịu trách nhiệm để chiết xuất các hệ số định tỷ lệ của các hệ số miền tần số 24 từ dòng dữ liệu 20. Phân giải quang phổ mà trong đó các hệ số tỷ lệ được chỉ định đến các hệ số miền tần số 24 thô hơn so với phân giải quang phổ tương đối mịn được hỗ trợ bởi biến đổi dài. Như được minh họa bằng các dấu ngoặc móc 30, các hệ số miền tần số 24 có thể được nhóm thành nhiều dải hệ số tỷ lệ. Việc chia nhỏ các dải hệ số tỷ lệ có thể được chọn dựa trên tư duy tâm thính học và có thể, ví dụ, trùng khớp với các dải gọi là Bark (hoặc tới hạn). Do bộ chiết xuất hệ số định tỷ lệ 14 là bất khả thi đối với việc chuyển đổi độ dài biến đổi, đúng như bộ chiết xuất hệ số miền tần số 12, bộ chiết xuất hệ số định tỷ lệ 14 giả định từng khung 26 được chia nhỏ thành nhiều dải hệ số tỷ lệ 30 mà bằng nhau, bất kể việc tín hiệu hóa chuyển đổi độ dài biến đổi, và chiết xuất cho từng dải hệ số tỷ lệ 30 này hệ số tỷ lệ 32. Tại phía bộ mã hóa, chức năng của các hệ số miền tần số 24 đối với các dải hệ số tỷ lệ 30 này được thực hiện trong trạng thái không khử đan xen được minh họa trên Fig.1. Kết quả là, trong phạm vi các khung 26 tương ứng với biến đổi phân tách liên quan, từng hệ số tỷ lệ 32 thuộc về nhóm được bao gồm bởi cả hai, các hệ số miền tần số 24 của biến đổi dẫn trước và các hệ số miền tần số 24 của biến đổi theo sau.

Bộ biến đổi ngược 16 được tạo cấu hình để nhận đối với từng khung 26 các hệ số miền tần số 24 tương ứng và các hệ số tỷ lệ 32 tương ứng và để các hệ số miền tần số 24 của khung 26, được định tỷ lệ theo các hệ số tỷ lệ 32, được thực hiện phép biến

đổi ngược để thu được các phần miền thời gian của tín hiệu âm thanh. Biến đổi chồng có thể được sử dụng bởi bộ biến đổi ngược 16 như, ví dụ, biến đổi cosin rời rạc được cải biên (modified discrete cosine transform - MDCT). Bộ tổng hợp 18 kết hợp các phần miền thời gian để thu được tín hiệu âm thanh như là bằng cách sử dụng, ví dụ, quy trình chồng lấp - bổ sung phù hợp dẫn đến, ví dụ, sự xóa bỏ việc tạo răng cưa miền thời gian trong các phần chồng lấp của các phần miền thời gian được tạo ra bởi bộ biến đổi ngược 16.

Tất nhiên, bộ biến ngược 16 đáp lại việc chuyển đổi độ dài biến đổi được đề cập trên đây được tạo tín hiệu trong dòng dữ liệu 20 đối với các khung 26. Sự hoạt động của bộ biến đổi ngược 16 được mô tả chi tiết hơn bằng Fig.2.

Fig.2 thể hiện chi tiết hơn về cấu trúc bên trong có thể có của bộ biến đổi ngược 16. Như được biểu thị trên Fig.2, bộ biến đổi ngược 16 nhận cho khung hiện thời các hệ số miền tần số 24 kết hợp với khung đó, cũng như các hệ số tỷ lệ tương ứng 32 để giải lượng tử hóa các hệ số miền tần số 24. Hơn nữa, bộ biến đổi ngược 16 được kiểm soát bằng sự tín hiệu hóa 34 mà có mặt trong dòng dữ liệu 20 đối với từng khung. Bộ biến đổi ngược 16 cũng có thể được kiểm soát nhờ các thành phần khác của dòng dữ liệu 20 một cách tùy ý được chứa trong đó. Trong phần mô tả sau đây, các chi tiết liên quan đến các thông số bổ sung này được mô tả.

Như được thể hiện trên Fig.2, bộ biến đổi ngược 16 theo Fig.2 bao gồm bộ giải lượng tử hóa 36, bộ khử đan xen có thể kích hoạt 38 và giai đoạn biến đổi ngược 40. Để dễ dàng hiểu phần mô tả sau đây, các hệ số miền tần số phản hồi 24 như được suy ra đối với khung hiện thời từ bộ chiết xuất hệ số miền tần số 12 được thể hiện được đánh số từ 0 đến $N-1$. Mặt khác, do bộ chiết xuất hệ số miền tần số 12 là không cần quan tâm đối với, tức là, hoạt động độc lập với, sự tín hiệu hóa 34, bộ chiết xuất hệ số miền tần số 12 cung cấp cho bộ biến đổi ngược 16 các hệ số miền tần số 24 bằng phương pháp tương tự bất kể khung hiện thời có dạng biến đổi phân tách, hoặc dạng 1 biến đổi, tức là, số lượng các hệ số miền tần số 24 là N trong trường hợp minh họa này và sự kết hợp của các chỉ số từ 0 đến $N-1$ với N hệ số miền tần số 24 cũng duy trì như vậy bất kể việc tín hiệu hóa 34. Trong trường hợp khung hiện thời là loại một biến đổi hoặc biến đổi dài, các chỉ số từ 0 đến $N-1$ tương ứng với việc sắp xếp của hệ số miền

tần số 24 từ tần số thấp hơn đến tần số cao nhất, và trong trường hợp khung hiện thời có dạng biến đổi phân tách, các chỉ số tương ứng với thứ tự cho các hệ số miền tần số khi được bố trí theo quang phổ theo thứ tự phổ của chúng, nhưng theo cách đan xen sao cho từng hệ số miền tần số 24 thứ hai thuộc về biến đổi theo sau, trong khi các hệ số khác thuộc về biến đổi dẫn trước.

Các bằng chứng tương tự đúng đối với các hệ số thang tỷ lệ 32. Do bộ chiết xuất hệ số thang tỷ lệ 14 hoạt động theo cách bất khả thi đối với sự tín hiệu hóa 34, số lượng và thứ tự cũng như các trị số của các hệ số thang tỷ lệ 32 đến từ bộ chiết xuất hệ số thang tỷ lệ 14 là độc lập với sự tín hiệu hóa 34, với các hệ số thang tỷ lệ 32 trên Fig.2 được biểu thị làm mẫu là S_0 đến S_M với chỉ số tương ứng với thứ tự kế tiếp nhau trong số các dải hệ số thang tỷ lệ mà trong đó các hệ số thang tỷ lệ này được kết hợp.

Theo cách tương tự với bộ chiết xuất hệ số miền tần số 12 và bộ chiết xuất hệ số thang tỷ lệ 14, bộ giải lượng tử hóa 36 có thể hoạt động bất khả thi đối với, hoặc độc lập với, sự tín hiệu hóa 34. Bộ giải lượng tử hóa 36 giải lượng tử hóa, hoặc định tỷ lệ, các hệ số miền tần số phản hồi 24 bằng cách sử dụng hệ số tỷ lệ kết hợp với dải hệ số tỷ lệ mà các hệ số miền tần số tương ứng thuộc về. Một lần nữa, mối liên quan của các hệ số miền tần số phản hồi 24 với các dải hệ số tỷ lệ riêng, và do đó sự kết hợp của các hệ số miền tần số phản hồi 24 với các hệ số tỷ lệ 32, là độc lập đối với sự tín hiệu hóa 34, và do đó bộ biến đổi ngược 16 đưa các hệ số miền tần số 24 trải qua việc định tỷ lệ theo các hệ số tỷ lệ 32 ở độ phân giải quang phổ mà là độc lập với sự tín hiệu hóa. Ví dụ, bộ giải lượng tử hóa 36, độc lập với sự tín hiệu hóa 34, chỉ định các hệ số miền tần số với các chỉ số 0 đến 3 cho dải hệ số tỷ lệ thứ nhất và do đó hệ số tỷ lệ thứ nhất S_0 , các hệ số miền tần số với các chỉ số 4 đến 9 cho dải hệ số tỷ lệ thứ hai và do đó hệ số tỷ lệ S_1 và tương tự. Các ranh giới hệ số tỷ lệ chỉ là để minh họa. Bộ giải lượng tử hóa 36 có thể, ví dụ, để giải lượng tử hóa các hệ số miền tần số 24 thực hiện phép nhân bằng cách sử dụng hệ số tỷ lệ kết hợp, tức là, tính toán hệ số miền tần số x_0 sẽ là $x_0 \cdot S_0$, x_1 sẽ là $x_1 \cdot S_0$, ..., x_3 sẽ là $x_3 \cdot S_0$, x_4 sẽ là $x_4 \cdot S_1$, ..., x_9 sẽ là $x_9 \cdot S_1$, và tương tự. Theo cách khác, bộ giải lượng tử hóa 36 có thể thực hiện việc nội suy các hệ số tỷ lệ được sử dụng thực tế cho việc giải lượng tử hóa của các hệ số miền tần số 24 từ phân giải quang phổ thô được xác định bằng các dải hệ số tỷ lệ. Việc nội suy có thể là độc lập

với sự tín hiệu hóa 34. Tuy nhiên, theo cách khác, sự nội suy độc lập này có thể phụ thuộc vào sự tín hiệu hóa để giải thích cho các vị trí lấy mẫu quang phổ theo thời gian khác nhau của các hệ số miền tần số 24 phụ thuộc vào khung hiện thời là loại biến đổi phân tách hoặc loại một biến đổi/biến đổi dài.

Fig.2 minh họa rằng cho đến phía đầu vào của bộ khử đan xen có thể kích hoạt 38, thứ tự giữa các hệ số miền tần số 24 duy trì như nhau và áp dụng tương tự, ít nhất về cơ bản là như nhau, đối với sự hoạt động tổng thể cho đến điểm đó. Fig.2 thể hiện rằng, ngược dòng của bộ khử đan xen có thể kích hoạt 38, các hoạt động khác có thể được thực hiện bởi bộ biến đổi ngược 16. Ví dụ, bộ biến đổi ngược 16 có thể được tạo cấu hình để thực hiện việc điền đầy nhiễu âm cho các hệ số miền tần số 24. Ví dụ, trong chuỗi gồm các hệ số miền tần số 24, các dải hệ số tỷ lệ, tức là, các nhóm gồm các hệ số miền tần số phản hồi theo thứ tự theo chỉ số từ 0 đến $N - 1$, có thể được nhận biết, trong đó tất cả hệ số miền tần số 24 của các dải hệ số tỷ lệ tương ứng được lượng tử hóa tới 0. Các hệ số miền tần số này có thể được điền đầy, ví dụ, bằng cách sử dụng việc tạo ra nhiễu âm nhân tạo như, ví dụ, bằng cách sử dụng bộ tạo số giả ngẫu nhiên. Cường độ/mức của nhiễu âm được điền đầy vào trong dải hệ số tỷ lệ được lượng tử hóa tới 0 có thể được điều chỉnh bằng cách sử dụng hệ số tỷ lệ thuộc dải hệ số tỷ lệ tương ứng do không cần hệ số tương tự để định tỷ lệ do tất cả các hệ số quang phổ trong đó bằng 0. Việc điền đầy nhiễu âm này được thể hiện trên Fig.2 ở số chỉ dẫn 40 và được mô tả chi tiết hơn trong phương án của sáng chế trong tài liệu EP2304719A1 mục tài liệu tham khảo [6].

Fig.2 còn thể hiện rằng, bộ biến đổi ngược 16 có thể được tạo cấu hình để hỗ trợ việc lập mã âm thanh nổi kết hợp và/hoặc sự dự báo âm thanh nổi liên kênh. Trong khuôn khổ của sự dự báo âm thanh nổi liên kênh, bộ biến đổi ngược 16 có thể, ví dụ, dự báo 42 phổ trong cách sắp xếp không đan xen được thể hiện bởi thứ tự của các chỉ số từ 0 đến $N - 1$ từ kênh khác của tín hiệu âm thanh. Tức là, có thể là các hệ số miền tần số 24 mô tả ảnh phổ của kênh của tín hiệu âm thanh nổi, và bộ biến đổi ngược 16 được tạo cấu hình để xử lý các hệ số miền tần số 24 như số dư dự báo của tín hiệu dự báo được suy ra từ kênh khác của tín hiệu âm thanh nổi này. Sự dự báo âm thanh nổi đa kênh này có thể, ví dụ, được thực hiện với một số độ kết hạt quang phổ

độc lập với sự tín hiệu hóa 34. Các thông số dự báo phức hợp 44 kiểm soát sự dự báo âm thanh nổi phức hợp 42 có thể, ví dụ, kích hoạt sự dự báo âm thanh nổi phức hợp 42 đối với các dải nhất định trong dải hệ số tỷ lệ được đề cập trên đây. Đối với từng dải hệ số tỷ lệ mà sự dự báo phức hợp được kích hoạt theo cách của các thông số dự báo phức hợp 44, các hệ số miền tần số được định tỷ lệ 24, được sắp xếp theo thứ tự 0 đến N-1, nằm trong dải hệ số tỷ lệ tương ứng, sẽ được lấy tổng với tín hiệu dự báo liên kênh thu được từ kênh khác của tín hiệu âm thanh nổi. Hệ số phức hợp được bao gồm trong các thông số dự báo phức hợp 44 đối với dải hệ số tỷ lệ tương ứng này có thể kiểm soát tín hiệu dự báo.

