



(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)



1-0030817

(51)⁷ G10L 19/03; G10L 21/0388; G10L
25/03; G10L 21/0272

(13) B

(21) 1-2017-03748

(22) 23/02/2016

(86) PCT/EP2016/053752 23/02/2016

(87) WO/2016/135132 01/09/2016

(30) 15156704.7 26/02/2015 EP; 15181118.9 14/08/2015 EP

(45) 25/01/2022 406

(43) 27/11/2017 356A

(73) FRAUNHOFER-GESELLSCHAFT ZUR FOERDERUNG DER ANGEWANDTEN
FORSCHUNG E.V. (DE)

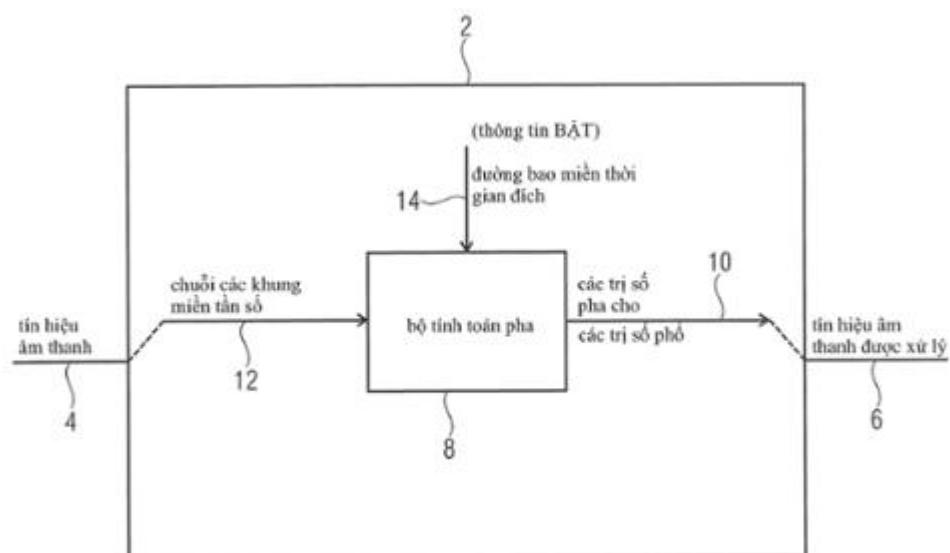
Hansastraße 27c, 80686 Muenchen, Germany

(72) DITTMAR, Christian (DE); MUELLER, Meinard (DE); DISCH, Sascha (DE).

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

(54) THIẾT BỊ VÀ PHƯƠNG PHÁP XỬ LÝ TÍN HIỆU ÂM THANH, BỘ GIẢI MÃ
ÂM THANH VÀ PHƯƠNG PHÁP GIẢI MÃ ÂM THANH, BỘ XỬ LÝ VÀ
PHƯƠNG PHÁP TÁCH NGUỒN ÂM THANH, BỘ XỬ LÝ VÀ PHƯƠNG PHÁP
NÂNG CAO BĂNG THÔNG

(57) Sáng chế đề cập đến thiết bị và phương pháp xử lý tín hiệu âm thanh, bộ giải mã âm thanh và phương pháp giải mã âm thanh, bộ xử lý và phương pháp tách nguồn âm thanh bộ xử lý và phương pháp nâng cao băng thông. Đối tượng của sáng chế là thiết bị (2) được mô tả bởi sơ đồ khối dạng giản lược để xử lý tín hiệu âm thanh (4) để thu được tín hiệu âm thanh được xử lý (6). Thiết bị (2) bao gồm bộ tính toán pha (8) để tính toán các trị số pha (10) cho các trị số phổ của chuỗi các khung miền tần số (12) biểu diễn các khung chồng lấp của tín hiệu âm thanh (4). Hơn nữa, bộ tính toán pha (8) được tạo cấu bình để tính toán các giá trị pha (10) dựa trên thông tin về đường bao miền thời gian đích (14) liên quan đến tín hiệu âm thanh được xử lý (6), để tín hiệu âm thanh được xử lý (6) có ít nhất trong phép tính gần đúng đường bao biên độ miền thời gian đích (14) và đường bao phổ được xác định bởi chuỗi các khung miền tần số (12).



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến thiết bị và phương pháp xử lý tín hiệu âm thanh để thu được tín hiệu âm thanh được xử lý. Các phương án còn thể hiện bộ giải mã âm thanh bao gồm thiết bị và bộ mã hóa âm thanh tương ứng, bộ xử lý tách nguồn âm thanh và bộ xử lý nâng cao băng thông, cả hai đều bao gồm thiết bị. Theo các phương án khác, sự phục hồi chuyển tiếp trong sự khôi phục tín hiệu và sự phục hồi chuyển tiếp trong sự khai triển âm thanh được thông báo theo điểm được thể hiện.

Tình trạng kỹ thuật của sáng chế

Nhiệm vụ tách hỗn hợp các nguồn âm thanh được xếp chồng thành các thành phần cấu thành đã làm tăng tầm quan trọng trong việc xử lý tín hiệu âm thanh kỹ thuật số. Trong xử lý tiếng nói, các thành phần này thường được sử dụng sự bày tỏ của các loa phát thanh đích bị cản trở bởi nhiễu âm hoặc người nói liên tục. Trong âm nhạc, các thành phần này có thể là nhạc cụ riêng hoặc giai điệu giọng hát, các nhạc cụ gõ, hoặc thậm chí các sự kiện nốt riêng lẻ. Các đề tài liên quan là khôi phục tín hiệu và bảo toàn chuyển tiếp và sự hợp thành âm thanh được thông báo điểm (tức là, sự tách nguồn).

Sự tách nguồn nhằm mục đích khai triển bản ghi âm nhạc phức điệu, đa âm sắc thành các tín hiệu thành phần như tiếng nói hát, nhạc điệu nhạc cụ, nhạc cụ gõ, hoặc các sự kiện nốt riêng lẻ xảy ra trong tín hiệu hỗn hợp. Ngoài việc trở thành bước quan trọng trong nhiều tác vụ phân tích và phục hồi âm nhạc, tách nguồn nhạc cũng là điều kiện tiên quyết cơ bản cho các ứng dụng như khôi phục âm nhạc, trộn tăng, và trộn lại. Đối với các mục đích này, độ chính xác cao về chất lượng cảm quan của các thành phần riêng lẻ là mong muốn. Phần lớn các công nghệ tách hiện có hoạt động trên sự biểu diễn thời gian - tần số (time-frequency - TF) của tín hiệu hỗn hợp. Các tín hiệu thành phần đích thường được khôi phục sử dụng phép biến đổi nghịch đảo phù hợp, mà lần lượt có thể đưa các thành phần lại có thể nghe được như nhiễu âm nhạc, các sự chuyển tiếp mờ hoặc tiếng vang trước. Các phương pháp hiện có phải chịu các thành phần lại có thể nghe

được dưới dạng nhiễu âm nhạc, sự nhiễu pha và các tiếng vang trước. Các thành phần lạ này thường làm nhiễu một chút đối với người nghe.

Có nhiều nghiên cứu gần đây về việc tách nguồn âm nhạc. Trong hầu hết các phương pháp, việc tách được thực hiện trong miền thời gian - tần số (time-frequency - TF) bằng cách cải biên ảnh phổ cường độ. Các tín hiệu miền thời gian tương ứng của các thành phần tách được suy ra bằng cách sử dụng thông tin pha ban đầu và áp dụng các phép biến đổi nghịch đảo phù hợp. Khi cố gắng đạt được chất lượng cảm quan tốt của các tín hiệu đơn được tách, nhiều tác giả quay lại các kỹ thuật khai triển được thông báo điểm. Điều này có lợi thế rằng việc tách có thể được hướng bằng thông tin về vị trí gần đúng của các tín hiệu thành phần trong thời gian (bắt đầu, kết thúc) và tần số (độ cao, âm sắc). Ít các tài liệu công bố giải quyết việc tách nguồn của các tín hiệu chuyển tiếp như trống. Các tài liệu khác đã tập trung vào việc tách sóng hài với các thành phần bộ gõ [5].

Hơn nữa, vấn đề của tiếng vang trước đã được chỉ ra trong lĩnh vực mã hóa âm thanh cảm quan, trong đó các tiếng vang trước thường được gây ra bởi sự sử dụng phép phân tích tương đối dài và các cửa sổ tổng hợp kết nối với thao tác trung gian của các ngăn TF như lượng tử hóa các cường độ phổ theo mô hình tâm thính học. Có thể xem xét tình trạng kỹ thuật để sử dụng sự chuyển đổi khối trong vùng lân cận các sự kiện chuyển tiếp [6]. Phương pháp thú vị được đề xuất trong tài liệu [13] trong đó các hệ số phổ được mã hóa bằng phép dự báo tuyến tính dọc theo trục tần số, tự động giảm các tiếng vang trước. Các công việc sau đó được đề xuất để khai triển tín hiệu thành các thành phần chuyển tiếp và dự và sử dụng các tham số mã hóa được tối ưu hóa cho mỗi dòng [3]. Sự bảo toàn chuyển tiếp cũng được kiểm tra trong bối cảnh các phương pháp cải biên định tỉ lệ thời gian trên cơ sở bộ mã hóa tiếng nói theo pha. Ngoài việc xử lý được tối ưu hóa các thành phần chuyển tiếp, một số tác giả đi theo nguyên tắc khóa pha hoặc tái khởi tạo pha trong các khung chuyển tiếp [8].