Hơn nữa, trong khuôn khổ lập mã âm thanh nổi kết hợp, bộ biến đổi ngược 16 có thể được tạo cấu hình để thực hiện việc giải mã giữa/bên (mid/side-MS) 46. Cụ thể là, bộ giải mã 10 theo Fig.1 có thể thực hiện các thao tác được mô tả trên đây 2 lần, một lần đối với kênh thứ nhất và lần còn lại đối với kênh thứ hai của tín hiệu âm thanh nổi, và được kiểm soát nhờ các thông số MS trong dòng dữ liệu 20, bộ biến đổi ngược 16 có thể giải mã MS hai kênh này hoặc để nguyên hai kênh này như chúng vốn là vậy, cụ thể là các kênh bên trái và kênh bên phải của tín hiệu âm thanh nổi. Các thông số MS 48 có thể chuyển đổi giữa việc lập mã MS tại mức khung hoặc thậm chí tại mức mịn hơn nhất định như trong các đơn vị của các dải hệ số tỷ lệ hoặc các nhóm của chúng. Trong trường hợp giải mã MS được kích hoạt, ví dụ, bộ biến đổi ngược 16 có thể tạo tổng của các hệ số miền tần số tương ứng 24 theo thứ tự của các hệ số 0 đến N-1, với các hệ số miền tần số tương ứng của kênh khác của tín hiệu âm thanh nổi, hoặc hiệu của chúng.

Fig.2 khi đó thể hiện rằng bộ khử đan xen có thể kích hoạt 38 đáp lại sự tín hiệu hóa 34 đối với khung hiện thời để, trong trường hợp khung hiện thời được tín hiệu hóa bằng sự tín hiệu hóa 34 để trở thành khung biến đổi phân tách, khử đan xen các hệ số miền tần số phản hồi sao cho thu được hai biến đổi, cụ thể là biến đổi dẫn trước 50 và biến đổi theo sau 52, và để các hệ số miền tần số được đan xen để tạo ra một biến đổi 54 trong trường hợp sự tín hiệu hóa 34 biểu thị khung hiện thời là khung biến đổi dài. Trong trường hợp khử đan xen, bộ khử đan xen 38 tạo thành một biến đổi trong số 50 và 52, biến đổi ngắn thứ nhất của hệ số miền tần số có các chỉ số bằng nhau, và biến

đôi ngắn khác của các hệ số miền tần số ở các vị trí chỉ số không bằng nhau. Ví dụ, các hệ số miền tần số có chỉ số bằng nhau có thể hình thành biến đổi dẫn trước (khi bắt đầu ở chỉ số 0), trong khi các hệ số khác hình thành biến đổi theo sau. Các biến đổi 50 và 52 trải qua biến đổi ngược có độ dài biến đổi ngắn hơn dẫn đến các phần miền thời gian lần lượt là 56 và 58. Bộ tổng hợp 18 theo Fig.1 định vị một cách chính xác các phần miền thời gian 56 và 58 theo thời gian, cụ thể là phần miền thời gian 56 tạo ra từ biến đổi dẫn trước 50 ở phía trước phần miền thời gian 58 tạo ra từ biến đổi theo sau 52, và thực hiện quy trình xếp chồng lặp - và - bổ sung ở đó, giữa và với các phần miền thời gian tạo ra từ các khung đi trước và theo sau của tín hiệu âm thanh. Trong trường hợp không khử đan xen, các hệ số miền tần số đến bộ đan xen 38 cấu tạo nên việc biến đổi dài 54 như chính là chúng, và giai đoạn biến đổi ngược 40 thực hiện biến đổi ngược trên đó để tạo ra phần miền thời gian 60 trải dài trên, và xa hơn, khung hiện thời 26 trong toàn bộ khoảng thời gian. Bộ tổng hợp 18 kết hợp phần miền thời gian 60 với các phần miền thời gian tương ứng tạo ra từ các khung đi trước và theo sau của tín hiệu âm thanh.

Bộ giải mã âm thanh miền tần số được mô tả trên đây giúp chuyển đổi độ dài biến đổi theo cách mà cho phép có thể tương thích với các bộ giải mã âm thanh miền tần số mà không đáp lại sự tín hiệu hóa 34. Cụ thể là, các bộ giải mã “loại cũ” này sẽ giả định nhầm rằng các khung mà thực tế được tín hiệu hóa bằng sự tín hiệu hóa 34 để là dạng biến đổi phân tách, là loại biến đổi dài. Nghĩa là, các bộ giải mã này sẽ nhầm lẫn trong việc để các hệ số miền tần số dạng phân tách được đan xen và thực hiện phép biến đổi ngược độ dài biến đổi dài. Tuy nhiên, chất lượng tạo thành của các khung bị ảnh hưởng bởi tín hiệu âm thanh được khôi phục sẽ vẫn khá hợp lý.

Nói cách khác, bất lợi về hiệu quả lập mã cũng vẫn khá hợp lý. Bất lợi về hiệu quả lập mã gây ra do sự tín hiệu hóa 34 bất kể khi các hệ số miền tần số và các hệ số tỷ lệ được mã hóa không cân nhắc ý nghĩa các của hệ số biến thiên và khai thác sự biến thiên này để làm tăng hiệu quả lập mã. Tuy nhiên, bất lợi này là tương đối nhỏ so với ưu điểm của phép tính tương thích ngược. Mệnh đề này cũng đúng đối với sự hạn chế để kích hoạt và khử kích hoạt bộ điền đầy nhiễu âm 40, sự dự báo âm thanh nổi phức hợp 42 và sự giải mã MS 46 chỉ trong các phần quang phổ liên tục (các dải hệ số tỷ lệ)

trong trạng thái khử đan xen được xác định bằng các chỉ số từ 0 đến $N-1$ trên Fig.2. Cơ hội để kiểm soát các công cụ lập mã này đặc biệt là đối với dạng khung (ví dụ, có 2 mức nhiễu âm) có thể có khả năng tạo ra các ưu điểm, nhưng các ưu điểm này được cố bù bởi ưu điểm ở việc có tính tương thích ngược.

Fig.2 thể hiện rằng, bộ giải mã theo Fig.1 thậm chí có thể được tạo cấu hình để hỗ trợ sự lập mã định hình nhiễu âm theo thời gian (Temporal Noise Shaping – TNS), tuy nhiên trong khi vẫn giữ tính tương thích ngược với các bộ giải mã không nhạy đối với sự tín hiệu hóa 34. Cụ thể là, Fig.2 minh họa khả năng thực hiện việc lọc TNS ngược sau sự dự báo âm thanh nội phức hợp 42 và sự giải mã MS 46 bất kỳ, nếu có. Để duy trì tính tương thích ngược, bộ biến đổi ngược 16 được tạo cấu hình để thực hiện việc lọc TNS ngược 62 trên dãy gồm N hệ số bất kể việc tín hiệu hóa 34 bằng cách sử dụng các hệ số TNS tương ứng 64. Nhờ cách này, dòng dữ liệu 20 lập mã các hệ số TNS 64 bằng nhau, bất kể việc tín hiệu hóa 34. Nghĩa là, số lượng các hệ số TNS và cách lập mã cho các hệ số tương tự là như nhau. Tuy nhiên, bộ biến đổi ngược 16 được tạo cấu hình để áp dụng theo cách khác nhau các hệ số TNS 64. Trong trường hợp khung hiện thời là khung biến đổi dài, việc lọc TNS ngược được thực hiện trên biến đổi dài 54, tức là các hệ số miền tần số kế tiếp nhau trong trạng thái đan xen, và trong trường hợp khung hiện thời được tín hiệu hóa bằng sự tín hiệu hóa 34 như là khung biến đổi phân tách, bộ biến đổi ngược 16 ngược các bộ lọc TNS 62 sự tương quan của biến đổi dẫn trước 50 và biến đổi theo sau 52, tức là chuỗi gồm các hệ số miền tần số của các chỉ số 0, 2, ..., $N-2$, 1, 3, 5, ..., $N-1$. Việc lọc TNS ngược 62 có thể, ví dụ, đòi hỏi bộ biến đổi ngược 16 áp dụng bộ lọc, mà chức năng truyền được thiết lập theo các hệ số TNS 64 theo chuỗi khử đan xen hoặc đan xen của các hệ số đã hoàn thành chuỗi xử lý bộ khử đan xen ngược dòng 38.

Do đó, bộ giải mã “loại cũ” mà xử lý một cách ngẫu nhiên các khung có dạng biến đổi phân tách như các khung biến đổi dài, áp dụng các hệ số TNS 64 mà được tạo ra bởi bộ mã hóa bằng cách phân tích sự tương quan của 2 biến đổi ngắn, cụ thể là 50 và 52, cho biến đổi 54 và do đó tạo ra, theo cách biến đổi ngược được áp dụng cho biến đổi 54, phần miền thời gian không chính xác 60. Tuy nhiên, thậm chí sự giảm chất lượng này ở các bộ giải mã này có thể chịu đựng được đối với người nghe trong

trường hợp hạn chế việc sử dụng của các khung biến đổi phân tách này trong các trường hợp mà tín hiệu thể hiện tiếng mưa hoặc tiếng vỗ tay hoặc âm thanh tương tự.

Để hoàn thiện, Fig.3 thể hiện rằng, việc lọc TNS ngược 62 của bộ biến đổi ngược 16 còn có thể được chèn vào trong chuỗi xử lý khác được thể hiện trên Fig.2. Ví dụ, việc lọc TNS ngược 62 có thể được định vị ngược chiều với sự dự báo âm thanh nổi phức hợp 42. Để giữ miền khử đan xen xuôi chiều và ngược chiều với việc lọc TNS ngược 62, Fig.3 thể hiện rằng trong trường hợp đó, các hệ số miền tần số 24 chỉ được khử đan xen sơ bộ 66, để thực hiện việc lọc TNS ngược 68 trong trạng thái tương quan được khử đan xen trong đó các hệ số miền tần số 24 như được xử lý trên đây theo chuỗi gồm các chỉ số 0, 2, 4, ..., N-2, 1, 3, ..., N-3, N-1, trong đó việc khử đan xen được đảo ngược 70 sao cho thu được các hệ số miền tần số trong phiên bản được lọc TNS ngược theo chuỗi đan xen của chúng 0, 1, 2, ..., N-1 một lần nữa. Vị trí của việc lọc TNS ngược 62 trong chuỗi các bước xử lý được thể hiện trên Fig.2 có thể được cố định hoặc có thể được tín hiệu hóa nhờ dòng dữ liệu 20 như, ví dụ, trên khung bằng cơ sở khung hoặc ở một số độ kết hạt khác.

Cần lưu ý rằng, với mục đích giảm sự mô tả, các phương án trên đây tập trung vào sự kề nhau của các khung biến đổi dài và các khung biến đổi phân tách. Tuy nhiên, các phương án của sáng chế còn có thể được mở rộng bằng sự giới thiệu các khung của dạng biến đổi khác như các khung gồm 8 biến đổi ngắn. Vì vậy, cần lưu ý rằng thuyết bất khả thi được đề cập trên đây, chỉ liên quan đến các khung riêng rẽ, theo cách tín hiệu hóa khác, từ các khung khác này của dạng biến đổi thứ ba bất kỳ sao cho bộ giải mã “loại cũ”, bằng cách kiểm tra sự tín hiệu hóa bao gồm trong tất cả các khung, xử lý ngẫu nhiên các khung biến đổi phân tách như các khung biến đổi dài, và chỉ các khung riêng rẽ từ các khung khác (tất cả ngoại trừ các khung biến đổi phân tách và các khung biến đổi dài) sẽ bao gồm sự tín hiệu hóa 34. Khi mà các khung khác này (tất cả ngoại trừ biến đổi phân tách và các khung biến đổi dài) liên quan, lưu ý rằng cách thức hoạt động của bộ chiết xuất 12 và 14 như việc lựa chọn phạm vi và tương tự như vậy có thể phụ thuộc vào tín hiệu hóa khác nữa, đó là, cách thức hoạt động này có thể khác với cách thức hoạt động được áp dụng đối với các khung biến đổi phân tách và các khung biến đổi dài.

Trước khi mô tả bộ mã hóa phù hợp thích hợp với các phương án của bộ giải mã được mô tả trên đây, việc thực hiện của các phương án trên đây được mô tả mà sẽ phù hợp do đó nâng cấp các bộ lập mã/bộ giải mã âm thanh trên cơ sở xHE-AAC để cho phép sự hỗ trợ việc phân tách biến đổi theo cách tương thích ngược.

Nghĩa là, sau đây mô tả cách thức thực hiện việc phân tách độ dài biến đổi trong bộ mã hóa - giải mã âm thanh mà được dựa trên MPEG-D xHE-AAC (USAC) với mục đích cải thiện chất lượng lập mã của các tín hiệu âm thanh nhất định với tốc độ bit thấp. Công cụ phân tách biến đổi được tín hiệu hóa một cách tương thích bán ngược sao cho các bộ giải mã xHE-AAC truyền thống có thể phân tích cú pháp và giải mã các dòng bit theo các phương án trên đây mà không có các lỗi âm thanh hoặc nhiễu âm rõ rệt. Như sẽ được thể hiện sau đây, sự tín hiệu hóa tương thích bán ngược này khai thác các trị số có thể không được sử dụng của việc kiểm soát phần tử cú pháp của khung, theo cách được lập mã có điều kiện, sự sử dụng việc điền đầy nhiễu âm. Trong khi các bộ giải mã xHE-AAC truyền thống không nhạy đối với các trị số có thể này của phần tử cú pháp điền đầy nhiễu âm tương ứng, các bộ giải mã âm thanh được cải thiện nhạy với các trị số này.