Vấn đề của sự khôi phục tín hiệu, cũng được biết như phép nghịch đảo ảnh phổ cường độ hoặc sự ước lượng pha là chủ đề được nghiên cứu tốt. Trong nghiên cứu cổ điển của họ [1], Griffin và Lim đã đề xuất thuật toán gọi là LSEE-MSTFTM cho sự khôi phục tín hiệu lặp, mà từ các ảnh phổ cường độ STFT được cải biên (modified STFT

magnitude - MSTFTM). Trong tài liệu [2], Le Roux và cộng sự đã phát triển cái nhìn khác về phương pháp này bằng cách mô tả sử dụng tiêu chuẩn đồng nhất TF. Bằng cách giữ các thuật toán cần thiết hoàn toàn trong miền TF, một vài sự đơn giản hóa và làm gần đúng có thể được đưa vào mà hạ thấp tải tính toán so với quy trình ban đầu. Vì các sự ước lượng pha thu được sử dụng LSEE-MSTFTM chỉ có thể hội tụ đến điều kiện tốt nhất cục bộ, một số công bố đã quan tâm việc tìm sự ước lượng ban đầu tốt đối với thông tin pha [3, 4]. Sturmel và Daudet [5] đã cung cấp bài báo chuyên sâu về các phương pháp khôi phục tín hiệu và chỉ ra các vấn đề chưa được giải quyết. Sự mở rộng LSEE-MSTFTM so với tốc độ hội tụ được đề xuất trong [6]. Các tác giả khác đã cố tạo công thức vấn đề ước lượng pha như sơ đồ tối ưu hóa lồi và đạt đến kết quả hứa hẹn được cản trở bởi độ phức tạp tính toán cao [7]. Tác phẩm khác [8] đã quan tâm đến việc áp dụng khung đồng nhất ảnh phổ vào sự khôi phục tín hiệu từ các ảnh phổ cường độ trên cơ sở sóng con.

Tuy nhiên, các phương pháp được mô tả cho việc khôi phục tín hiệu có chung vấn đề là sự thay đổi nhanh của tín hiệu âm thanh, ví dụ, là diễn hình cho sự chuyển tiếp, có thể phải chịu do các thành phần được mô tả trước đó như, ví dụ, các tiếng vang trước.

Do đó, phương pháp được cải thiện là cần thiết.

Tài liệu tham khảo

[1] Daniel W. Griffin and Jae S. Lim, “Signal estimation from modified short-time Fourier transform”, IEEE Transactions on Acoustics, Speech and Signal Processing, vol. 32, no. 2, pp. 236–243, April 1984.

[2] Jonathan Le Roux, Nobutaka Ono, and Shigeki Sagayama, “Explicit consistency constraints for STFT spectrograms and their application to phase reconstruction” in Proceedings of the ISCA Tutorial and Research Workshop on Statistical And Perceptual Audition, Brisbane, Australia, September 2008, pp. 23–28.

[3] Xinglei Zhu, Gerald T. Beauregard, and Lonce L. Wyse, “Real-time signal estimation from modified short-time Fourier transform magnitude spectra”, IEEE Transactions on Audio, Speech, and Language Processing, vol. 15, no. 5, pp. 1645–1653, July 2007.

[4] Jonathan Le Roux, Hirokazu Kameoka, Nobutaka Ono, and Shigeki Sagayama, “Phase initialization schemes for faster spectrogram-consistency-based signal reconstruction” in Proceedings of the Acoustical Society of Japan Autumn Meeting, September 2010, number 3-10-3.

[5] Nicolas Sturmel and Laurent Daudet, “Signal reconstruction from STFT magnitude: a state of the art” in Proceedings of the International Conference on Digital Audio Effects (DAFx), Paris, France, September 2011, pp. 375–386.

[6] Nathanaël Perraudin, Peter Balazs, and Peter L. Søndergaard, “A fast Griffin-Lim algorithm” in Proceedings IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA), New Paltz, NY, USA, October 2013, pp. 1–4.

[7] Dennis L. Sun and Julius O. Smith III, “Estimating a signal from a magnitude spectrogram via convex optimization” in Proceedings of the Audio Engineering Society (AES) Convention, San Francisco, USA, October 2012, Preprint 8785.

[8] Tomohiko Nakamura and Hiokazu Kameoka, “Fast signal reconstruction from magnitude spectrogram of continuous wavelet transform based on spectrogram consistency” in Proceedings of the International Conference on Digital Audio Effects (DAFx), Erlangen, Germany, September 2014, pp. 129–135.

[9] Volker Gnann and Martin Spiertz, “Inversion of shorttime fourier transform magnitude spectrograms with adaptive window lengths” in Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, (ICASSP), Taipei, Taiwan, April 2009, pp. 325–328.

[10] Jonathan Le Roux, Hirokazu Kameoka, Nobutaka Ono, and Shigeki Sagayama, “Fast signal reconstruction from magnitude STFT spectrogram based on spectrogram consistency” in Proceedings International Conference on Digital Audio Effects (DAFx), Graz, Austria, September 2010, pp. 397–403.

Bản chất kỹ thuật của sáng chế

Mục đích của sáng chế là cung cấp khái niệm được cải thiện để xử lý tín hiệu âm thanh.

Theo phương án, thiết bị xử lý tín hiệu âm thanh để thu được tín hiệu âm thanh

được xử lý có thể có: bộ tính toán pha để tính toán các giá trị pha cho các giá trị phổ của chuỗi các khung miền tần số biểu diễn các khung chồng lấp của tín hiệu âm thanh, trong đó bộ tính toán pha được tạo cấu hình để tính toán các giá trị pha dựa trên thông tin về đường bao miền thời gian đích liên quan đến tín hiệu âm thanh được xử lý, để tín hiệu âm thanh được xử lý bao gồm ít nhất trong phép tính gần đúng đường bao miền thời gian đích và đường bao phổ được xác định bởi chuỗi các khung miền tần số.

Theo phương án khác, bộ mã hóa âm thanh để mã hóa tín hiệu âm thanh, bộ mã hóa có thể có: bộ xử lý tín hiệu âm thanh được tạo cấu hình để mã hóa tín hiệu âm thanh miền thời gian sao cho tín hiệu âm thanh được mã hóa có sự biểu diễn của chuỗi các khung miền tần số của tín hiệu âm thanh và sự biểu diễn của đường bao miền thời gian, và bộ xác định đường bao được tạo cấu hình để xác định đường bao từ tín hiệu âm thanh miền thời gian, trong đó bộ xác định đường bao còn được tạo cấu hình để so sánh đường bao được định trước để xác định sự biểu diễn của đường bao miền thời gian đích trên cơ sở so sánh.

Theo phương án khác, tín hiệu âm thanh có thể có: sự biểu diễn của chuỗi các khung miền tần số của tín hiệu âm thanh miền thời gian và sự biểu diễn của đường bao miền thời gian.

Theo phương án khác, bộ xử lý tách nguồn âm thanh có thể có: thiết bị theo sáng chế, và màn chắn phổ để chắn phổ của tín hiệu âm thanh ban đầu để thu được tín hiệu âm thanh được cải biên được nhập vào thiết bị để xử lý, trong đó tín hiệu âm thanh được xử lý là tín hiệu nguồn được tách liên quan đến đường bao miền thời gian đích.

Theo phương án khác, bộ xử lý nâng cao băng thông để xử lý tín hiệu âm thanh được mã hóa, bộ xử lý nâng cao băng thông có thể có: bộ xử lý nâng cao để tạo ra tín hiệu nâng cao từ băng tín hiệu âm thanh nằm trong tín hiệu được mã hóa, và thiết bị theo sáng chế để xử lý, trong đó bộ xử lý nâng cao được tạo cấu hình để trích đường bao miền thời gian đích từ sự biểu diễn được mã hóa nằm trong tín hiệu được mã hóa hoặc từ băng tín hiệu âm thanh nằm trong tín hiệu được mã hóa.

Theo phương án khác, phương pháp xử lý tín hiệu âm thanh để thu được tín hiệu âm thanh được xử lý, phương pháp có thể có các bước: tính toán các giá trị pha cho các giá trị phổ của chuỗi các khung miền tần số biểu diễn các khung chồng lấp của tín hiệu

âm thanh, trong đó các giá trị pha được tính toán dựa trên thông tin về đường bao miền thời gian đích liên quan đến tín hiệu âm thanh được xử lý, để tín hiệu âm thanh được xử lý bao gồm ít nhất trong phép tính gần đúng đường bao miền thời gian đích và đường bao phổ được xác định bởi chuỗi các khung miền tần số.

Theo phương án khác, phương pháp giải mã âm thanh, phương pháp có thể có phương pháp xử lý tín hiệu âm thanh để thu được tín hiệu âm thanh được xử lý, phương pháp có thể có các bước: tính toán các giá trị pha cho các giá trị phổ của chuỗi các khung miền tần số biểu diễn các khung chồng lấp của tín hiệu âm thanh, trong đó các giá trị pha được tính toán dựa trên thông tin về đường bao miền thời gian đích liên quan đến tín hiệu âm thanh được xử lý, để tín hiệu âm thanh được xử lý bao gồm ít nhất trong phép tính gần đúng đường bao miền thời gian đích và đường bao phổ được xác định bởi chuỗi các khung miền tần số; nhận tín hiệu được mã hóa, tín hiệu được mã hóa bao gồm sự biểu diễn của chuỗi các khung miền tần số, và sự biểu diễn của đường bao miền thời gian đích.

Theo phương án khác, phương pháp tách nguồn âm thanh, phương pháp có thể có: phương pháp xử lý tín hiệu âm thanh để thu được tín hiệu âm thanh được xử lý, phương pháp có thể có các bước: tính toán các giá trị pha cho các giá trị phổ của chuỗi các khung miền tần số biểu diễn các khung chồng lấp của tín hiệu âm thanh, trong đó các giá trị pha được tính toán dựa trên thông tin về đường bao miền thời gian đích liên quan đến tín hiệu âm thanh được xử lý, để tín hiệu âm thanh được xử lý bao gồm ít nhất trong phép tính gần đúng đường bao miền thời gian đích và đường bao phổ được xác định bởi chuỗi các khung miền tần số, và tạo màn chắn phổ của tín hiệu âm thanh ban đầu để thu được tín hiệu âm thanh được cải biên được nhập vào thiết bị để xử lý; trong đó tín hiệu âm thanh được xử lý là tín hiệu nguồn được tách liên quan đến đường bao miền thời gian đích.

Theo phương án khác, phương pháp nâng cao băng thông của tín hiệu âm thanh được mã hóa, phương pháp có thể có: tạo ra tín hiệu nâng cao từ băng tín hiệu âm thanh nằm trong tín hiệu được mã hóa; phương pháp xử lý tín hiệu âm thanh để thu được tín hiệu âm thanh được xử lý, phương pháp có các bước: tính toán các giá trị pha cho các giá trị phổ của chuỗi các khung miền tần số biểu diễn các khung chồng lấp của tín hiệu

âm thanh, trong đó các giá trị pha được tính toán dựa trên thông tin về đường bao miền thời gian đích liên quan đến tín hiệu âm thanh được xử lý, để tín hiệu âm thanh được xử lý bao gồm ít nhất trong phép tính gần đúng đường bao miền thời gian đích và đường bao phổ được xác định bởi chuỗi các khung miền tần số; trong đó tạo ra bao gồm trích đường bao miền thời gian đích từ sự biểu diễn được mã hóa nằm trong tín hiệu được mã hóa hoặc từ băng tín hiệu âm thanh nằm trong tín hiệu được mã hóa.

Theo phương án khác, phương pháp mã hóa âm thanh, phương pháp có thể có các bước: mã hóa tín hiệu âm thanh miền thời gian sao cho tín hiệu âm thanh được mã hóa bao gồm sự biểu diễn chuỗi các khung miền tần số của tín hiệu âm thanh miền thời gian và sự biểu diễn đường bao miền thời gian đích; và xác định đường bao từ tín hiệu âm thanh, và so sánh đường bao với tập hợp các đường bao được định trước để xác định sự biểu diễn đường bao miền thời gian đích trên cơ sở so sánh.

Phương án khác có thể có vật ghi lưu trữ số không tạm thời bao gồm chương trình máy tính được chứa trên đó để thực hiện phương pháp xử lý tín hiệu âm thanh để thu được tín hiệu âm thanh được xử lý, phương pháp có các bước: tính toán các giá trị pha cho các giá trị phổ của chuỗi các khung miền tần số biểu diễn các khung chồng lấp của tín hiệu âm thanh, trong đó các giá trị pha được tính toán dựa trên thông tin về đường bao miền thời gian đích liên quan đến tín hiệu âm thanh được xử lý, để tín hiệu âm thanh được xử lý bao gồm ít nhất trong phép tính gần đúng đường bao miền thời gian đích và đường bao phổ được xác định bởi chuỗi các khung miền tần số, khi chương trình máy tính nêu trên được chạy bởi máy tính.

Phương án khác có thể có vật ghi lưu trữ số không tạm thời bao gồm chương trình máy tính được chứa trên đó để thực hiện phương pháp giải mã âm thanh, phương pháp có: phương pháp xử lý tín hiệu âm thanh để thu được tín hiệu âm thanh được xử lý, phương pháp có các bước: tính toán các giá trị pha cho các giá trị phổ của chuỗi các khung miền tần số biểu diễn các khung chồng lấp của tín hiệu âm thanh, trong đó các giá trị pha được tính toán dựa trên thông tin về đường bao miền thời gian đích liên quan đến tín hiệu âm thanh được xử lý, để tín hiệu âm thanh được xử lý bao gồm ít nhất trong phép tính gần đúng đường bao miền thời gian đích và đường bao phổ được xác định bởi chuỗi các khung miền tần số; nhận tín hiệu được mã hóa, tín hiệu được mã hóa bao gồm

sự biểu diễn của chuỗi các khung miền tần số, và sự biểu diễn của đường bao miền thời gian đích, khi chương trình máy tính nêu trên được chạy bởi máy tính.

Phương án khác có thể có vật ghi lưu trữ số không tạm thời bao gồm chương trình máy tính được chứa trên đó để thực hiện phương pháp tách nguồn âm thanh, phương pháp có: phương pháp xử lý tín hiệu âm thanh để thu được tín hiệu âm thanh được xử lý, phương pháp có các bước: tính toán các giá trị pha cho các giá trị phổ của chuỗi các khung miền tần số biểu diễn các khung chồng lấp của tín hiệu âm thanh, trong đó các giá trị pha được tính toán dựa trên thông tin về đường bao miền thời gian đích liên quan đến tín hiệu âm thanh được xử lý, để tín hiệu âm thanh được xử lý bao gồm ít nhất trong phép tính gần đúng đường bao miền thời gian đích và đường bao phổ được xác định bởi chuỗi các khung miền tần số, và tạo màn chắn phổ của tín hiệu âm thanh ban đầu để thu được tín hiệu âm thanh được cải biên được nhập vào thiết bị để xử lý; trong đó tín hiệu âm thanh được xử lý là tín hiệu nguồn được tách liên quan đến đường bao miền thời gian đích, khi chương trình máy tính nêu trên được chạy bởi máy tính.

Phương án khác có thể có vật ghi lưu trữ số không tạm thời bao gồm chương trình máy tính được chứa trên đó để thực hiện phương pháp nâng cao băng thông của tín hiệu âm thanh được mã hóa, phương pháp có: tạo ra tín hiệu nâng cao từ băng tín hiệu âm thanh nằm trong tín hiệu được mã hóa; phương pháp xử lý tín hiệu âm thanh để thu được tín hiệu âm thanh được xử lý, phương pháp có các bước: tính toán các giá trị pha cho các giá trị phổ của chuỗi các khung miền tần số biểu diễn các khung chồng lấp của tín hiệu âm thanh, trong đó các giá trị pha được tính toán dựa trên thông tin về đường bao miền thời gian đích liên quan đến tín hiệu âm thanh được xử lý, để tín hiệu âm thanh được xử lý bao gồm ít nhất trong phép tính gần đúng đường bao miền thời gian đích và đường bao phổ được xác định bởi chuỗi các khung miền tần số; trong đó tạo ra bao gồm trích đường bao miền thời gian đích từ sự biểu diễn được mã hóa nằm trong tín hiệu được mã hóa hoặc từ băng tín hiệu âm thanh nằm trong tín hiệu được mã hóa, khi chương trình máy tính nêu trên được chạy bởi máy tính.

Phương án khác có thể có vật ghi lưu trữ số không tạm thời bao gồm chương trình máy tính được chứa trên đó để thực hiện phương pháp mã hóa âm thanh, phương pháp có các bước: mã hóa tín hiệu âm thanh miền thời gian sao cho tín hiệu âm thanh được

mã hóa bao gồm sự biểu diễn chuỗi các khung miền tần số của tín hiệu âm thanh miền thời gian và sự biểu diễn đường bao miền thời gian đích; và xác định đường bao từ tín hiệu âm thanh, và so sánh đường bao với tập hợp các đường bao được định trước để xác định sự biểu diễn đường bao miền thời gian đích trên cơ sở so sánh, khi chương trình máy tính nêu trên được chạy bởi máy tính.