Cụ thể là, việc thực hiện được mô tả dưới đây giúp, cùng với các phương án được mô tả trên đây, đưa ra độ dài biến đổi trung gian nhằm lập mã các tín hiệu như tiếng mưa hoặc tiếng vỗ tay, tốt hơn là khối dài phân tách, tức là 2 biến đổi nối tiếp nhau, mỗi nửa hoặc một phần tư của độ dài quang phổ của khối dài, với sự chồng lấp thời gian tối đa giữa các biến đổi này nhỏ hơn sự chồng lấp theo thời gian tối đa giữa các khối dài liên tiếp. Để cho phép các dòng bit được lập mã với sự phân tách biến đổi, tức là sự tín hiệu hóa 34, có thể được đọc và phân tích bằng các bộ giải mã truyền thống xHE-AAC, việc phân tách nên được sử dụng theo cách tương thích bán ngược: sự có mặt của công cụ phân tách biến đổi này không nên làm cho bộ giải mã truyền thống dừng - hoặc thậm chí không bắt đầu - việc giải mã. Khả năng đọc của các dòng bit này bởi cấu trúc hạ tầng xHE-AAC còn có thể tạo thuận tiện cho việc chiếm lĩnh thị trường. Để đạt được mục tiêu vừa được đề cập về tính tương thích bán ngược nhằm sử dụng sự phân tách biến đổi trong phạm vi của xHE-AAC hoặc các biến thể có thể của nó, sự phân tách biến đổi được tín hiệu hóa sự tín hiệu hóa việc điền đầy nhiễu

âm của xHE-AAC. Theo các phương án được mô tả trên đây, để gắn sự phân tách biến đổi vào trong các bộ lập mã/bộ giải mã xHE-AAC, thay cho chuỗi cửa sổ ngừng - bắt đầu miền tần số (frequency-domain - FD) sự biến đổi phân tách bao gồm hai biến đổi riêng rẽ, nửa độ dài có thể được sử dụng. Các biến đổi nửa độ dài nối tiếp theo thời gian được đan xen vào khối dạng ngừng - bắt đầu đơn lẻ theo cách từng hệ số đối với các bộ giải mã mà không hỗ trợ sự phân tách biến đổi, tức là các bộ giải mã xHE-AAC truyền thống. Việc tín hiệu hóa nhờ điền đầy nhiễu âm tín hiệu hóa được thực hiện như được mô tả dưới đây. Cụ thể là, điền đầy nhiễu âm thông tin phụ 8 bit có thể được sử dụng để chuyển tải biến đổi phân tách. Việc tín hiệu hóa thông qua sự tín hiệu hóa điền đầy nhiễu âm được thực hiện như mô tả sau đây, Cụ thể, thông tin phụ điền đầy nhiễu âm 8 bit có thể được sử dụng để truyền sự phân tách biến đổi. Điều này là khả thi bởi vì tiêu chuẩn MPEG-D trong tài liệu tham khảo [4] nêu rằng, tất cả 8 bit được truyền dẫn thậm chí nếu mức nhiễu âm được áp dụng bằng 0. Trong tình huống đó, một số bit điền đầy nhiễu âm có thể được sử dụng lại cho sự phân tách biến đổi, tức là đối với sự tín hiệu hóa 34.

Tính tương thích bán ngược đối với sự phân tích và tái tạo dòng bit bằng các bộ giải mã truyền thống xHE-AAC có thể được đảm bảo như sau. Sự phân tách biến đổi được tạo tín hiệu nhờ mức nhiễu âm bằng 0, tức là tất cả 3 bit điền đầy nhiễu âm đầu tiên bằng 0, được theo sau bởi 5 bit khác 0 (mà theo lệ thường thể hiện độ lệch nhiễu âm) chứa thông tin phụ liên quan đến việc phân tách biến đổi cũng như mức nhiễu âm hao hụt. Do bộ giải mã xHE-AAC truyền thống bỏ qua trị số của độ lệch 5 bit nếu mức nhiễu âm 3 bit bằng 0, sự có mặt của sự tín hiệu hóa phân tách biến đổi 34 chỉ có hiệu quả đối với việc điền đầy nhiễu âm trong bộ giải mã truyền thống: việc điền đầy nhiễu âm bị ngắt do 3 bit đầu tiên bằng 0, và bộ duy trì của sự hoạt động giải mã hoạt động như mong muốn. Cụ thể là, biến đổi phân tách được xử lý giống như khối ngừng - bắt đầu truyền thống với biến đổi ngược toàn bộ độ dài (do việc đan xen hệ số được đề cập trên đây) và sự khử đan xen không được thực hiện. Sau đây, bộ giải mã truyền thống vẫn thực hiện việc giải mã "từ từ" dòng dữ liệu/dòng bit tăng cường 20 bởi vì bộ giải mã này không cản chận âm thanh tín hiệu đầu ra 22 hoặc thậm chí hủy bỏ việc giải mã khi tiếp cận khung của dạng phân tách biến đổi. Tuy nhiên, bộ giải mã truyền thống như vậy không có khả năng tạo ra sự khôi phục chính xác cho các khung biến đổi phân

tách, dẫn đến chất lượng bị giảm đi trong các khung chịu tác động so với việc giải mã bằng bộ giải mã thích hợp, ví dụ theo Fig. 1. Tuy nhiên, giả sử việc phân tách biến đổi được sử dụng như mong muốn, tức là chỉ trên đầu vào tạm thời hoặc rất nhiều ở tốc độ bit thấp, chất lượng thông qua bộ giải xHE-AAC mã sẽ tốt hơn nếu các khung bị tác động bị nhiều do việc chặn âm thanh hoặc gây ra các lỗi phát lại rõ rệt khác.

Cụ thể là, sự mở rộng của bộ lập mã/bộ giải mã xHE-AAC theo hướng phân tách biến đổi có thể như sau.

Theo phần mô tả trên đây, dụng cụ mới được sử dụng cho xHE-AAC có thể được gọi là phân tách biến đổi (transform splitting - TS). Đó sẽ là dụng cụ mới trong bộ lập mã miền tần số (frequency-domain - FD) của xHE-AAC hoặc, ví dụ, MPEG-H 3D-Audio được dựa trên USAC trong tài liệu tham khảo [4]. Sự phân tách biến đổi sau đó sẽ có thể sử dụng cho các đoạn tín hiệu tạm thời nhất định như phương án thay thế cho biến đổi dài đều đặn (mà dẫn đến nhiễu theo thời gian, đặc biệt là tiếng vang trước, ở tốc độ bit thấp) hoặc 8 biến đổi ngắn (mà dẫn đến các lỗ hổng quang phổ và các giả âm bong bóng ở tốc độ bit thấp). TS có thể sau đó được tín hiệu hóa một cách tương thích bán ngược bằng cách đan xen hệ số FD vào trong biến đổi dài mà có thể được phân tích chính xác bằng bộ giải mã MPEG-D USAC truyền thống.

Phần mô tả của công cụ này sẽ là tương tự với phần mô tả trên đây. Khi TS được kích hoạt trong biến đổi dài, 2 MDCT nửa độ dài được dùng thay cho một MDCT toàn bộ độ dài, và các hệ số của 2 MDCT, tức là 50 và 52, được truyền dẫn theo cách đan xen từng đường dẫn. Việc truyền dẫn đan xen đã được sử dụng, ví dụ, trong trường hợp biến đổi bắt đầu (dừng) FD, với các hệ số của MDCT đầu tiên theo thời gian được đặt tại các chỉ số chẵn và các hệ số của MDCT thứ hai theo thời gian được đặt tại các chỉ số lẻ (trong đó việc lập chỉ số bắt đầu từ 0), nhưng bộ giải mã không có khả năng xử lý các biến đổi dừng - bắt đầu sẽ không thể phân tích chính xác dòng dữ liệu. Nghĩa là, tùy theo các phạm vi khác nhau được sử dụng cho việc lập mã entropi các hệ số miền tần số dùng cho biến đổi dừng - bắt đầu này, cú pháp biến thiên được tạo dòng lên các biến đổi một nửa, bộ giải mã không có khả năng hỗ trợ các cửa sổ dừng - bắt đầu sẽ phải bỏ qua các khung cửa sổ dừng - bắt đầu tương ứng.

Tóm tắt ngắn gọn lại phương án được mô tả trên đây, điều này có nghĩa là bộ giải mã theo Fig.1 có thể, tùy theo mô tả trên đây, có khả năng, theo cách khác, hỗ trợ thêm độ dài biến đổi, tức là sự chia nhỏ hơn của các khung 26 nhất định thành thậm chí nhiều hơn 2 biến đổi bằng cách sử dụng sự tín hiệu hóa mà mở rộng sự tín hiệu hóa 34. Tuy nhiên, đối với sự kề nhau của các phần chia nhỏ của biến đổi của các khung 26, ngoài biến đổi phân tách được kích hoạt bằng cách sử dụng sự tín hiệu hóa 34, bộ chiết xuất hệ số FD 12 và bộ chiết xuất hệ số định tỷ lệ 14 sẽ nhảy với sự tín hiệu hóa này trong đó chế độ hoạt động của chúng sẽ thay đổi phụ thuộc vào sự tín hiệu hóa ngoài đó ngoài sự tín hiệu hóa 34. Hơn nữa, việc truyền dẫn theo dòng của các hệ số TNS, các tham số MS và các tham số dự báo phức hợp, được biến đổi theo dạng biến đổi được tín hiệu hóa ngoài dạng biến đổi phân tách theo 56 và 59, sẽ yêu cầu rằng từng bộ giải mã phải có khả năng đáp lại, tức là hiểu được, sự tín hiệu hóa bằng cách lựa chọn giữa “các dạng biến đổi đã biết” này hoặc các khung bao gồm dạng biến đổi dài theo 60, và các dạng biến đổi khác, ví dụ như dạng chia nhỏ các khung thành 8 biến đổi ngắn như trong trường hợp AAC. Trong trường hợp đó, “sự tín hiệu hóa đã biết” này sẽ nhận biết các khung mà sự tín hiệu hóa 34 tín hiệu hóa dạng biến đổi phân tách, như các khung có dạng biến đổi dài sao cho các bộ giải mã không có khả năng hiểu được sự tín hiệu hóa 34, xử lý các khung này như các khung biến đổi dài hơn là các khung có các dạng khác, như các khung dạng 8 biến đổi ngắn.

Quay lại việc mô tả của khả năng mở rộng của xHE-AAC, các giới hạn hoạt động nhất định có thể được cung cấp để gắn dụng cụ TS vào khuôn khổ lập mã này. Ví dụ, TS có thể được cho phép chỉ được sử dụng trong cửa sổ dài - bắt đầu hoặc ngừng - bắt đầu FD. Nghĩa là, chuỗi cửa sổ phản tử cú pháp cơ bản có thể được yêu cầu phải bằng 1. Ngoài ra, do sự tín hiệu hóa tương thích bán ngược, nó có thể là yêu cầu rằng, TS có thể chỉ được áp dụng khi việc điền đầy nhiều âm hệ số cú pháp là một hoạt động trong ngăn cú pháp `UsacCoreConfig()`. Khi TS được tín hiệu hóa để được kích hoạt, tất cả các dụng cụ FD ngoại trừ TNS và MDCT ngược hoạt động trên tập hợp đan xen (dài) của các hệ số TS. Điều này cho phép việc sử dụng lại độ lệch dài hệ số tỷ lệ và các bảng bộ lập mã số học biến đổi dài cũng như các hình dạng cửa sổ và các độ dài chồng lấp.

Dưới đây, các thuật ngữ và các định nghĩa được thể hiện mà được sử dụng sau đây để giải thích cách mà tiêu chuẩn USAC được mô tả trong tài liệu tham khảo [4] có thể được mở rộng để đưa ra chức năng TS tương thích ngược, trong đó đôi khi sự tham chiếu được thực hiện cho các phần trong tiêu chuẩn đó cho độ giả quan tâm.

Mục dữ liệu mới có thể là:

biến đổi phân tách dấu hiệu nhị phân biểu thị TS có được sử dụng trong khung và kênh hiện thời hay không

Các mục hỗ trợ mới có thể là:

chuỗi cửa sổ dạng chuỗi cửa sổ FD đối với khung và kênh hiện thời (phần 6.2.9)

độ lệch nhiều âm độ lệch điền đầy nhiều âm để biến đổi các hệ số tỷ lệ của các dải được lượng tử hóa tới 0 (phần 7.2)

mức nhiều âm mức điền đầy nhiều âm thể hiện biên độ của nhiều âm quang phổ bổ sung (phần 7.2)

độ dài biến đổi một nửa một nửa của `coreCoderFrameLength` (`ccfl`, độ dài biến đổi, phần 6.1.1)

đường dẫn thông thấp một nửa một nửa của số lượng các đường dẫn MDCT được truyền đối với kênh hiện thời.

Việc giải mã của biến đổi bắt đầu (ngừng-) FD bằng cách sử dụng sự phân tách biến đổi (TS) trong khuôn khổ USAC có thể được thực hiện theo các bước hoàn toàn kế tiếp nhau, như sau:

Thứ nhất, việc giải mã chuyển đổi phân tách và đường dẫn thông thấp một nửa có thể được thực hiện.

biến đổi phân tách thực tế sẽ không thể hiện hệ số dòng bit độc lập mà được suy ra từ hệ số điền đầy nhiều âm, độ lệch nhiều âm và mức nhiều âm, và trong trường hợp `UsacChannelPairElement()`, chỉ số cửa sổ thông thường trong `StereoCoreToolInfo()`. Nếu `noiseFilling == 0`, `split_transform` bằng 0. Mặt khác,

```

nếu (noiseFilling != 0) && (noise_level == 0)) {
    split_transform = (noise_offset & 16) / 16;
    noise_level = (noise_offset & 14) / 2;
    noise_offset = (noise_offset & 1) * 16;
}
khác {
    split_transform = 0;;
}

```

Nói cách khác, nếu `noise_level == 0`, `noise_offset` bao gồm dấu hiệu `split_transform` được theo sau bởi 4 bit dữ liệu điền đầy nhiễu âm, mà sau đó được sắp xếp lại. Do sự hoạt động này làm thay đổi các trị số của `noise_level` và `noise_offset`, việc này phải được thực hiện trước quá trình điền đầy nhiễu âm của phần 7.2. Hơn nữa, nếu `common_window == 1` trong `UsacChannelPairElement()`, `split_transform` được xác định chỉ trong kênh (thứ nhất) bên trái; `split_transform` của kênh bên phải được thiết lập bằng với (tức là được sao chép từ) `split_transform` của kênh trái, và mã giả trên đây không được thực hiện trong kênh bên phải.

`half_lowpass_line` được xác định từ bảng kết thúc dải hệ số tỷ lệ "dài", `swb_offset_long_window`, và `max_sfb` của kênh hiện thời, hoặc trong trường hợp âm thanh nổi và `common_window == 1`, `max_sfb_ste`.

```

max_sfb_ste trong hệ số với StereoCoreToolInfo() và common_window == 1,
lowpass_sfb =

```

`max_sfb` theo cách khác. Dựa trên dấu hiệu `igFilling`, `half_lowpass_line` được suy ra:

```

nếu (igFilling != 0) {
    lowpass_sfb = max(lowpass_sfb, ig_stop_sfb);
}
half_lowpass_line = swb_offset_long_window[lowpass_sfb] / 2;

```

Sau đó, như là bước thứ hai, việc khử đan xen của phổ nửa độ dài để định hình nhiễu âm theo thời gian sẽ được thực hiện.

Sau khi giải lượng tử hóa quang phổ, điền đầy nhiễu âm, và áp dụng hệ số tỷ lệ và trước khi áp dụng định hình nhiễu âm theo thời gian (Temporal Noise Shaping - TNS), các hệ số TS trong spec[] được khử đan xen bằng cách sử dụng đệm trợ giúp[]:

```

đối với (i = 0, i2 = 0; i < half_lowpass_line; i += 1, i2 += 2) {
    spec[i] = spec[i2];          /* isolate 1st window */
    buffer[i] = spec[i2+1];      /* isolate 2nd window */
}

đối với (i = 0; i < half_lowpass_line; i += 1) {
    spec[i+half_lowpass_line] = buffer[i]; /* copy 2nd window */
}

```

Việc khử đan xen tại chỗ đặt một cách hiệu quả 2 phổ TS nửa độ dài trên đỉnh của nhau, và

dụng cụ TNS giờ đây hoạt động như thông thường trên phổ giả toàn bộ độ dài tạo thành.

Tham khảo đến phần trên đây, quy trình như vậy được mô tả đối với Fig.3.

Sau đó, như bước thứ ba, việc đan xen lại theo thời gian vào sẽ được sử dụng cùng với với 2 MDCT ngược kế tiếp nhau.

Nếu common_window == 1 trong khung hiện thời hoặc việc giải mã âm thanh nổi được thực hiện sau việc giải mã TNS (tns_on_lr == 0 trong phần 7.8), spec[] phải được đan xen lại theo thời gian vào trong phổ toàn bộ độ dài:

```

đối với (i = 0; i < half_lowpass_line; i += 1) {
    buffer[i] = spec[i]; /* copy 1st window */
}

đối với (i = 0, i2 = 0; i < half_lowpass_line; i += 1, i2 += 2) {

```

```

spec[i2] = buffer[i];                                /* merge 1st window */
spec[i2+1] = spec[i+half_lowpass_line]; /* merge 2nd window */
}

```

Phổ giả tạo thành được sử dụng để giải mã âm thanh nổi (phần 7.7) và để cập nhật `dmx_re_prev[ ]`

(các phần 7.7.2 và A.1.4). Trong trường hợp `tns_on_lr == 0`, các phổ toàn bộ độ dài được giải mã âm thanh nổi một lần nữa

được khử đan xen bằng cách lặp lại quy trình của phần A.1.3.2. Cuối cùng, 2 MDCT ngược được tính toán.

với `ccfl` và hình dạng cửa sổ của kênh của khung hiện thời và cuối cùng. Xem phần 7.9 và Fig.1.

Một số biến đổi có thể được tạo ra cho việc giải mã âm thanh nổi dự báo phức hợp của xHE-AAC.

Phương pháp tín hiệu hóa tương thích bán ngược ẩn có thể được sử dụng theo cách khác để gắn TS vào trong xHE-AAC.

Phần trên mô tả phương pháp sử dụng một bit trong dòng bit để tín hiệu hóa sự sử dụng của việc phân tách biến đổi theo sáng chế, có mặt trong chuyển đổi phân tách, cho bộ giải mã theo sáng chế. Cụ thể là, việc tín hiệu hóa này (hãy gọi nó là việc tín hiệu hóa tương thích bán ngược ẩn) cho phép dữ liệu dòng bit truyền thống sau đây - ở đây là thông tin phụ điền đầy nhiễu âm được sử dụng độc lập với tín hiệu theo sáng chế: Trong phương án này, dữ liệu điền đầy nhiễu âm không phụ thuộc vào dữ liệu phân tách biến đổi, và ngược lại. Ví dụ, dữ liệu điền đầy nhiễu âm bao gồm tất cả giá trị 0 (`noise_level = noise_offset` được thiết lập = 0) có thể được truyền trong khi chuyển đổi phân tách có thể giữ trị số bất kỳ (là chỉ báo nhị phân, hoặc là 0 hoặc 1).

Trong các trường hợp mà tính độc lập giới hạn này giữa dữ liệu dòng bit truyền thống và dữ liệu dòng bit theo sáng chế không được yêu cầu và tín hiệu theo sáng chế là quyết định nhị phân, việc truyền ẩn bit tín hiệu hóa có thể tránh được, và quyết định nhị phân này có thể được tín hiệu hóa bởi sự có mặt hoặc vắng mặt của cái có thể được

gọi là tín hiệu hóa tương thích bán ngược ản. Xét đến phương án trên đây một lần nữa làm ví dụ, việc sử dụng để phân tách biến đổi có thể được truyền đơn giản bằng cách sử dụng việc tín hiệu hóa theo sáng chế: Nếu `noise_level` bằng 0 và, đồng thời, `noise_offset` khác 0, khi đó `split_transform` được thiết lập bằng 1. Nếu cả hai `noise_level` và `noise_offset` khác 0, `split_transform` được thiết lập bằng 0. Sự phụ thuộc của tín hiệu ản theo sáng chế vào tín hiệu điền đầy nhiễu âm truyền thống tăng lên khi cả hai `noise_level` và `noise_offset` bằng 0. Trong trường hợp này, không rõ ràng liệu việc tín hiệu hóa truyền thống hoặc tín hiệu hóa ản theo sáng chế sẽ được sử dụng. Để tránh sự mơ hồ này, giá trị của `split_transform` phải được xác định trước. Trong ví dụ này, thích hợp để xác định `split_transform` = 0 nếu dữ liệu điền đầy nhiễu âm bao gồm tất cả mức 0, do đây là điều mà bộ mã hóa truyền thống không phân tách biến đổi sẽ tín hiệu hóa khi sự điền đầy nhiễu âm không được sử dụng trong khung.

Vấn đề mà còn lại cần phải giải quyết là trong trường hợp tín hiệu hóa tương thích bán ngược ản là cách để tín hiệu hóa `split_transform` == 1 và không điền đầy nhiễu âm đồng thời. Như đã được giải thích, dữ liệu điền đầy nhiễu âm không được tất cả bằng 0, và nếu độ lớn nhiễu bằng 0 được yêu cầu, `noise_level` ((`noise_offset` & 14)/2 như trên đây) phải bằng 0. Điều này khiến cho chỉ ((`noise_offset` & 1)*16 như trên đây) lớn hơn 0 là độ phân giải. May là, trị số của `noise_offset` được bỏ qua nếu sự điền đầy nhiễu âm không được thực hiện trong bộ giải mã được dựa trên USAC trong tài liệu tham khảo [4], do đó phương pháp này trở thành khả thi trong phương án này. Do đó, việc tín hiệu hóa `split_transform` trong lập mã giả như trên đây có thể được biến đổi như sau, bằng cách sử dụng bit tín hiệu hóa TS được lưu để truyền 2 bit (4 trị số) thay cho 1 bit cho `noise_offset`:

```

nếu ((noiseFilling != 0) && (noise_level == 0) && (noise_offset set != 0)) {
    split_transform = 1;
    noise_level = (noise_offset set & 28) / 4;
    noise_offset set = (noise_offset set & 3) * 8;
}
khác {

```

```
split_transform = 0;
}
```

Do đó, bằng cách áp dụng phương án này, việc mô tả USAC có thể được mở rộng bằng cách sử dụng phần mô tả sau.

Sự mô tả công cụ sẽ hầu như tương tự. Đó là,

Khi việc phân tách biến đổi (transform splitting - TS) được kích hoạt trong biến đổi dài, 2 MDCT nửa độ dài được dùng thay cho một MDCT toàn bộ độ dài. Các hệ số của 2 MDCT được truyền theo cách đan xen từng đường như biến đổi miền tần số (frequency domain - FD) truyền thống, với các hệ số của MDCT đầu tiên theo thời gian được đặt tại các chỉ số chẵn và các hệ số của MDCT thứ hai theo thời gian được đặt tại các chỉ số lẻ.

Các giới hạn hoạt động có thể bắt buộc vì TS chỉ có thể được sử dụng trong cửa sổ FD dài - bắt đầu hoặc ngừng - bắt đầu (`window_sequence == 1`) và TS chỉ có thể được áp dụng khi `noiseFilling` bằng 1 trong `UsacCoreConfig( )`. Khi TS được tín hiệu hóa, tất cả các công cụ FD ngoại trừ TNS và MDCT ngược hoạt động trên tập hợp đan xen (dài) của các hệ số TS. Điều này cho phép việc sử dụng lại độ lệch nhiễu âm dải hệ số tỷ lệ và các bảng bộ lập mã số học biến đổi dài cũng như các dạng cửa sổ và các độ dài chồng lấp.

Các thuật ngữ và các định nghĩa được sử dụng sau đây liên quan tới các Hệ số Trợ giúp sau đây

<code>common_window</code>	biểu thị nếu kênh 0 và kênh 1 của CPE sử dụng các tham số cửa sổ đồng dạng (tham khảo ISO/IEC 23003-3:2012 phần 6.2.5.1.1).
<code>window_sequence</code>	dạng chuỗi cửa sổ FD đối với khung và kênh hiện thời (tham khảo ISO/IEC 23003-3:2012 phần 6.2.9).
<code>tns_on_lr</code>	Biểu thị chế độ hoạt động cho việc lọc TNS (tham khảo ISO/IEC 23003-3:2012 phần 7.8.2).

noiseFilling	Dấu hiệu này tín hiệu hóa sự sử dụng của việc điền đầy nhiễu âm của các lỗ hổng quang phổ trong bộ lập mã nhân FD (tham khảo ISO/IEC 23003-3:2012 phần 6.1.1.1).
noise_offset	độ lệch điền đầy nhiễu âm để biến đổi các hệ số tỷ lệ của các dải được lượng tử hóa tới 0 (tham khảo ISO/IEC 23003-3:2012 phần 7.2).
noise_level	mức điền đầy nhiễu âm thể hiện độ lớn của nhiễu âm quang phổ bổ sung (tham khảo ISO/IEC 23003-3:2012 phần 7.2).
split_transform	dấu hiệu nhị phân biểu thị TS có được sử dụng trong khung và kênh hiện thời hay không.
half_transform_length	một nửa của coreCoderFrameLength (ccfl, độ dài biến đổi, tham khảo ISO/IEC 23003-3:2012 phần 6.1.1).
half_lowpass_line	một nửa của số lượng các đường dẫn MDCT được truyền đối với kênh hiện thời.

Quy trình giải mã liên quan tới TS có thể được mô tả như sau. Cụ thể là, việc giải mã sự biến đầu bắt đầu FD (dùng-) với TS được thực hiện trong 3 bước kế tiếp nhau như sau.

Trước tiên, việc giải mã split_transform và half_lowpass_line được thực hiện. Hệ số trợ giúp split_transform không thể hiện hệ số dòng bit độc lập mà được suy ra từ các hệ số điền đầy nhiễu âm, noise_offset và noise_level, và trong trường hợp UsacChannelPairElement(), dấu hiệu common_window trong stereoCoreToolInfo(). Nếu nạp nhiễu == 0, split_transform là 0. Theo cách khác,

```

nếu ((noiseFilling != 0) && (noise_level == 0)) {
    split_transform = 1;
    noise_level = (noise_offset set & 28) / 4;
    noise_offset được cài đặt = (noise_offset set & 3) * 8;

```

```

}

khác {

    split_transform = 0;

}

```

Nói cách khác, nếu `noise_level == 0`, `noise_offset` bao gồm dấu hiệu `split_transform` được theo sau bởi 4 bit của dữ liệu điền đầy nhiễu âm, mà sau đó được sắp xếp lại. Do hoạt động này làm thay đổi các giá trị của `noise_level` và `noise_offset` được cài đặt, hoạt động này phải được thực hiện trước quy trình điền đầy nhiễu âm của ISO/IEC 23003-3:2012 phần 7.2.

Hơn nữa, nếu `common_window == 1` trong `UsacChannelPairElement()`, `split_transform` chỉ được xác định trong kênh (thứ nhất) bên trái; `split_transform` của kênh bên phải được thiết lập bằng với (tức là được sao chép từ) `split_transform` của kênh trái, và sự lặp mã giả trên đây không được thực hiện trong kênh bên phải.

Hệ số trợ giúp `half_lowpass_line` được xác định từ bảng độ lệch nhiễu âm dải hệ số tỷ lệ “dài”, `swb_offset_long_window`, và `max_sfb` của kênh hiện thời, hoặc trong trường hợp âm thanh nổi và `common_window == 1`, `max_sfb_ste`.

$$\text{lowpass_sfb} = \begin{cases} \text{max_sfb_ste} & \text{in elements with StereoCoreToolInfo() and } \mathbf{common_window} == \\ \text{max_sfb} & \text{otherwise.} \end{cases}$$

Dựa trên dấu hiệu `igFilling`, `half_lowpass_line` được suy ra:

```

nếu (igFilling != 0) {

    lowpass_sfb = max(lowpass_sfb, ig_stop_sfb);

}

```

```

half_lowpass_line = swb_offset_long_window[lowpass_sfb] / 2;

```

Sau đó, việc khử đan xen của các phổ nửa độ dài để định hình nhiễu theo thời gian được thực hiện.

Sau khi giải lượng tử hóa phổ, điền đầy nhiễu âm, và áp dụng hệ số tỷ lệ và trước khi áp dụng định hình nhiễu âm theo thời gian (temporal noise shaping - TNS), các hệ số TS trong `spec[ ]` được khử đan xen bằng cách sử dụng đệm trợ giúp `[ ]`:

```

đối với (i = 0, i2 = 0; i < half_lowpass_line; i += 1, i2 += 2) {
    spec[i] = spec[i2]; /* isolate 1st window */
    buffer[i] = spec[i2+1]; /* isolate 2nd window */
}

đối với (i = 0; i < half_lowpass_line; i += 1) {
    spec[i+half_lowpass_line] = buffer[i]; /* copy 2nd window */
}

```

Việc khử đan xen tại chỗ đặt một cách hiệu quả 2 phổ TS nửa độ dài trên đỉnh của nhau, và công cụ TNS giờ đây hoạt động như thông thường trên phổ giả toàn bộ độ dài tạo thành.