Sáng chế dựa trên việc tìm ra rằng đường bao biên độ miền thời gian đích có thể được áp dụng cho các giá trị phổ của chuỗi các khung miền tần số theo thời gian hoặc miền thời gian. Nói cách khác, pha của tín hiệu có thể được hiệu chỉnh sau sự xử lý tín hiệu sử dụng phép chuyển đổi thời gian - tần số và tần số - thời gian, trong đó biên độ hoặc cường độ của tín hiệu này vẫn được duy trì hoặc được giữ (không đổi). Pha có thể được khôi phục sử dụng, ví dụ, thuật toán lặp như thuật toán được đề xuất bởi Griffin và Lim. Tuy nhiên, sử dụng đường bao miền thời gian đích cải thiện một cách đáng kể chất lượng của sự khôi phục pha, mà dẫn đến số lượng các phép lặp giảm nếu thuật toán lặp được sử dụng. Đường bao miền thời gian đích có thể được tính toán hoặc được làm gần đúng.

Các phương án thể hiện thiết bị để xử lý tín hiệu âm thanh để thu được tín hiệu âm thanh được xử lý. Thiết bị có thể bao gồm bộ tính toán pha để tính toán các giá trị pha cho các giá trị phổ của chuỗi các khung miền tần số biểu diễn các khung chồng lấp của tín hiệu âm thanh. Bộ tính toán pha có thể được tạo cấu hình để tính toán các giá trị pha dựa trên thông tin về đường bao miền thời gian đích liên quan đến tín hiệu âm thanh được xử lý, để tín hiệu âm thanh được xử lý có ít nhất trong phép tính gần đúng đường bao miền thời gian đích và đường bao miền phổ được xác định bởi chuỗi các khung miền tần số. Thông tin về đường bao biên độ miền thời gian đích có thể được áp dụng vào chuỗi các khung miền tần số trong miền thời gian hoặc tần số.

Để vượt qua hạn chế nêu trên của các phương pháp đã biết, các phương án thể hiện kỹ thuật, phương pháp cho thiết bị để bảo toàn tốt hơn các thành phần chuyển tiếp trong các tín hiệu nguồn được khôi phục. Cụ thể, mục đích có thể là giảm các tiếng vang trước mà làm hỏng độ rõ lúc bắt đầu của các sự kiện nốt từ trống và các bộ gõ cũng như piano và ghi-ta.

Các phương án còn thể hiện sự mở rộng hoặc cải thiện cho quy trình khôi phục tín

hiệu bởi Griffin và Lim [1] mà, ví dụ, bảo toàn tốt hơn các thành phần tín hiệu chuyển tiếp. Phương pháp ban đầu ước lượng thông tin pha cần thiết cho sự khôi phục miền thời gian từ cường độ STFT (STFT magnitude - STFTM) bằng cách đi qua đi lại giữa STFT và tín hiệu miền thời gian, chỉ cập nhật thông tin pha, trong khi giữ STFTM cố định. Sự mở rộng được đề xuất hoặc cải thiện thao tác các khôi phục miền thời gian trung gian để giảm tiếng vang trước mà có khả năng đứng trước các chuyển tiếp.

Theo phương án thứ nhất, thông tin về đường bao miền thời gian đích được áp dụng vào chuỗi các khung miền tần số trong miền thời gian. Do đó, phép biến đổi Fourier thời gian ngắn được cải biên (modified Short-Time Fourier Transform - MSTFT) có thể được suy ra từ chuỗi các khung miền tần số. Dựa trên phép biến đổi Fourier thời gian ngắn được cải biên, phép biến đổi Fourier thời gian ngắn nghịch đảo có thể được thực hiện. Vì phép biến đổi Fourier thời gian ngắn nghịch đảo (Inverse Short-Time Fourier Transform - ISTFT) thực hiện quy trình chồng lấp và cộng, các giá trị cường độ và các giá trị pha của MSTFT ban đầu được thay đổi (cập nhật, thích ứng hoặc điều chỉnh). Điều này dẫn đến sự khôi phục miền thời gian trung gian của tín hiệu âm thanh. Hơn nữa, đường bao miền thời gian đích có thể được áp dụng cho sự khôi phục miền thời gian trung gian. Điều này có thể, ví dụ, được thực hiện bằng cách nhân chấp tín hiệu miền thời gian bởi phản hồi xung hoặc bằng cách nhân phổ với hàm biến đổi. Sự khôi phục miền thời gian trung gian của tín hiệu âm thanh có (gần đúng) đường bao miền thời gian đích có thể được chuyển đổi thời gian - tần số sử dụng phép biến đổi Fourier thời gian ngắn (Short-Time Fourier Transform - STFT). Do đó, sự phân tích chồng lấp và/hoặc các cửa sổ tổng hợp có thể được sử dụng.

Thậm chí nếu việc điều biến của đường bao miền thời gian đích không được áp dụng, STFT của sự biểu diễn miền thời gian trung gian của tín hiệu âm thanh có thể sẽ khác với MSTFT sớm hơn do quy trình chồng lấp và cộng trong ISTFT và STFT. Điều này có thể được thực hiện trong thuật toán lặp, trong đó, đối với MSTFT được cập nhật, giá trị pha của toán tử STFT được sử dụng và biên độ tương ứng hoặc giá trị cường độ được loại bỏ. Thay vào đó, như là giá trị cường độ hoặc biên độ đối với MSTFT được cập nhật, các giá trị biên độ ban đầu có thể được sử dụng, vì các tác giả sáng chế giả định rằng giá trị cường độ (hoặc biên độ) được khôi phục chỉ có thông tin pha sai. Do đó, trong mỗi bước lặp, các giá trị pha được tạo thích ứng cho các giá trị pha (ban đầu)

đúng.

Theo phương án thứ hai, đường bao biên độ miền thời gian đích có thể được áp dụng vào chuỗi các khung miền tần số trong miền tần số. Do đó, các bước được thực hiện trước đó trong miền thời gian có thể được chuyển (biến đổi, áp dụng hoặc chuyển đổi) thành miền tần số. Cụ thể, điều này có thể là phép biến đổi miền thời gian của cửa sổ tổng hợp của ISTFT và cửa sổ phân tích của STFT. Điều này có thể dẫn đến sự biểu diễn tần số của các khung lân cận mà sẽ chồng lấp khung hiện thời sau ISTFT và STFT đã được biến đổi thành miền thời gian. Tuy nhiên, phần này được dịch chuyển đến vị trí đúng trong khung hiện thời, và sự bổ sung được thực hiện để suy ra sự biểu diễn miền tần số trung gian của tín hiệu âm thanh. Hơn nữa, đường bao miền thời gian đích có thể được biến đổi thành miền tần số, ví dụ sử dụng STFT, sao cho sự biểu diễn tần số của đường bao miền thời gian đích có thể được áp dụng cho sự biểu diễn miền tần số trung gian. Một lần nữa, quy trình này có thể được thực hiện lặp được lặp lại sử dụng pha được cập nhật của sự biểu diễn miền tần số trung gian có (trong sự gần đúng) đường bao của đường bao miền thời gian đích. Hơn nữa, cường độ ban đầu của MSTFT được sử dụng, vì các tác giả sáng chế giả định rằng cường độ đã được khôi phục một cách hoàn hảo.

Sử dụng thiết bị nêu trên, nhiều phương án khác có thể được giả định có các khả năng khác để suy ra đường bao miền thời gian đích. Các phương án thể hiện bộ giải mã âm thanh bao gồm thiết bị được đề cập trước đó. Bộ giải mã âm thanh có thể nhận tín hiệu âm thanh từ bộ mã hóa âm thanh (được kết hợp). Bộ mã hóa âm thanh có thể phân tích tín hiệu âm thanh để suy ra đường bao miền thời gian đích, ví dụ, đối với mỗi khung thời gian của tín hiệu âm thanh. Đường bao miền thời gian đích được suy ra có thể được so sánh với danh sách được định trước gồm các đường bao miền thời gian đích minh họa. Đường bao miền thời gian đích được định trước mà gần nhất với đường bao miền thời gian đích được tính toán của tín hiệu âm thanh có thể được kết hợp với chuỗi nhất định gồm các bit, ví dụ, chuỗi gồm bốn bit để phân bố 16 đường bao miền thời gian đích khác nhau. Bộ giải mã âm thanh có thể bao gồm các đường bao miền thời gian đích được định trước giống nhau, ví dụ, bảng mã hoặc bảng tìm kiếm, và có thể xác định (đọc, tính toán hoặc tính) đường bao miền thời gian đích được định trước (được mã hóa) bằng chuỗi các bit được truyền dẫn từ bộ mã hóa.