Cuối cùng, việc đan xen lại theo thời gian và hai MDCT ngược kế tiếp nhau có thể được sử dụng:

Nếu `common_window == 1` trong khung hiện thời hoặc việc giải mã âm thanh nổi được thực hiện sau khi giải mã TNS (`tns_on_lr == 0` trong phần 7.8), `spec[ ]` phải được đan xen lại theo thời gian vào trong phổ toàn bộ độ dài:

```

đối với (i = 0; i < half_lowpass_line; i += 1) {
    buffer[i] = spec[i]; /* copy 1st window */
}

đối với (i = 0, i2 = 0; i < half_lowpass_line; i += 1, i2 += 2) {
    spec[i2] = buffer[i]; /* merge 1st window */
    spec[i2+1] = spec[i+half_lowpass_line]; /* merge 2nd window */
}

```

Phổ giả tạo thành được sử dụng để giải mã âm thanh nổi (ISO/IEC 23003-3:2012 phần 7.7) và để cập nhật $dmx_re_prev[]$ (ISO/IEC 23003-3:2012 phần 7.7.2) và trong trường hợp $tns_on_lr == 0$, các phổ toàn bộ độ dài âm thanh nổi được khử đan xen một lần nữa bằng cách lặp lại quy trình của phần. Cuối cùng, 2 MDCT ngược được tính toán với $ccfl$ và hình dạng cửa sổ của kênh của khung hiện thời và cuối cùng.

Việc xử lý đối với TS tuân thủ phần mô tả đã nêu trong ISO/IEC 23003-3:2012 phần "7.9 Dàn lọc và chuyển đổi khối." Phần bổ sung sau đây nên được tính đến.

Các hệ số TS trong $spec[]$ được khử đan xen bằng cách sử dụng đệm trợ giúp[] với N , độ dài cửa sổ được dựa trên trị số chuỗi cửa sổ:

```

đối với (i = 0, i2 = 0; i < N/2; i += 1, i2 += 2) {
    spec[0][i] = spec[i2]; /* isolate 1st window */
    buffer[i] = spec[i2+1]; /* isolate 2nd window */
}
đối với (i = 0; i < N/2; i += 1) {
    spec[1][i] = buffer[i]; /* copy 2nd window */
}

```

IMDCT đối với phổ TS nửa độ dài khi đó được xác định là:

$$x_{(0,1),n} = \frac{2}{N} \sum_{k=0}^{\frac{N}{4}-1} spec[(0,1)][k] \cos\left(\frac{4\pi}{N}(n+n_0)\left(k+\frac{1}{2}\right)\right) \quad \text{for } 0 \leq n < \frac{N}{2}$$

Việc tạo cửa sổ tiếp theo và các bước chuyển đổi khối được xác định trong các phần nhỏ sau.

Việc phân tách biến đổi với STOP_START_SEQUENCE sẽ giống như phần mô tả sau:

STOP_START_SEQUENCE kết hợp với sự phân tách biến đổi được mô tả trên Fig.2. Chuỗi này bao gồm 2 cửa sổ nửa độ dài 56, 58 bổ sung và chồng lấp với độ dài $N_{1/2}$ bằng 1024 (960, 768). N_s được thiết lập bằng 256 (240, 192) tương ứng.

Các cửa sổ (0,1) đối với 2 IMDCT nửa độ dài được đưa ra như sau:

$$W_{(0,1)}(n) = \begin{cases} 0.0, & \text{for } 0 \leq n < \frac{N-l/2-N_s}{4} \\ W_{(0,1),LEFT,N_s}(n - \frac{N-l/2-N_s}{4}), & \text{for } \frac{N-l/2-N_s}{4} \leq n < \frac{N-l/2+N_s}{4} \\ 1.0, & \text{for } \frac{N-l/2+N_s}{4} \leq n < \frac{3N-l/2-N_s}{4} \\ W_{(0,1),RIGHT,N_s}(n + \frac{N_s}{2} - \frac{3N-l/2-N_s}{4}), & \text{for } \frac{3N-l/2-N_s}{4} \leq n < \frac{3N-l/2+N_s}{4} \\ 0.0, & \text{for } \frac{3N-l/2+N_s}{4} \leq n < N-l/2 \end{cases}$$

trong đó đối với các cửa sổ IMDCT thứ nhất

$$W_{0,LEFT,N_s}(n) = \begin{cases} W_{KBD_LEFT,N_s}(n), & \text{if window_shape_previous_block} == 1 \\ W_{SIN_LEFT,N_s}(n), & \text{if window_shape_previous_block} == 0 \end{cases},$$

$$W_{0,RIGHT,N_s}(n) = \begin{cases} W_{KBD_RIGHT,N_s}(n), & \text{if window_shape} == 1 \\ W_{SIN_RIGHT,N_s}(n), & \text{if window_shape} == 0 \end{cases}$$

được áp dụng và đối với các cửa sổ IMDCT thứ hai

$$W_{1,LEFT,N_s}(n) = \begin{cases} W_{KBD_LEFT,N_s}(n), & \text{if window_shape} == 1 \\ W_{SIN_LEFT,N_s}(n), & \text{if window_shape} == 0 \end{cases},$$

$$W_{1,RIGHT,N_s}(n) = \begin{cases} W_{KBD_RIGHT,N_s}(n), & \text{if window_shape} == 1 \\ W_{SIN_RIGHT,N_s}(n), & \text{if window_shape} == 0 \end{cases}$$

được áp dụng.

Sự chồng lấp và bổ sung giữa 2 cửa sổ nửa độ dài tạo ra các trị số miền thời gian được tạo cửa sổ $z_{i,n}$ được mô tả như sau. Ở đây, N_l được thiết lập bằng 2048 (1920, 1536), N_s bằng 256 (240, 192) tương ứng:

$$Z_{i,n}(n) = \begin{cases} 0.0, & \text{for } 0 \leq n < N_s \\ x_{0,n-N_s} \cdot W_0(n - N_s), & \text{for } N_s \leq n < \frac{2N-l-N_s}{4} \\ x_{0,n-N_s} \cdot W_0(n - N_s) + x_{1,n-(N-l/2-N_s)} \cdot W_1(n - (N-l/2 - N_s)), & \text{for } \frac{N-l+N_s}{4} \leq n < \frac{2N-l+N_s}{4} \\ x_{1,n-(N-l/2-N_s)} \cdot W_1(n - (N-l/2 - N_s)), & \text{for } \frac{2N-l+N_s}{4} \leq n < N \end{cases}$$

Sự Phân tách Biến đổi với LONG_START_SEQUENCE sẽ giống như phần mô tả sau:

LONG_START_SEQUENCE kết hợp với sự phân tách biến đổi được mô tả trên Fig.4. Chuỗi này bao gồm 3 cửa sổ được xác định như sau, trong đó $N_l/2$ được thiết lập bằng 1024 (960, 768), N_s được thiết lập bằng 256 (240, 192) tương ứng.

$$W_0(n) = \begin{cases} 1.0, & \text{for } 0 \leq n < \frac{3N_l/2 - N_s}{4} \\ W_{0,RIGHT,N_s}(n + \frac{N_s}{2} - \frac{3N_l/2 - N_s}{4}), & \text{for } \frac{3N_l/2 - N_s}{4} \leq n < \frac{3N_l/2 + N_s}{4} \\ 0.0, & \text{for } \frac{3N_l/2 + N_s}{4} \leq n < N_l/2 \end{cases}$$

$$W_1(n) = \begin{cases} 0.0, & \text{for } 0 \leq n < \frac{N_l/2 - N_s}{4} \\ W_{1,LEFT,N_s}(n - \frac{N_l/2 - N_s}{4}), & \text{for } \frac{N_l/2 - N_s}{4} \leq n < \frac{N_l/2 + N_s}{4} \\ 1.0, & \text{for } \frac{N_l/2 + N_s}{4} \leq n < \frac{3N_l/2 - N_s}{4} \\ W_{1,RIGHT,N_s}(n + \frac{N_s}{2} - \frac{3N_l/2 - N_s}{4}), & \text{for } \frac{3N_l/2 - N_s}{4} \leq n < \frac{3N_l/2 + N_s}{4} \\ 0.0, & \text{for } \frac{3N_l/2 + N_s}{4} \leq n < N_l/2 \end{cases}$$

Các nửa cửa sổ bên trái/bên phải được đưa ra bởi:

$$W_{1,LEFT,N_s}(n) = \begin{cases} W_{KBD_LEFT,N_s}(n), & \text{if window_shape} == 1 \\ W_{SIN_LEFT,N_s}(n), & \text{if window_shape} == 0 \end{cases},$$

$$W_{(0,1),RIGHT,N_s}(n) = \begin{cases} W_{KBD_RIGHT,N_s}(n), & \text{if window_shape} == 1 \\ W_{SIN_RIGHT,N_s}(n), & \text{if window_shape} == 0 \end{cases}$$

Cửa sổ thứ ba bằng một nửa trái của LONG_START_WINDOW:

$$W_2(n) = \begin{cases} W_{LEFT,N_l}(n), & \text{for } 0 \leq n < N_l/2 \\ 1.0, & \text{for } N_l/2 \leq n < N_l \end{cases}$$

với

$$W_{LEFT,N_l}(n) = \begin{cases} W_{KBD_LEFT,N_l}(n), & \text{if window_shape_previous_block} == 1 \\ W_{SIN_LEFT,N_l}(n), & \text{if window_shape_previous_block} == 0 \end{cases}$$

Sự chồng lấp và bổ sung giữa 2 cửa sổ nửa độ dài tạo ra các trị số miền thời gian được tạo cửa sổ trung gian $\tilde{Z}_{i,n}$ được mô tả như sau. Ở đây, N_l được thiết lập bằng 2048 (1920, 1536), N_s đến 256 (240, 192) tương ứng.

$$\tilde{Z}_{i,n}(n) = \begin{cases} -1 \cdot x_{0,2N_s-n-1} \cdot W_0(2N_s-n-1), & \text{for } 0 \leq n < N_s \\ x_{0,n-N_s} \cdot W_0(n-N_s), & \text{for } N_s \leq n < \frac{2N_l-N_s}{4} \\ x_{0,n-N_s} \cdot W_0(n-N_s) + x_{1,n-(N_l/2-N_s)} \cdot W_1(n-(N_l/2-N_s)), & \text{for } \frac{N_l+N_s}{4} \leq n < \frac{2N_l+N_s}{4} \\ x_{1,n-(N_l/2-N_s)} \cdot W_1(n-(N_l/2-N_s)), & \text{for } \frac{2N_l+N_s}{4} \leq n < N \end{cases}$$

Các trị số miền thời gian được tạo cửa sổ cuối cùng $Z_{i,n}$ thu được bằng cách áp dụng W_2 :

$$Z_{i,n}(n) = \tilde{Z}_{i,n}(n) \cdot W_2(n), \quad \text{for } 0 \leq n < N_l$$

Bất kể việc tín hiệu hóa tương thích bán ngược rõ ràng hoặc ẩn được sử dụng hay không, mà cả hai quá trình này được mô tả trên đây, một số biến đổi có thể là cần thiết cho việc giải mã âm thanh nổi dự báo phức hợp của xHE-AAC để thu được sự hoạt động có ý nghĩa trên các phổ được đan xen.

Sự biến đổi trong việc giải mã âm thanh nổi dự báo phức hợp có thể được áp dụng như sau.

Do các công cụ âm thanh nổi FD hoạt động trên phổ giả được đan xen khi TS hoạt động trong cặp kênh, sự thay đổi là không cần thiết cho việc xử lý Sự dự báo Phức hợp hoặc M/S cơ bản. Tuy nhiên, sự suy ra của trộn giảm $dmx_re_prev[]$ của khung trước đó và sự tính toán trộn giảm MDST $dmx_im[]$ trong ISO/IEC 23003-3:2012 phần 7.7.2 cần được làm thích ứng nếu TS được sử dụng trong kênh của khung cuối cùng hoặc khung hiện thời:

- use_prev_frame phải bằng 0 nếu mức độ hoạt động TS được thay đổi trong kênh từ khung cuối cùng đến khung hiện thời. Nói cách khác, $dmx_re_prev[]$ phải không được sử dụng trong trường hợp đó do sự chuyển đổi độ dài biến đổi.

- Nếu TS đã được kích hoạt, `dmx_re_prev[ ]` và `dmx_re[ ]` đặc trưng hóa các phổ giả đan xen và phải được khử đan xen thành 2 phổ nửa độ dài TS tương ứng của chúng để tính toán MDST chính xác.
- Tùy thuộc vào mức độ hoạt động TS, 2 trộn giảm MDST nửa độ dài được tính toán bằng cách sử dụng các hệ số bộ lọc được làm thích ứng (các Bảng 1 và 2) và được đan xen vào trong phổ toàn bộ độ dài `dmx_im[ ]` (giống như `dmx_re[ ]`).
- `window_sequence`: Các ước lượng trộn giảm MDST được tính toán đối với từng nhóm cặp cửa sổ. `use_prev_frame` được đánh giá chỉ đối với cặp đầu tiên của hai cặp cửa sổ một nửa. Đối với cặp cửa sổ còn lại, cặp cửa sổ đi trước luôn luôn được sử dụng trong các ước lượng MDST, mà ngụ ý `use_prev_frame = 1`.
- Các dạng cửa sổ: Các thông số ước lượng MDST đối với cửa sổ hiện thời, mà là các hệ số bộ lọc như được mô tả dưới đây, phụ thuộc vào hình dạng của các nửa cửa sổ bên trái và bên phải. Đối với cửa sổ thứ nhất, điều này có nghĩa là các thông số bộ lọc là chức năng của các dấu hiệu hình dạng cửa sổ của khung hiện thời và khung trước đó. Cửa sổ còn lại chỉ chịu tác động bởi hình dạng cửa sổ hiện thời.