Theo các phương án khác, thiết bị nêu trên có thể là phần của bộ xử lý tách nguồn âm thanh. Bộ xử lý tách nguồn âm thanh có thể sử dụng phép làm gần đúng thô của đường bao miền thời gian đích, vì tín hiệu âm thanh ban đầu chỉ có một nguồn trong số nhiều nguồn của tín hiệu âm thanh (thường) không có sẵn. Do đó, đặc biệt với sự khôi phục chuyển tiếp, phần khung hiện thời lên đến vị trí chuyển tiếp ban đầu có thể bị bắt buộc là 0. Điều này có thể giảm một cách hiệu quả các tiếng vang trước ở phía trước sự chuyển tiếp thường được kết hợp do thuật toán xử lý tín hiệu. Hơn nữa, sự bắt đầu chung có thể được sử dụng như phép làm gần đúng đối với đường bao miền thời gian đích, ví dụ, sự bắt đầu giống nhau cho mỗi khung. Theo phương án khác, sự bắt đầu khác nhau có thể được sử dụng cho các thành phần khác nhau của tín hiệu âm thanh, ví dụ, được suy ra từ danh sách được định trước gồm các sự bắt đầu. Ví dụ, đường bao miền thời gian đích hoặc sự bắt đầu của piano có thể khác với đường bao miền thời gian đích hoặc sự bắt đầu của ghita, chũm chọe, hoặc tiếng nói. Do đó, nguồn hiện thời hoặc thành phần cho tín hiệu âm thanh có thể được phân tích, ví dụ, để phát hiện loại thông tin âm thanh (nhạc cụ, tiếng nói, v.v.) để xác định (theo lý thuyết) phép làm gần đúng phù hợp nhất của đường bao miền thời gian đích. Theo phương án khác, loại thông tin âm thanh có thể được đặt trước (bởi người sử dụng), nếu sự tách nguồn âm thanh, ví dụ, được dự định để tách một hoặc nhiều nhạc cụ (ví dụ, ghi ta, chũm chọe, sáo, hoặc piano) hoặc tiếng nói từ phần còn lại của tín hiệu âm thanh. Dựa trên thiết lập trước, sự bắt đầu tương ứng cho việc ghi rãnh âm thanh được tách biệt hoặc cô lập có thể được chọn.

Theo các phương án, bộ xử lý nâng cao băng thông có thể sử dụng thiết bị được đề cập trước đó. Bộ xử lý nâng cao băng thông sử dụng bộ mã hóa lỗi để mã hóa sự biểu diễn độ phân giải cao của một hoặc nhiều băng của tín hiệu âm thanh. Hơn nữa, các băng mà không được mã hóa sử dụng bộ mã hóa lỗi có thể được làm gần úng trong bộ giải mã nâng cao băng thông sử dụng tham số của bộ mã hóa nâng cao băng thông. Đường bao miền thời gian đích có thể được truyền dẫn, ví dụ, như tham số, bằng bộ mã hóa. Tuy nhiên, theo phương án được ưu tiên, đường bao miền thời gian đích không được truyền dẫn (như tham số) bởi bộ mã hóa. Do đó, đường bao miền thời gian đích có thể được suy ra trực tiếp từ phần được giải mã lỗi hoặc (các) băng tần của tín hiệu âm thanh. Hình dạng hoặc đường bao của phần được giải mã lỗi của tín hiệu âm thanh là sự làm gần đúng tốt cho đường bao miền thời gian của tín hiệu âm thanh ban đầu. Tuy

nhiên, các thành phần tần số cao có thể bị thiếu trong phần được giải mã lỗi của tín hiệu âm thanh dẫn đến đường bao miền thời gian đích mà có thể ít được nêu bật hơn so với đường bao ban đầu. Ví dụ, đường bao miền thời gian đích có thể giống với phiên bản được lọc thông thấp của tín hiệu âm thanh hoặc phần của tín hiệu âm thanh. Tuy nhiên, phép làm gần đúng của đường bao miền thời gian đích từ tín hiệu âm thanh được giải mã lỗi có thể (về trung bình) chính xác hơn so với, ví dụ, sử dụng bảng mã trong đó thông tin về đường bao miền thời gian đích có thể được truyền dẫn từ bộ mã hóa nâng cao bằng thông đến bộ giải mã nâng cao bằng thông.

Theo các phương án khác, sự mở rộng hiệu quả của thuật toán khôi phục tín hiệu lặp đi lặp lại được đề xuất bởi Griffin và Lim sẽ được thể hiện. Sự mở rộng thể hiện bước trung gian nằm trong sự khôi phục lặp đi lặp lại sử dụng phép biến đổi Fourier thời gian ngắn được cải biên. Bước trung gian có thể bắt buộc hình dạng mong muốn hoặc được định trước của tín hiệu mà sẽ được khôi phục. Do đó, đường bao được định trước có thể được áp dụng cho tín hiệu (miền thời gian) được khôi phục, ví dụ, sử dụng phép điều biến biên độ, trong mỗi bước của phép lặp. Cách khác, đường bao có thể được áp dụng cho tín hiệu được khôi phục sử dụng phép nhân chập STFT và đường bao trong miền thời gian - tần số. Phương pháp thứ hai có thể có lợi hoặc có hiệu quả hơn, vì STFT nghịch đảo và STFT có thể được mô phỏng (được thực hiện, được biến đổi hoặc được chuyển) trong miền thời gian - tần số và do đó, các bước này không cần được thực hiện rõ ràng. Hơn nữa, các sự rút gọn, như, ví dụ, sự xử lý chọn lọc theo chuỗi có thể được nhận ra. Hơn nữa sự khởi tạo của các pha (của bước MSTFT thứ nhất) có các giá trị có nghĩa là lợi thế, vì đạt được sự chuyển đổi nhanh hơn.

Trước các phương án được mô tả cụ thể sử dụng các hình vẽ kèm theo, các tác giả sáng chế chỉ ra rằng các phần tử giống nhau hoặc tương đương về chức năng được cho các số tham chiếu giống nhau trong các hình vẽ và phần mô tả được lặp lại cho các phần tử được cung cấp các số tham chiếu giống nhau được đưa ra. Do đó, các phần mô tả được đề xuất cho các phần tử có các số tham chiếu giống nhau được trao đổi lẫn nhau

Mô tả vắn tắt các hình vẽ

Các phương án của sáng chế sẽ được mô tả sau đây viện dẫn đến các hình vẽ đính kèm, trong đó:

Fig.1 thể hiện sơ đồ khối dạng giản lược của thiết bị xử lý tín hiệu âm thanh để thu được tín hiệu âm thanh được xử lý;

Fig.2 thể hiện sơ đồ khối dạng giản lược của thiết bị theo phương án khác sử dụng sự xử lý miền thời gian - tần số hoặc sự xử lý miền tần số;

Fig.3 thể hiện thiết bị theo phương án khác trong sơ đồ khối dạng giản lược sử dụng sự xử lý miền thời gian - tần số;

Fig.4 thể hiện sơ đồ khối dạng giản lược của thiết bị theo phương án khác sử dụng sự xử lý miền tần số.

Fig.5 thể hiện sơ đồ khối dạng giản lược của thiết bị theo phương án khác sử dụng sự xử lý miền thời gian - tần số;

Các hình vẽ từ Fig.6a đến Fig.6d thể hiện biểu đồ giản lược của sự khôi phục chuyển tiếp theo phương án;

Fig.7 thể hiện sơ đồ khối dạng giản lược của thiết bị theo phương án khác sử dụng sự xử lý miền tần số;

Fig.8 thể hiện sơ đồ miền thời gian dạng giản lược minh họa một phân đoạn của tín hiệu âm thanh;

Các hình vẽ từ Fig.9a đến Fig.9c minh họa các sơ đồ của các tín hiệu thành phần chùm chọi khác được tách từ vòng lặp trống minh họa;

Fig.10a và Fig.10b thể hiện sự minh họa dưới dạng giản đồ của hỗn hợp tín hiệu bộ gõ chứa ba nhạc cụ như nguồn cho sự tách nguồn của các vòng lặp trống;

Fig.11a thể hiện sự phát triển của ước số không đồng nhất được chuẩn hóa so với số các lần lặp;

Fig.11b thể hiện sự phát triển của năng lượng tiếng vang trước so với số các lần lặp;

Fig.12a thể hiện sơ đồ dạng giản lược của sự phát triển của ước số không đồng nhất được chuẩn hóa so với số các lần lặp;

Fig.12b thể hiện sự phát triển của năng lượng tiếng vang trước so với số các lần lặp;

Fig.13 thể hiện sơ đồ dạng giản lược của kết quả phép khai triển NMF điển hình, minh họa các mẫu được tách (ba đồ thị phía bên trái) thực chất giống các phiên bản nguyên mẫu của các sự kiện bắt đầu trong V (sơ đồ phía dưới bên phải).

Fig.14a thể hiện sơ đồ dạng giản lược của sự phát triển của ước số đồng nhất được chuẩn hóa so với số các lần lặp;

Fig.14b thể hiện sơ đồ dạng giản lược của sự phát triển của năng lượng tiếng vang trước so với số các lần lặp;

Fig.15 thể hiện bộ mã hóa âm thanh để mã hóa tín hiệu âm thanh theo phương án;

Fig.16 thể hiện bộ giải mã âm thanh bao gồm thiết bị và giao diện đầu vào;

Fig.17 thể hiện tín hiệu âm thanh bao gồm sự biểu diễn của chuỗi các khung miền tần số và sự biểu diễn đường bao miền thời gian đích;

Fig.18 thể hiện sơ đồ khối dạng giản lược của bộ xử lý tách nguồn âm thanh theo phương án;

Fig.19 thể hiện sơ đồ khối dạng giản lược của bộ xử lý nâng cao băng thông theo phương án;

Fig.20 thể hiện sơ đồ miền tần số dạng giản lược minh họa sự nâng cao băng thông;

Fig.21 thể hiện sự biểu diễn dạng giản lược của sự khôi phục miền thời gian (trung gian);

Fig.22 thể hiện sơ đồ khối dạng giản lược của phương pháp xử lý tín hiệu âm thanh để thu được tín hiệu âm thanh được xử lý;

Fig.23 thể hiện sơ đồ khối dạng giản lược của phương pháp giải mã âm thanh;

Fig.24 thể hiện sơ đồ khối dạng giản lược của phương pháp tách nguồn âm thanh;

Fig.25 thể hiện sơ đồ khối dạng giản lược của phương pháp nâng cao băng thông của tín hiệu âm thanh được mã hóa;

Fig.26 thể hiện sơ đồ khối dạng giản lược của phương pháp mã hóa âm thanh;

Mô tả chi tiết sáng chế

Trong phần sau đây, các phương án của sáng chế sẽ được mô tả chi tiết hơn. Các

phần tử được thể hiện trong các hình vẽ tương ứng có chức năng giống nhau hoặc tương tự nhau sẽ có các kí hiệu tham chiếu giống nhau.

Fig.1 thể hiện sơ đồ khối dạng giản lược của thiết bị 2 xử lý tín hiệu âm thanh 4 để thu được tín hiệu âm thanh được xử lý 6. Thiết bị 2 bao gồm bộ tính toán pha 8 để tính toán các giá trị pha 10 cho các giá trị phổ của chuỗi các khung miền tần số 12 biểu diễn các khung chồng lấp của tín hiệu âm thanh 4. Hơn nữa, bộ tính toán pha 8 được tạo cấu hình để tính toán các giá trị pha 10 dựa trên thông tin về đường bao miền thời gian đích 14 liên quan đến tín hiệu âm thanh được xử lý 6, để tín hiệu âm thanh được xử lý 6 có ít nhất trong phép tính gần đúng đường bao biên độ miền thời gian đích 14 và đường bao phổ được xác định bởi chuỗi các khung miền tần số 12. Do đó, bộ tính toán pha 8 có thể được tạo cấu hình để nhận thông tin về đường bao miền thời gian đích hoặc để trích xuất thông tin về đường bao miền thời gian đích từ (sự biểu diễn của) đường bao miền thời gian đích.

Các giá trị phổ của chuỗi các khung miền tần số 10 có thể được tính toán sử dụng phép biến đổi Fourier thời gian ngắn (Short-Time Fourier Transform - STFT) của tín hiệu âm thanh 4. Do đó, STFT có thể sử dụng các cửa sổ phân tích có phạm vi chồng lấp là, ví dụ, 50%, 67%, 75%, hoặc thậm chí nhiều hơn. Nói cách khác, STFT có thể sử dụng kích thước bước nhảy, ví dụ, là một nửa, một phần ba, hoặc một phần tư chiều dài của cửa sổ phân tích.

Thông tin về đường bao miền đích 14 có thể được suy ra sử dụng các phương pháp khác liên quan đến phương án hiện thời hoặc đã được sử dụng. Trong môi trường mã hóa, ví dụ, bộ mã hóa có thể phân tích tín hiệu âm thanh (ban đầu) (trước khi mã hóa) và truyền dẫn, ví dụ, bảng mã hoặc chỉ số bảng tìm kiếm đến bộ giải mã biểu diễn đường bao miền đích được định trước gần với đường bao miền đích được tính toán. Bộ giải mã, có chung bảng mã hoặc bảng tìm kiếm như bộ mã hóa có thể suy ra đường bao miền thời gian đích sử dụng chỉ số bảng mã được nhận.

Trong môi trường nâng cao băng thông, đường bao của sự biểu diễn được giải mã lỗi của tín hiệu âm thanh có thể là sự biểu diễn được giải mã tốt cho đường bao miền thời gian đích ban đầu.

Sự nâng cao băng thông che phủ bất kỳ dạng nào của sự nâng cao băng thông của

tín hiệu được xử lý so với băng thông của tín hiệu đầu vào trước khi xử lý. Một cách nâng cao băng thông là sự thực hiện điền đầy khoảng trống, như điền đầy khoảng trống thông minh như được mô tả trong WO2015010948 hoặc sự điền đầy khoảng trống bán tham số, trong đó các khoản trống phổ là tín hiệu đầu vào được điền đầy hoặc “được nâng cao” bằng các phần phổ khác của tín hiệu đầu vào có hoặc không có sự giúp đỡ của thông tin tham số được truyền dẫn. Cách khác nâng cao băng thông là phép lặp băng phổ (spectral band replication - SBR) như được sử dụng trong HE-AAC (MPEG 4) hoặc các quy trình liên quan, trong đó băng phía trên tần số giao chéo được tạo ra bởi sự xử lý. Ngược lại với sự thực hiện điền đầy khoảng trống, băng thông của tín hiệu lỗi trong SBR được giới hạn, trong khi sự thực hiện điền đầy khoảng trống có tín hiệu lỗi băng đầy đủ. Do đó, sự nâng cao băng thông biểu diễn sự mở rộng băng thông đến các tần số cao hơn tần số giao chéo hoặc sự mở rộng tần số đến các khoảng cách phổ được định vị, so với tần số, dưới tần số tối đa của tín hiệu lỗi.

Hơn nữa, trong môi trường tách nguồn, đường bao miền thời gian đích có thể được làm gần đúng. Điều này có thể là 0 đệm lên vị trí ban đầu của sự chuyển tiếp hoặc sử dụng các sự bắt đầu (khác nhau) như phép làm gần đúng hoặc ước lượng thô của đường bao miền thời gian đích. Nói cách khác, đường bao miền thời gian đích được làm gần đúng có thể được suy ra từ đường bao miền thời gian hiện thời của tín hiệu miền thời gian trung gian bằng cách buộc đường bao miền thời gian hiện thời bằng 0 từ lúc bắt đầu khung hoặc phần của tín hiệu âm thanh lên đến vị trí ban đầu của sự chuyển tiếp. Theo các phương án khác, đường bao miền thời gian hiện thời được điều biến bởi một hoặc nhiều sự bắt đầu (được định trước). Sự bắt đầu có thể được cố định cho (toàn bộ) sự xử lý của tín hiệu âm thanh hoặc, nói cách khác, được chọn một lần trước (hoặc đối với) sự xử lý khung (thời gian) thứ nhất hoặc phần của tín hiệu âm thanh.

(Sự làm gần đúng) của đường bao miền thời gian đích có thể được sử dụng để tạo hình dạng của tín hiệu âm thanh được xử lý, ví dụ, sử dụng phép điều biến biên độ hoặc phép nhân, sao cho tín hiệu âm thanh được xử lý có ít nhất gần ứng đường bao miền thời gian đích. Tuy nhiên, đường bao phổ của tín hiệu âm thanh được xử lý được xác định bởi chuỗi các khung miền tần số, vì đường bao miền thời gian đích bao gồm chủ yếu các thành phần tần số thấp khi được so sánh với phổ của chuỗi các khung miền tần số, sao cho phần lớn tần số duy trì không đổi.

Fig.2 thể hiện sơ đồ khối dạng giản lược của thiết bị 2 theo phương án khác. Thiết bị của Fig.2 thể hiện bộ tính toán pha 8 bao gồm bộ xử lý phép lặp 16 để thực hiện thuật toán lặp để tính toán, bắt đầu từ các giá trị pha ban đầu 18, các giá trị pha 10 cho các giá trị phổ sử dụng đích tối ưu hóa yêu cầu sự đồng nhất của các khối chồng lấp trong phạm vi chồng lấp. Hơn nữa, bộ xử lý lặp 16 được tạo cấu hình để sử dụng, trong bước lặp khác, sự ước lượng pha được cập nhật 20, phụ thuộc vào đường bao miền thời gian đích. Nói cách khác, sự tính toán các giá trị pha 10 có thể được thực hiện sử dụng thuật toán lặp được thực hiện bởi bộ xử lý phép lặp 16. Do đó, các trị số cường độ của chuỗi các khung miền tần số có thể được biết đến và duy trì không đổi. Bắt đầu từ các giá trị pha ban đầu 18, bộ xử lý phép lặp có thể cập nhật lặp đi lặp lại các giá trị pha đối với các giá trị phổ sử dụng, sau mỗi phép lặp, sự ước lượng pha được cập nhật 20 để thực hiện các phép lặp.