Bảng 1 – Các Tham số Bộ lọc MDST đối với Cửa sổ Hiện thời (`filter_coefs`)

Chuỗi Cửa sổ Hiện thời	Nửa Trái: Dạng Sin Nửa Phải: Dạng Sin	Nửa Trái: Dạng KBD Nửa Phải: Dạng KBD
LONG_START_SEQUENCE STOP_START_SEQUENCE	[0,185618f, -0,000000f, 0,627371f, 0,000000f, -0,627371f, 0,000000f, - 0,185618f]	[0,203599f, -0,000000f, 0,633701f, 0,000000f, -0,633701f, 0,000000f, - 0,203599f]
Chuỗi Cửa sổ Hiện thời	Nửa Trái: Dạng sin Nửa Phải: dạng KBD	Nửa Trái: dạng KBD Nửa Phải: Dạng sin
LONG_START_SEQUENCE STOP_START_SEQUENCE	[0,194609f, 0,006202f, 0,630536f, 0,000000f, -0,630536f, -0,006202f, - 0,194609f]	[0,194609f, -0,006202f, 0,630536f, 0,000000f, -0,630536f, 0,006202f, - 0,194609f]

Bảng 2 - MDST Các thông số lọc đối với Cửa sổ có trước (filter_coefs_prev)

Chuỗi Cửa sổ Hiện thời	Nửa trái của cửa sổ Hiện thời: Dạng Sin	Nửa trái của cửa sổ Hiện thời: Dạng KBD
LONG_START_SEQUENCE STOP_START_SEQUENCE	[0,038498, 0,039212, 0,039645, 0,039790, 0,039645, 0,039212, 0,038498]	[0,038498, 0,039212, 0,039645, 0,039790, 0,039645, 0,039212, 0,038498]

Cuối cùng, Fig.5 thể hiện, để hoàn thiện, bộ giải mã âm thanh miền tần số có thể hỗ trợ chuyển việc đổi độ dài biến đổi phù hợp với các phương án được nêu ra trên đây. Đó là, bộ mã hóa theo Fig.5 mà nhìn chung được biểu thị bằng cách sử dụng số tham chiếu 100 có khả năng mã hóa tín hiệu âm thanh 102 vào trong dòng dữ liệu 20 theo cách sao cho bộ giải mã theo Fig.1 và các biến thể tương ứng được mô tả trên đây có khả năng tận dụng khuôn phân tách biến đổi đối với một số khung, trong khi các bộ giải mã “loại cũ” vẫn có khả năng xử lý các khung TS mà không phân tích các lỗi hoặc tương tự.

Bộ mã hóa 100 theo Fig.5 bao gồm bộ biến đổi 104, bộ định tỷ lệ ngược 106, bộ chèn hệ số miền tần số 108 và bộ chèn vào hệ số tỷ lệ 110. Bộ biến đổi 104 nhận tín hiệu âm thanh 102 để mã hóa và được tạo cấu hình để đưa các phần miền thời gian của tín hiệu âm thanh trải qua sự biến đổi để thu được các hệ số miền tần số cho các khung của tín hiệu âm thanh. Cụ thể là, như rõ hơn từ phần trình bày trên đây, bộ biến đổi 104 quyết định trên cơ sở từng khung một mà trong đó phần chia nhỏ hơn của các khung 26 này vào trong các biến đổi - hoặc các cửa sổ biến đổi - được sử dụng. Như được mô tả trên đây, các khung 26 có thể có độ dài bằng nhau và sự biến đổi có thể là biến đổi chồng bằng cách sử dụng các biến đổi chồng lấp có độ dài khác nhau. Fig.5 minh họa rằng, ví dụ, khung 26a phải trải qua một biến đổi dài, khung 26b phải trải qua sự phân tách biến đổi, tức là 2 biến đổi của một nửa độ dài, và khung 26c được thể hiện phải trải qua nhiều hơn 2, tức là $2^n > 2$, thậm chí các biến đổi ngắn hơn của 2^{-n} độ dài biến đổi dài. Như được mô tả trên đây, nhờ cách này, bộ mã hóa 100 có khả năng tương thích với phân giải quang phổ theo thời gian của ảnh phổ được thể hiện bởi sự

biến đổi chồng được thực hiện bởi bộ biến đổi 104 cho nội dung âm thanh biến thiên theo thời gian hoặc dạng nội dung âm thanh của tín hiệu âm thanh 102.

Nghĩa là, các hệ số miền tần số được tạo ra tại đầu ra của bộ biến đổi 104 biểu diễn cho ảnh phổ của tín hiệu âm thanh 102. Bộ định tỷ lệ ngược 106 được kết nối với đầu ra của bộ biến đổi 104 và được tạo cấu hình để định tỷ lệ ngược, và lượng tử hóa một cách đồng thời, các hệ số miền tần số theo các hệ số tỷ lệ. Đặc biệt là, bộ định tỷ lệ ngược hoạt động trên các hệ số tần số mà các hệ số này thu được bởi bộ biến đổi 104. Nghĩa là, bộ định tỷ lệ ngược 106 phải, cần thiết, nhận thức về sự chuyển đổi độ dài biến đổi hoặc sự chuyển đổi cách thức biến đổi cho các khung 26. Cần lưu ý rằng, bộ định tỷ lệ ngược 106 cần thiết để xác định các hệ số tỷ lệ. Bộ định tỷ lệ ngược 106 là, cho mục đích này, ví dụ, phần của vòng phản hồi mà đánh giá ngưỡng che giấu tâm thính học được xác định đối với tín hiệu âm thanh 102 để nhiều âm sự lượng tử hóa được giới thiệu bởi sự lượng tử hóa và dần dần được thiết lập theo các hệ số tỷ lệ, dưới ngưỡng tâm thính học của sự phát hiện khi mà tuân thủ hoặc không tuân thủ một số giới hạn bit.

Tại đầu ra của bộ định tỷ lệ ngược 106, các hệ số tỷ lệ và các hệ số miền tần số được định tỷ lệ ngược và được lượng tử hóa là đầu ra và bộ chèn hệ số tỷ lệ 110 được tạo cấu hình để chèn các hệ số tỷ lệ vào dòng dữ liệu 20, trong khi bộ chèn hệ số miền tần số 108 được tạo cấu hình để chèn các hệ số miền tần số của các khung của tín hiệu âm thanh, được định tỷ lệ ngược và được lượng tử hóa theo các hệ số tỷ lệ, vào dòng dữ liệu 20. Theo cách tương ứng với bộ giải mã, cả bộ chèn 108 và 110 hoạt động bất kể cách thức biến đổi liên quan với các khung 26 trong phạm vi sự kề nhau của các khung 26a theo cách thức biến đổi dài và các khung 26b theo cách thức biến đổi phân tách được đề cập.

Nói cách khác, các bộ chèn 110 và 108 hoạt động độc lập với sự tín hiệu hóa 34 được đề cập trên đây mà bộ biến đổi 104 được tạo cấu hình để tín hiệu hóa, hoặc chèn vào, dòng dữ liệu 20 lần lượt đối với các khung 26a và 26b.

Nói cách khác, trong phương án trên đây, bộ biến đổi 104 sắp xếp một cách thích hợp các hệ số biến đổi của các khung biến đổi dài và biến đổi phân tách, cụ thể bằng cách sắp xếp hoặc đan xen nối tiếp phẳng, và bộ chèn làm việc thực sự độc lập

với 109. Tuy nhiên trong nghĩa tổng quan hơn, bộ chèn đáp ứng được yêu cầu nếu tính độc lập của bộ chèn hệ số miền tần số từ sự tín hiệu hóa bị hạn chế đối với việc chèn của chuỗi gồm các hệ số miền tần số của từng khung biến đổi dài và biến đổi phân tách của tín hiệu âm thanh, được định tỷ lệ ngược theo các hệ số tỷ lệ, vào dòng dữ liệu trong đó, phụ thuộc vào sự tín hiệu hóa, chuỗi gồm các hệ số miền tần số được tạo thành bằng cách sắp xếp một cách tuần tự các hệ số miền tần số của một biến đổi của khung tương ứng theo cách không đan xen trong trường hợp khung là khung biến đổi dài, và bằng việc đan xen các hệ số miền tần số của nhiều hơn một biến đổi của khung tương ứng trong trường hợp khung tương ứng là khung biến đổi phân tách.

Trong phạm vi bộ chèn hệ số miền tần số 108 có liên quan, thực tế là bộ chèn tương tự hoạt động độc lập với sự tín hiệu hóa 34 một mặt phân biệt giữa các khung 26a và mặt khác các khung 26b, có nghĩa là bộ chèn 108 chèn các hệ số miền tần số của các khung của tín hiệu âm thanh, được định tỷ lệ ngược theo các hệ số tỷ lệ, vào dòng dữ liệu 20 theo chuỗi kế tiếp nhau trong trường hợp một biến đổi được thực hiện trong khung tương ứng, theo cách không đan xen, và chèn các hệ số miền tần số của các khung tương ứng bằng cách sử dụng việc đan xen trong trường hợp nhiều hơn một biến đổi được thực hiện đối với khung tương ứng, cụ thể là 2 trong ví dụ ở Fig.5. Tuy nhiên, như đã nêu trên đây, cách thức phân tách biến đổi còn có thể được thực hiện khác nhau để phân tách một biến đổi thành nhiều hơn 2 biến đổi.

Cuối cùng, nên chú ý rằng bộ mã hóa theo Fig.5 còn có thể được làm thích ứng để thực hiện tất cả các phương cách lập mã bổ sung khác được nêu ra trên đây đối với Fig.2 như lập mã MS, sự dự báo âm lập thể phức hợp 42 và TNS với, cho mục đích này, sự xác định các tham số tương ứng 44, 48 và 64 của chúng.

Mặc dù một số khía cạnh đã được mô tả trong phạm vi của thiết bị, rõ ràng rằng, các khía cạnh này còn thể hiện sự 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. Tương tự, các khía cạnh được mô tả trong phạm vi của bước phương pháp còn thể hiện sự mô tả của khối hoặc mục hoặc đặc điểm tương ứng của thiết bị tương ứng. Một số 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 tử. Trong một số phương án, một hoặc nhiều bước quan trọng nhất trong phương pháp có thể được thực hiện bởi thiết bị này.

Tùy theo các yêu cầu thực hiện nhất định, các phương án của sáng chế có thể được thực hiện bởi phần cứng hoặc trong phần mềm. Ứng dụng có thể được thực hiện bằng cách sử dụng vật ghi lưu trữ dạng số, ví dụ đĩa mềm, DVD, Blu-Ray, CD, ROM, PROM, EPROM, EEPROM hoặc bộ nhớ FLASH, có các tín hiệu kiểm soát có thể đọc được bằng điện tử được lưu trữ trên đó, mà tương thích (hoặc có khả năng tương thích) với hệ thống máy tính có thể lập trình sao cho phương pháp tương ứng được thực hiện. Do đó, vật ghi lưu trữ dạng số có thể đọc được bằng máy tính.

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

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

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

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

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

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

Phương án nữa 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.

Một 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) chương trình máy tính để thực hiện một trong các phương pháp được mô tả ở đây tới thiết bị nhận. Thiết bị 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 để truyền chương trình máy tính sang thiết bị 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 số 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 bên trên chỉ mang tính minh họa cho các nguyên lý của sáng chế. Cần hiểu rằng các sự biến thể và biến đổi của các phương án và các chi tiết được mô tả ở đây sẽ là rõ ràng đối với người có hiểu biết trung bình trong lĩnh vực kỹ thuật tương ứng. Do đó, sáng chế chỉ bị giới hạn bởi phạm vi của các điểm yêu cầu bảo hộ sắp đưa ra dưới đây và không bị giới hạn bởi các chi tiết cụ thể được biểu diễn bằng cách mô tả và giải thích của các phương án ở đây.

YÊU CẦU BẢO HỘ

1. Bộ giải mã âm thanh miền tần số hỗ trợ chuyển đổi độ dài biến đổi, bộ giải mã này bao gồm:

bộ chiết xuất hệ số miền tần số (12) được tạo cấu hình để chiết xuất các hệ số miền tần số (24) của các khung của tín hiệu âm thanh từ dòng dữ liệu;

bộ chiết xuất hệ số thang tỷ lệ (14) được tạo cấu hình để chiết xuất các hệ số thang tỷ lệ từ dòng dữ liệu;

bộ biến đổi ngược (16) được tạo cấu hình để đưa các hệ số miền tần số của các khung, mà được định tỷ lệ theo các hệ số thang tỷ lệ, đến phép biến đổi ngược để thu được các phần miền thời gian của tín hiệu âm thanh;

bộ tổng hợp (18) được tạo cấu hình để kết hợp các phần miền thời gian để thu được tín hiệu âm thanh,

trong đó, bộ biến đổi ngược đáp lại việc tín hiệu hóa trong các khung của tín hiệu âm thanh để, tùy thuộc vào việc tín hiệu hóa,

tạo thành một biến đổi bằng cách sắp xếp một cách liên tiếp các hệ số miền tần số của khung tương ứng, được định tỷ lệ theo các hệ số thang tỷ lệ, theo cách không khử đan xen và thực hiện một biến đổi là biến đổi ngược của độ dài biến đổi thứ nhất, hoặc

tạo thành nhiều hơn một biến đổi bằng việc khử đan xen các hệ số miền tần số của khung tương ứng, được định tỷ lệ theo các hệ số thang tỷ lệ, và thực hiện mỗi biến đổi trong số nhiều hơn một biến đổi là biến đổi ngược của độ dài biến đổi thứ hai, ngắn hơn so với độ dài biến đổi thứ nhất,

trong đó, bộ chiết xuất hệ số miền tần số và bộ chiết xuất hệ số thang tỷ lệ hoạt động độc lập với việc tín hiệu hóa,

trong đó bộ biến đổi ngược được tạo cấu hình để:

thực hiện việc lọc định hình nhiễu âm theo thời gian ngược (62) trên chuỗi gồm N hệ số bất kể việc tín hiệu hóa bằng cách áp dụng hàm chuyển đổi bộ lọc mà được

thiết lập theo các hệ số định hình nhiễu âm theo thời gian (temporal noise shaping TNS) (64) trên chuỗi gồm N hệ số, bằng cách:

trong việc tạo thành một biến đổi, áp dụng việc lọc định hình nhiễu âm theo thời gian ngược bằng cách sử dụng các hệ số miền tần số được sắp xếp một cách liên tục theo cách không khử đan xen như chuỗi gồm N hệ số, và

trong việc tạo thành nhiều hơn một biến đổi, áp dụng việc lọc định hình nhiễu âm theo thời gian ngược cho các hệ số miền tần số bằng cách sử dụng các hệ số miền tần số được sắp xếp một cách liên tục theo cách khử đan xen theo đó nhiều hơn một biến đổi được nối chuỗi theo quang phổ như chuỗi gồm N hệ số.

2. Bộ giải mã âm thanh miền tần số theo điểm 1, trong đó bộ chiết xuất hệ số thang tỷ lệ (14) được tạo cấu hình để chiết xuất các hệ số thang tỷ lệ từ dòng dữ liệu với độ phân giải quang phổ -theo thời gian mà độc lập với việc tín hiệu hóa.

3. Bộ giải mã âm thanh miền tần số theo điểm 1 hoặc 2, trong đó bộ chiết xuất hệ số miền tần số (12) sử dụng giải mã entropi dựa trên phạm vi hoặc bảng mã để chiết xuất các hệ số miền tần số từ dòng dữ liệu, với việc chỉ định, đối với từng hệ số miền tần số, phạm vi hoặc bảng mã giống nhau cho hệ số miền tần số tương ứng bất kể việc tín hiệu hóa.

4. Bộ giải mã âm thanh miền tần số theo một điểm bất kỳ trong số các điểm từ 1 đến 3, trong đó bộ biến đổi ngược được tạo cấu hình để các hệ số miền tần số được thực hiện việc định tỷ lệ theo các hệ số thang tỷ lệ với độ phân giải quang phổ độc lập với việc tín hiệu hóa.

5. Bộ giải mã âm thanh miền tần số theo một điểm bất kỳ trong số các điểm nêu trên, trong đó bộ biến đổi ngược được tạo cấu hình để các hệ số miền tần số được thực hiện việc điền đầy nhiễu âm, với các hệ số miền tần số được sắp xếp một cách liên tục theo cách không khử đan xen, và với độ phân giải quang phổ độc lập với việc tín hiệu hóa.

6. Bộ giải mã âm thanh miền tần số theo một điểm bất kỳ trong số các điểm nêu trên, trong đó bộ biến đổi ngược được tạo cấu hình để hỗ trợ việc lập mã âm lập thể ghép có hoặc không có sự dự báo âm lập thể liên kênh và để sử dụng các hệ số miền tần số như là phổ tổng (giữa) hoặc hiệu (bên) hoặc số dư dự báo của dự báo âm lập thể liên kênh,

với các hệ số miền tần số được sắp xếp theo cách không khử đan xen, bất kể việc tín hiệu hóa.

7. Bộ giải mã âm thanh miền tần số theo một điểm bất kỳ trong số các điểm nêu trên, trong đó số lượng của nhiều hơn một biến đổi là bằng 2, và độ dài biến đổi thứ nhất gấp đôi độ dài biến đổi thứ hai.

8. Bộ giải mã âm thanh miền tần số theo một điểm bất kỳ trong số các điểm nêu trên, trong đó phép biến đổi ngược là phép biến đổi cosin rời rạc được cải biên (modified discrete cosine transform - MDCT) ngược.

9. Phương pháp giải mã âm thanh miền tần số hỗ trợ chuyển đổi độ dài biến đổi, phương pháp bao gồm các bước:

chiết xuất các hệ số miền tần số của các khung của tín hiệu âm thanh từ dòng dữ liệu;

chiết xuất các hệ số thang tỷ lệ từ dòng dữ liệu;

thực hiện biến đổi ngược các hệ số miền tần số của các khung, được định tỷ lệ theo các hệ số thang tỷ lệ, để thu được các phần miền thời gian của tín hiệu âm thanh;

kết hợp các phần miền thời gian để thu được tín hiệu âm thanh,

trong đó, việc thực hiện biến đổi ngược đáp ứng với việc tín hiệu hóa trong các khung của tín hiệu âm thanh để, phụ thuộc vào việc tín hiệu hóa, bao gồm bước:

tạo thành một biến đổi bằng cách sắp xếp một cách liên tục các hệ số miền tần số của khung tương ứng theo cách không khử đan xen và thực hiện một biến đổi là biến đổi ngược của độ dài biến đổi thứ nhất, hoặc

tạo thành nhiều hơn một biến đổi bằng việc khử đan xen các hệ số miền tần số của khung tương ứng và thực hiện từng biến đổi trong số nhiều hơn một biến đổi là biến đổi ngược của độ dài biến đổi thứ hai, ngắn hơn so với độ dài biến đổi thứ nhất,

trong đó, việc chiết xuất các hệ số miền tần số và việc chiết xuất các hệ số thang tỷ lệ là độc lập với việc tín hiệu hóa,

trong đó việc thực hiện biến đổi ngược bao gồm các bước:

thực hiện việc lọc định hình nhiều âm theo thời gian ngược (62) trên chuỗi gồm N hệ số bất kể việc tín hiệu hóa bằng cách áp dụng bộ lọc hàm của bộ biến đổi của bộ lọc mà được thiết lập theo các hệ số TNS (64) trên chuỗi gồm N hệ số bằng cách:

trong việc tạo thành một biến đổi, áp dụng việc lọc định hình nhiều âm theo thời gian ngược bằng cách sử dụng các hệ số miền tần số được sắp xếp một cách liên tục theo cách không khử đan xen như chuỗi gồm N hệ số, và

trong việc tạo thành nhiều hơn một biến đổi, áp dụng việc lọc định hình nhiều âm theo thời gian ngược cho các hệ số miền tần số bằng cách sử dụng các hệ số miền tần số được sắp xếp một cách liên tục theo cách không khử đan xen theo đó nhiều hơn một biến đổi được nối chuỗi theo quang phổ như chuỗi gồm N hệ số.

10. Bộ mã hóa âm thanh miền tần số hỗ trợ chuyển đổi độ dài biến đổi, bộ mã hóa này bao gồm:

bộ biến đổi (104) được tạo cấu hình để biến đổi các phần miền thời gian của tín hiệu âm thanh để thu được các hệ số miền tần số của các khung của tín hiệu âm thanh;

bộ định tỷ lệ ngược (106) được tạo cấu hình để định tỷ lệ ngược các hệ số miền tần số theo các hệ số thang tỷ lệ;

bộ chèn hệ số miền tần số (108) được tạo cấu hình để chèn các hệ số miền tần số của các khung của tín hiệu âm thanh, được định tỷ lệ ngược theo các hệ số thang tỷ lệ, vào trong dòng dữ liệu; và

bộ chèn hệ số thang tỷ lệ (110) được tạo cấu hình để chèn các hệ số thang tỷ lệ vào trong dòng dữ liệu,

trong đó, bộ biến đổi được tạo cấu hình để chuyển đổi các khung của các tín hiệu âm thanh ít nhất giữa bước

thực hiện một biến đổi độ dài biến đổi thứ nhất cho khung tương ứng, và

thực hiện nhiều hơn một biến đổi của độ dài biến đổi thứ hai, ngắn hơn so với độ dài biến đổi thứ nhất, cho khung tương ứng,

trong đó, bộ biến đổi còn được tạo cấu hình để báo hiệu việc chuyển đổi bằng việc tín hiệu hóa trong các khung của dòng dữ liệu;

trong đó, bộ chèn hệ số miền tần số được tạo cấu hình để:

tùy thuộc vào việc tín hiệu hóa, tạo thành chuỗi gồm các hệ số miền tần số bằng cách

sắp xếp một cách liên tục các hệ số miền tần số của một biến đổi của khung tương ứng theo cách không đan xen trong trường hợp một biến đổi được thực hiện cho khung tương ứng, và

đan xen các hệ số miền tần số của nhiều hơn một biến đổi của khung tương ứng trong trường hợp nhiều hơn một biến đổi được thực hiện cho khung tương ứng,

theo cách độc lập với việc tín hiệu hóa, chèn, cho khung tương ứng, chuỗi gồm các hệ số miền tần số của khung tương ứng tín hiệu âm thanh, được định tỷ lệ ngược theo các hệ số thang tỷ lệ, vào trong dòng dữ liệu,

trong đó bộ chèn hệ số thang tỷ lệ hoạt động độc lập với việc tín hiệu hóa,

trong đó bộ mã hóa được tạo cấu hình để:

thực hiện việc định hình nhiễu âm theo thời gian ngược trên chuỗi gồm N hệ số để xác định các hệ số TNS (64) theo cách bất kể việc tín hiệu hóa trong đó:

trong trường hợp thực hiện một biến đổi, các hệ số miền tần số được sắp xếp một cách liên tục theo cách không đan xen được sử dụng như chuỗi gồm N hệ số, và

trong trường hợp thực hiện nhiều hơn một biến đổi, các hệ số miền tần số được sắp xếp một cách liên tục theo cách không đan xen theo đó nhiều hơn một biến đổi được nối chuỗi theo quang phổ được sử dụng như chuỗi gồm N hệ số.

11. Phương pháp mã hóa âm thanh miền tần số hỗ trợ chuyển đổi độ dài biến đổi, phương pháp bao gồm các bước:

thực hiện biến đổi các phần miền thời gian của tín hiệu âm thanh để thu được các hệ số miền tần số của các khung của tín hiệu âm thanh;

định tỷ lệ ngược cho các hệ số miền tần số theo các hệ số thang tỷ lệ;

chèn các hệ số miền tần số của các khung của tín hiệu âm thanh, được định tỷ lệ ngược theo các hệ số thang tỷ lệ, vào trong dòng dữ liệu; và

chèn các hệ số thang tỷ lệ vào trong dòng dữ liệu,

trong đó, việc thực hiện biến đổi chuyển đổi các khung của tín hiệu âm thanh ít nhất giữa các công đoạn:

thực hiện một biến đổi của độ dài biến đổi thứ nhất cho khung tương ứng, và

thực hiện nhiều hơn một biến đổi của độ dài biến đổi thứ hai, ngắn hơn độ dài biến đổi thứ nhất, cho khung tương ứng,

trong đó, phương pháp bao gồm việc báo hiệu chuyển đổi bằng việc tín hiệu hóa trong các khung của dòng dữ liệu;

trong đó việc chèn các hệ số miền tần số được thực hiện bởi sự phụ thuộc vào việc tín hiệu hóa, chuỗi gồm các hệ số miền tần số được tạo thành bằng cách

sắp xếp một cách liên tục các hệ số miền tần số của một biến đổi của khung tương ứng theo cách không đan xen trong trường hợp một biến đổi được thực hiện cho khung tương ứng, và

bằng cách đan xen các hệ số miền tần số của nhiều hơn một biến đổi của khung tương ứng trong trường hợp nhiều hơn một biến đổi được thực hiện cho khung tương ứng,

theo cách độc lập với việc tín hiệu hóa, chèn, cho khung tương ứng, chuỗi gồm các hệ số miền tần số của khung tương ứng của tín hiệu âm thanh, được định tỷ lệ ngược theo các hệ số thang tỷ lệ, vào trong dòng dữ liệu

trong đó, việc chèn các hệ số thang tỷ lệ được thực hiện độc lập với việc tín hiệu hóa,

trong đó phương pháp bao gồm các bước:

thực hiện việc định hình nhiễu âm theo thời gian trên chuỗi gồm N hệ số để xác định các hệ số TNS (64) theo cách bất kể việc tín hiệu hóa, trong đó:

trong trường hợp thực hiện một biến đổi, các hệ số miền tần số được sắp xếp một cách liên tục theo cách không khử đan xen được sử dụng như chuỗi gồm N hệ số, và

trong trường hợp thực hiện nhiều hơn một biến đổi, các hệ số miền tần số được sắp xếp một cách liên tục theo cách không khử đan xen theo đó nhiều hơn một biến đổi được nối chuỗi theo quang phổ được sử dụng như chuỗi gồm N hệ số.

12. Vật ghi đọc được bằng máy tính bao gồm chương trình máy tính có mã chương trình để thực hiện phương pháp theo điểm 9, khi chạy trên máy tính.

13. Vật ghi đọc được bằng máy tính bao gồm chương trình máy tính có mã chương trình để thực hiện phương pháp theo điểm 11, khi chạy trên máy tính.