Đích tối ưu hóa có thể, ví dụ, là số các phép lặp. Theo các phương án khác, đích tối ưu hóa có thể là ngưỡng, trong đó các giá trị pha được cập nhật chỉ đến mức nhỏ khi được so với các giá trị pha của bước lặp trước đó, hoặc đích tối ưu hóa có thể là sự chênh lệch của cường độ không đổi (ban đầu) của chuỗi các khung miền tần số khi được so sánh với cường độ của các giá trị phổ sau quy trình lặp. Do đó, các giá trị phổ có thể được cải thiện hoặc cập nhật sao cho phổ tần số riêng lẻ của các phần này của các khung của tín hiệu âm thanh bằng nhau hoặc ít nhất chỉ khác ở mức nhỏ. Nói cách khác, tất cả các phần khung của các khung chồng lấp của tín hiệu âm thanh chồng lấp một khung khác nên có sự biểu diễn tần số giống nhau hoặc tương tự nhau.

Theo các phương án, bộ tính toán pha được tạo cấu hình để thực hiện thuật toán lặp phù hợp với quy trình khôi phục tín hiệu lặp đi lặp lại bởi Griffin và Lim. Các phương án (cụ thể hơn) khác được thể hiện với các hình vẽ sau đây. Tại đó, bộ xử lý lặp sẽ chia nhỏ hoặc được thay thế bởi chuỗi các khối xử lý, cụ thể các bộ chuyển đổi tần số thành thời gian 22, bộ điều biến biên độ 24, và bộ chuyển đổi thời gian thành tần số 26. Để thuận lợi hơn, bộ xử lý phép lặp 16 thường được chỉ ra (không rõ ràng) trong các hình vẽ, tuy nhiên, các khối xử lý nêu trên thực hiện các toán tử giống nhau như bộ xử lý phép lặp 16, hoặc bộ xử lý phép lặp giám sát hoặc điều khiển điều kiện kết thúc (hoặc điều kiện thoát) của sự xử lý phép lặp, như, ví dụ, đích tối ưu hóa. Hơn nữa, bộ xử lý phép lặp có thể thực hiện các toán tử theo sự xử lý miền tần số được thể hiện, ví dụ, trên

Fig.4 và Fig.7.

Fig.3 thể hiện thiết bị theo phương án khác trong sơ đồ khối dạng giản lược. Thiết bị 2 bao gồm bộ chuyển đổi tần số thành thời gian 22, bộ điều biến biên độ 24, và bộ chuyển đổi thời gian thành tần số 26, trong đó phép chuyển đổi tần số thành thời gian và/hoặc phép chuyển đổi thời gian thành tần số có thể thực hiện quy trình chồng lấp và cộng. Bộ chuyển đổi tần số thành thời gian 22 có thể tính toán sự khôi phục miền thời gian trung gian 28 của tín hiệu âm thanh 4 từ chuỗi các khung miền tần số 12 và sự ước lượng giá trị pha ban đầu 18 hoặc các ước lượng giá trị pha 10 của bước lặp trước đó. Bộ điều biến biên độ 24 có thể điều biến sự khôi phục miền thời gian 28 sử dụng (thông tin về) đường bao miền thời gian đích 14 để thu được tín hiệu âm thanh được điều biến biên độ 30. Hơn nữa, bộ chuyển đổi thời gian thành tần số được tạo cấu hình để chuyển đổi tín hiệu được điều biến biên độ 30 thành chuỗi các khung miền tần số 32 có các giá trị pha 10. Do đó, bộ tính toán pha 8 được tạo cấu hình để sử dụng, cho bước lặp tiếp theo, các giá trị pha 10 (của chuỗi các khung miền tần số) và các giá trị phổ của chuỗi các khung miền tần số (mà không phải chuỗi khác gồm các khung miền tần số). Nói cách khác, bộ tính toán pha sử dụng các giá trị pha của chuỗi các khung miền tần số 32 khác sau mỗi bước lặp. Các giá trị cường độ của chuỗi khác gồm các khung miền tần số có thể được loại bỏ hoặc không được sử dụng cho sự xử lý khác. Hơn nữa, bộ tính toán pha 8 sử dụng các giá trị cường độ của chuỗi (ban đầu) gồm các khung miền tần số 12, vì các tác giả sáng chế giả định rằng các giá trị cường độ đã được khôi phục (một cách hoàn hảo).

Khái quát hơn, bộ tính toán pha 8 được tạo cấu hình để áp dụng phép điều biến biên độ, ví dụ, trong bộ điều biến biên độ 22, cho sự khôi phục miền thời gian trung gian 28 của tín hiệu âm thanh 4, dựa trên đường bao miền thời gian đích 14. Phép điều biến biên độ có thể được thực hiện sử dụng phép điều biến băng phụ đơn, phép điều biến băng phụ đôi có hoặc không có sự truyền dẫn vật mang được chặn hoặc sử dụng phép nhân của đường bao miền thời gian đích với sự khôi phục miền thời gian trung gian của tín hiệu âm thanh. Sự ước lượng giá trị pha ban đầu có thể là giá trị pha của tín hiệu âm thanh, giá trị được chọn (tùy ý) như, ví dụ, 0, giá trị ngẫu nhiên, hoặc ước lượng của pha của băng tần của tín hiệu âm thanh, hoặc pha của nguồn tín hiệu âm thanh, ví dụ, khi sử dụng phép tách nguồn âm thanh.

Theo các phương án khác, bộ tính toán pha 8 được tạo cấu hình để phát ra sự khôi phục miền thời gian trung gian 28 của tín hiệu âm thanh 4 như tín hiệu âm thanh được xử lý 6, khi điều kiện xác định phép lặp (ví dụ điều kiện kết thúc phép lặp) thỏa mãn. Điều kiện xác định phép lặp có thể liên quan gần với đích tối ưu hóa và có thể xác định độ lệch tối đa của đích tối ưu hóa đến giá trị tối ưu hóa hiện thời. Hơn nữa, điều kiện xác định phép lặp có thể là số lần lặp (tối đa), độ lệch (tối đa) của cường độ của chuỗi khác gồm các khung miền tần số 32 khi được so sánh với cường độ của chuỗi các khung miền tần số 12, hoặc tác động cập nhật (tối đa) của các giá trị pha 10, nằm giữa khung hiện thời và khung trước đó.

Fig.4 thể hiện sơ đồ khối dạng giản lược của thiết bị 2 theo phương án, mà có thể là phương án hay thế khi được so sánh với phương án của Fig.3 Bộ tính toán pha 8 được tạo cấu hình để áp dụng phép nhân chập 34 của sự biểu diễn phổ 14' và ít nhất một đường bao miền thời gian đích 14 và ít nhất một sự biểu diễn miền tần số trung gian, hoặc các phần hoặc các băng được lựa chọn hoặc chỉ phần thông cao hoặc chỉ một vài phần dải giữa của ít nhất một đường bao miền thời gian đích 14 hoặc ít nhất một sự biểu diễn miền tần số trung gian 28' của tín hiệu âm thanh 4. Nói cách khác, sự xử lý của Fig.3 có thể được thực hiện trong miền tần số thay vì miền thời gian. Do đó, đường bao miền thời gian đích 14, cụ thể hơn, sự biểu diễn tần số 14' của đường bao này, có thể được áp dụng cho sự biểu diễn miền tần số 28' sử dụng phép nhân chập thay vì phép điều biến biên độ. Tuy nhiên, ý tưởng là một lần nữa sử dụng cường độ (ban đầu) của chuỗi các khung miền tần số cho mỗi phép lặp và hơn nữa, sau khi sử dụng giá trị pha ban đầu 18 trong bước lặp thứ nhất, sử dụng các ước lượng giá trị pha được cập nhật 10 đối với mỗi bước lặp khác. Nói cách khác, bộ tính toán pha được tạo cấu hình để sử dụng các giá trị pha 10 thu được bởi phép nhân chập 34 như các ước lượng giá trị pha được cập nhật đối với bước lặp tiếp theo. Hơn nữa, thiết bị có thể bao gồm bộ chuyển đổi đường bao đích 36 để chuyển đổi đường bao miền thời gian đích thành miền phổ. Hơn nữa, thiết bị 2 có thể bao gồm bộ chuyển đổi tần số thành thời gian 38 để tính toán sự khôi phục miền thời gian 28 từ sự khôi phục miền tần số trung gian 28' sử dụng các ước lượng giá trị pha 10 thu được từ bước lặp gần nhất và chuỗi các khung miền tần số 12. Nói cách khác, sự biểu diễn miền tần số trung gian 28' có thể bao gồm các giá trị cường độ của chuỗi các khung miền tần số và các giá trị pha 10 của các ước lượng giá

trị pha được cập nhật. Sự khôi phục miền thời gian 28 có thể là tín hiệu âm thanh được xử lý 6 hoặc ít nhất phần của tín hiệu âm thanh được xử lý 6. Phần có thể đề cập, ví dụ, đến số lượng các băng tần giảm khi được so sánh với tổng số các băng tần của tín hiệu âm thanh được xử lý hoặc tín hiệu âm thanh 4.