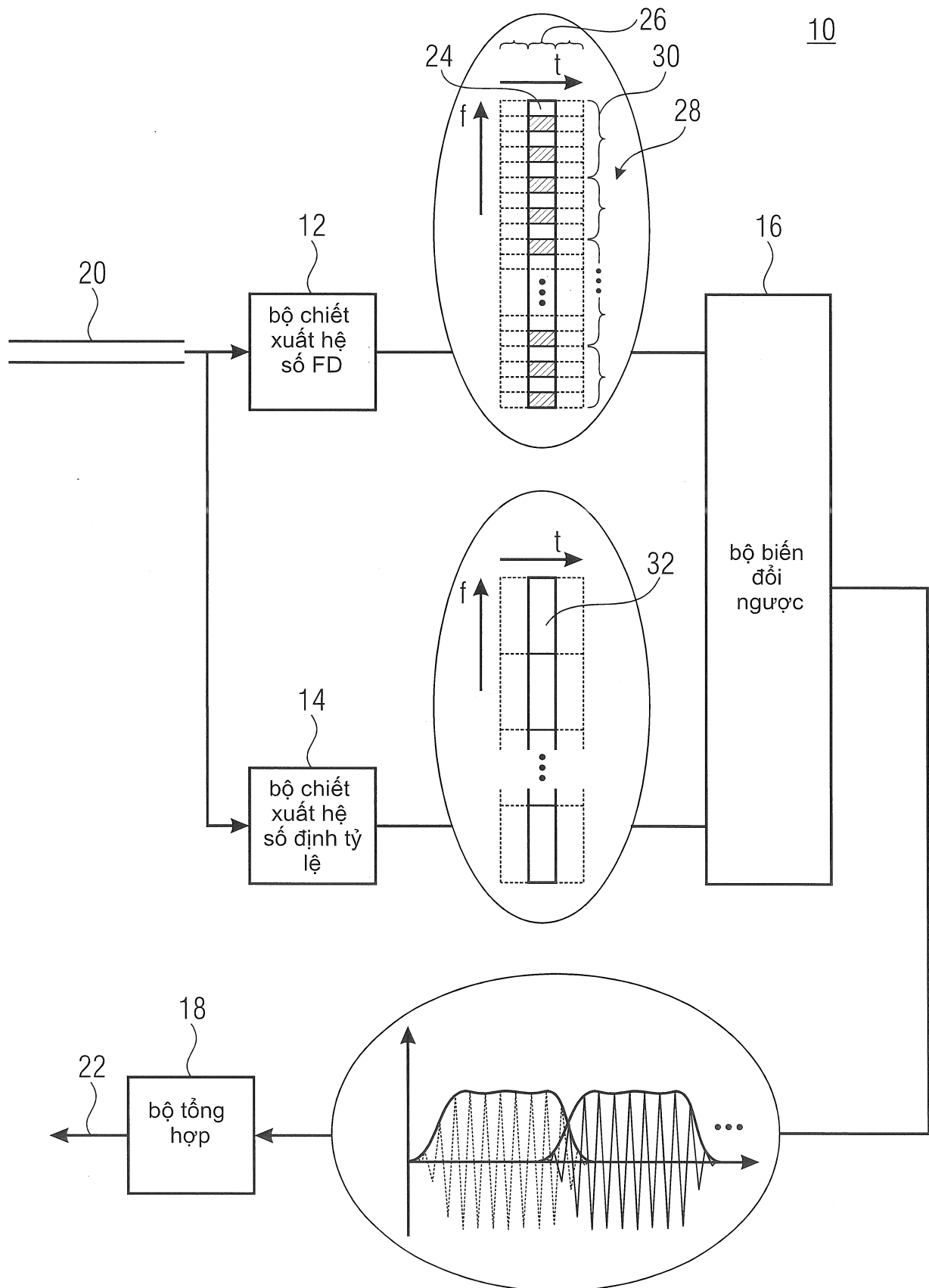


FIG 1

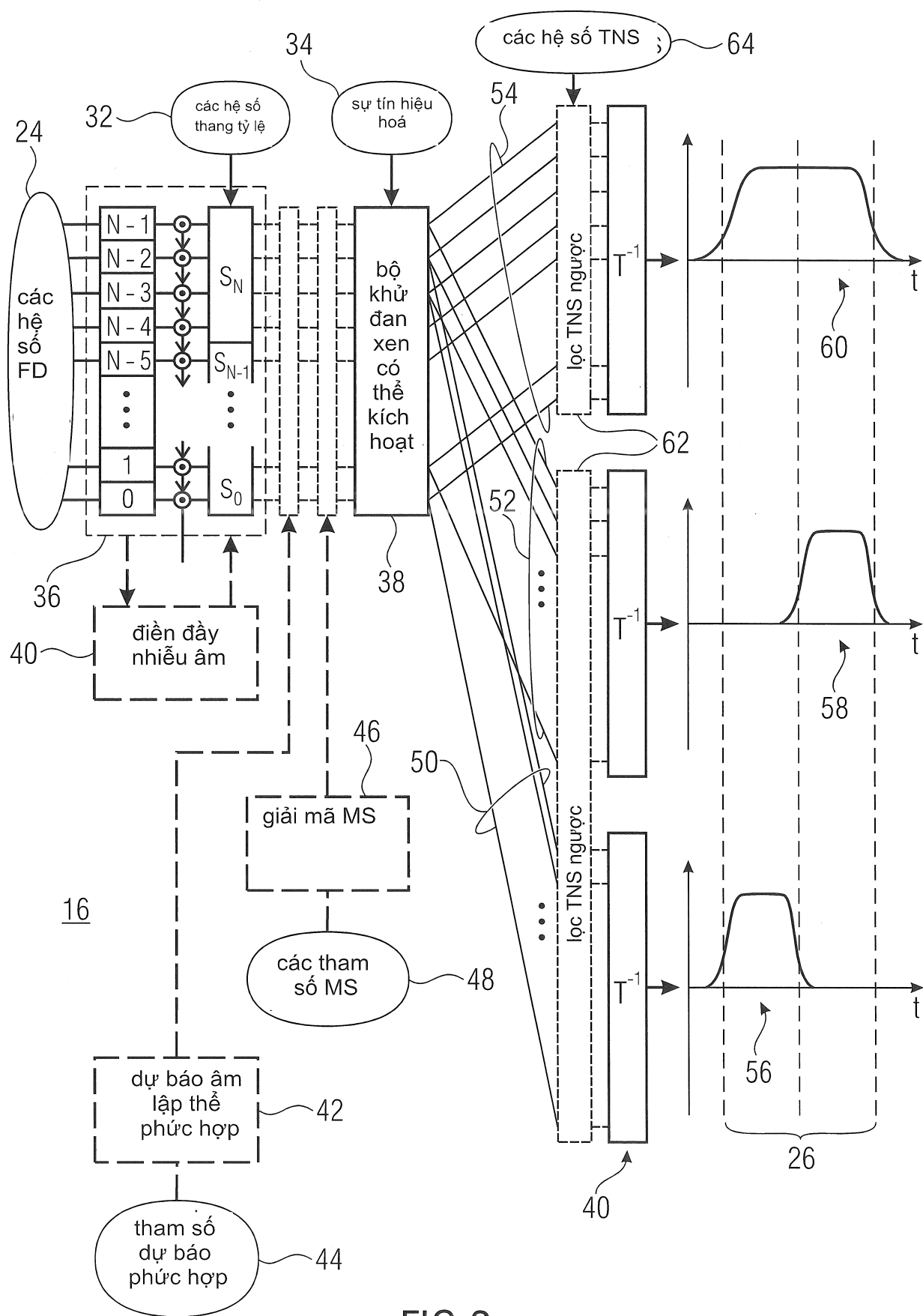


FIG 2

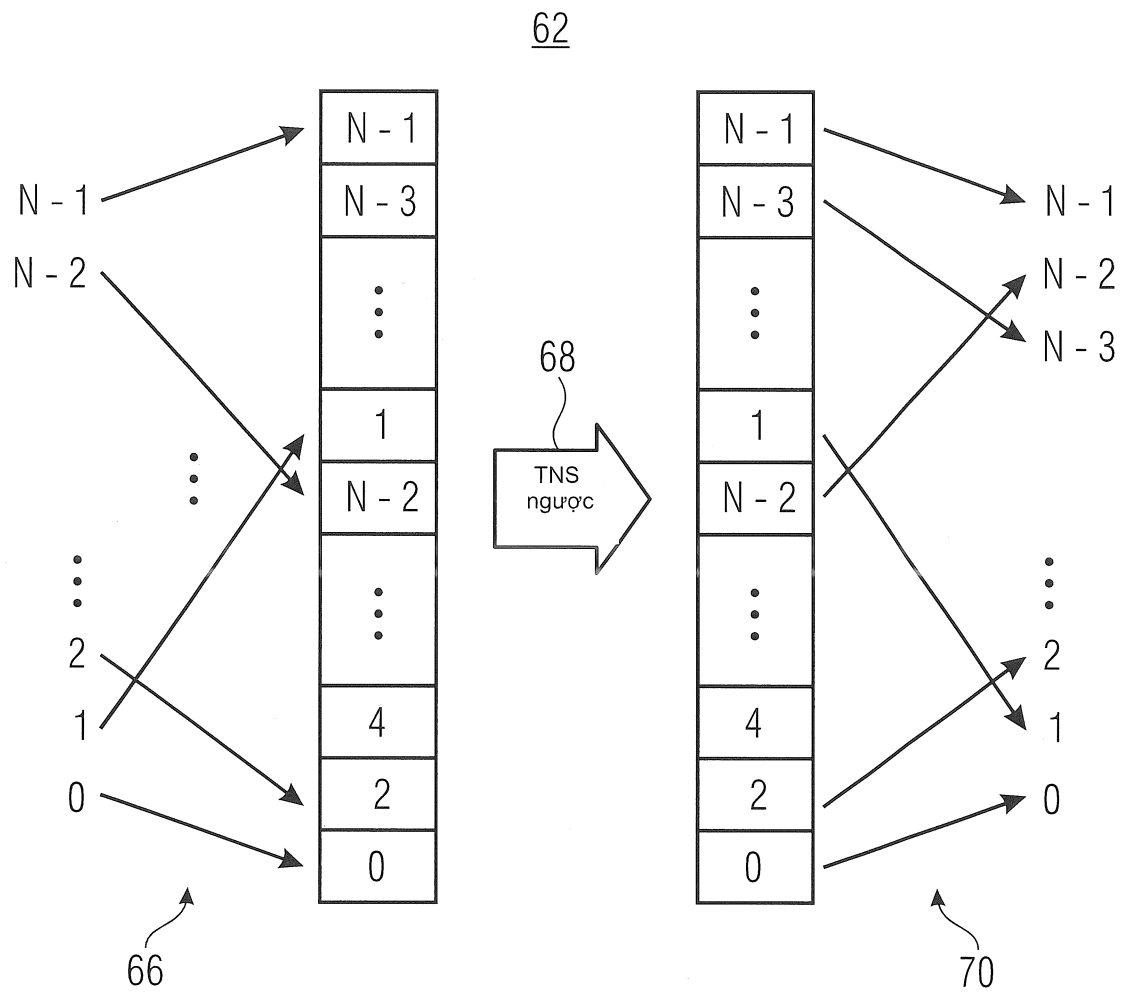


FIG 3

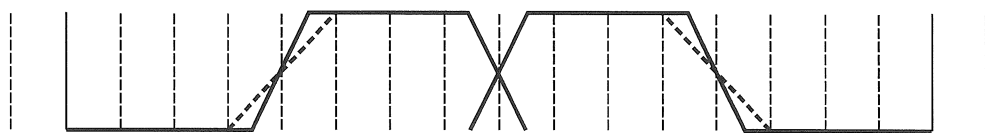


FIG 4

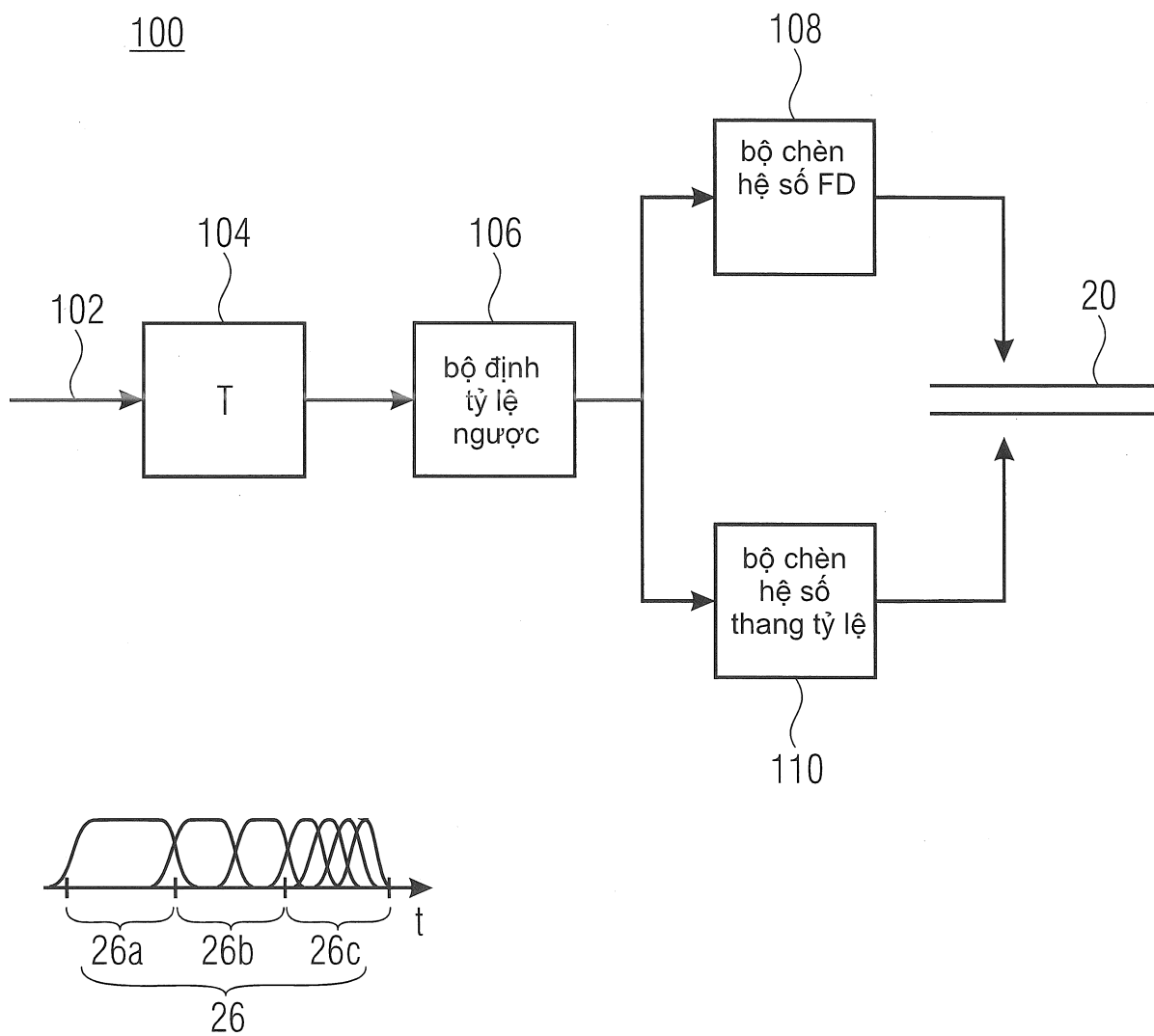


FIG 5