Theo các phương án khác, các bộ tính toán pha 8 bao gồm bộ xử lý phép nhân chập 40. Bộ xử lý phép nhân chập 40 có thể áp dụng hạch nhân chập, hạch dịch chuyển, và/hoặc thuật toán khung thêm vào giữa để thu được sự biểu diễn miền tần số trung gian 28' của tín hiệu âm thanh 4. Nói cách khác, bộ xử lý phép nhân chập có thể xử lý chuỗi các khung miền tần số 12, trong đó bộ xử lý nhân chập 40 có thể được tạo cấu hình để áp dụng phép tương đương miền tần số của quy trình chồng lấp và cộng miền thời gian cho chuỗi các khung miền tần số 12 trong miền tần số để xác định sự khôi phục miền tần số trung gian. Theo các phương án khác, bộ xử lý phép nhân chập được tạo cấu hình để xác định, dựa trên khung miền tần số hiện thời, phần của các khung miền tần số liên kế mà đóng góp cho khung miền tần số hiện thời sau phép chồng lấp và cộng miền thời gian được thực hiện trong miền tần số. Hơn nữa, bộ xử lý phép nhân chập 40 có thể còn xác định vị trí chồng lấp của phần của khung miền tần số liên kế nằm trong khung tần số hiện thời và để thực hiện sự bổ sung các vị trí của các khung miền tần số liên kế với khung miền tần số hiện thời tại các vị trí chồng lấp. Theo phương án khác, bộ xử lý phép nhân chập 40 được tạo cấu hình để biến đổi thời gian thành tần số phép tổng hợp miền thời gian và cửa sổ phân tích miền thời gian để xác định phần của khung miền tần số liên kế, mà đóng góp cho khung miền tần số hiện thời sau phép chồng lấp và cộng miền thời gian được thực hiện trong miền tần số. Hơn nữa, bộ xử lý nhân chập còn được tạo cấu hình để dịch chuyển phần của khung miền tần số liên kế để chồng lấp vị trí nằm trong khung miền tần số hiện thời và để áp dụng phần của khung miền tần số liên kế đến khung hiện thời tại vị trí chồng lấp.

Nói cách khác, quy trình miền thời gian được thể hiện trên Fig.3 có thể được chuyển (biến đổi, áp dụng hoặc chuyển đổi) thành miền tần số. Do đó, các cửa sổ tổng hợp và phân tích của bộ chuyển đổi tần số thành thời gian 22 và bộ chuyển đổi thời gian thành tần số 26 có thể được chuyển (biến đổi, áp dụng hoặc chuyển đổi) thành miền tần số. Sự biểu diễn miền tần số (kết quả) của các cửa sổ tổng hợp và phân tích xác định (hoặc cắt) các phần của các khung liên kế đến khung hiện thời mà có thể đã chồng lấp

trong quy trình chồng lấp và cộng trong miền thời gian. Hơn nữa, các phần cắt được dịch chuyển đến vị trí đúng nằm trong khung hiện thời và được cộng vào khung hiện thời sao cho phép biến đổi tần số thành thời gian miền thời gian và phép biến đổi thời gian thành tần số được thực hiện trong miền tần số. Điều này là có lợi, vì phép biến đổi tín hiệu rõ ràng có thể bị bỏ qua hoặc không được thực hiện, mà có thể tăng hiệu quả tính toán của bộ tính toán pha 8 và thiết bị 2.

Fig.5 thể hiện sơ đồ khối dạng giản lược của thiết bị 2 theo phương án khác tập trung vào sự khôi phục tín hiệu của các kênh hoặc các băng được tách của tín hiệu âm thanh 4. Do đó, tín hiệu âm thanh 4 trong miền thời gian có thể được biến đổi thành chuỗi các khung miền tần số 12 biểu diễn các khung chồng lấp của tín hiệu âm thanh 4 sử dụng bộ chuyển đổi thời gian - tần số, ví dụ STFT 42. Từ đó, bộ ước lượng cường độ cải biên 44' có thể trích cường độ 44 của chuỗi các khung miền tần số hoặc các thành phần hoặc các tín hiệu thành phần của chuỗi các khung miền tần số. Hơn nữa, ước lượng pha ban đầu 18 có thể được tính toán từ chuỗi các khung miền tần số 12 sử dụng bộ ước lượng pha ban đầu 18' hoặc bộ ước lượng pha ban đầu 18' có thể chọn, ví dụ, ước lượng pha tùy ý 18, mà không suy ra từ chuỗi các khung miền tần số 12. Dựa trên cường độ 44 của chuỗi các khung miền tần số 12 và ước lượng pha ban đầu 18, MSTFT 12' có thể được tính toán như chuỗi ban đầu gồm các khung miền tần số 12'' có cường độ được khôi phục (hoàn hảo) 44 mà duy trì không đổi trong sự xử lý khác, và chỉ ước lượng pha ban đầu 18. Ước lượng pha ban đầu 18 được cập nhật sử dụng bộ tính toán pha 8.

Trong bước khác, bộ chuyển đổi tần số thành thời gian 22, ví dụ, STFT nghịch đảo (inverse STFT - ISTFT), có thể tính toán sự khôi phục miền thời gian trung gian 28 của chuỗi (ban đầu) gồm các khung miền tần số 12''. Sự khôi phục miền thời gian trung gian 28 có thể được điều biến biên độ, ví dụ, được nhân, với đường bao đích, hoặc chính xác hơn, đường bao miền thời gian đích 14. Bộ chuyển đổi thời gian thành tần số 26, ví dụ, STFT, có thể tính toán chuỗi khác gồm các khung miền tần số 32 có các giá trị pha 10. MSTFT 12' có thể sử dụng bộ ước lượng pha được cập nhật 10 và cường độ 44 của chuỗi các khung miền tần số 12 trong chuỗi được cập nhật gồm các khung miền tần số. Thuật toán lặp này có thể được thực hiện hoặc được lặp L lần trong, ví dụ, bộ xử lý phép lặp 16, mà có thể thực hiện các bước xử lý được đề cập trên đây của bộ tính toán pha 8. Ví dụ, sau quy trình phép lặp được hoàn thiện, sự khôi phục miền thời gian 28'' được

suy ra từ sự khôi phục miền thời gian trung gian 28.

Nói cách khác, sau đây, mô hình ký hiệu và tín hiệu được thể hiện và phương pháp khôi phục tín hiệu được sử dụng được mô tả. Sau đó, phép mở rộng để bảo toàn chuyển tiếp trong phương pháp LSEE-MSTFTM được thể hiện kết hợp với ví dụ minh họa.

Tín hiệu miền thời gian giá trị thực, rời rạc $x : \mathbb{Z} \rightarrow \mathbb{R}$ được xem xét để là hỗn hợp của các tín hiệu thành phần hiện thời. Mục đích là khai triển x thành tín hiệu đích chuyển tiếp $x^t : \mathbb{Z} \rightarrow \mathbb{R}$ và tín hiệu thành phần dư $x^r : \mathbb{Z} \rightarrow \mathbb{R}$ sao cho

$$x \approx x^t + x^r. \quad (1')$$

Lưu ý rằng phép khai triển được đặt như phép làm gần đúng, vì tiêu điểm là chất lượng cảm quan cải thiện của tín hiệu chuyển tiếp x^t và được chấp nhận rằng siêu vị trí của x^t và x^r có thể không sử dụng chính xác x ban đầu. Bây giờ, giả định rằng x^t chứa chính xác một chuyển tiếp, mà vị trí theo thời gian $n_0 \in \mathbb{Z}$ của chuyển tiếp này đã biết. Để $\mathcal{X}(m, k)$ với $m, k \in \mathbb{Z}$ là ngắn TF được tạo giá trị phức tại khung thời gian thứ m và hệ số phổ thứ k của phép biến đổi Fourier thời gian ngắn (Short-Time Fourier Transform - STFT). Hệ số được tính toán bởi

$$\mathcal{X}(m, k) := \sum_{n=0}^{N-1} x(n + mH)w(n) \exp(-2\pi i kn/N), \quad (2')$$

trong đó $w : [0 : N - 1] \rightarrow \mathbb{R}$ là hàm cửa sổ thích hợp có kích thước khối $N \in \mathbb{N}$ và $H \in \mathbb{N}$ là tham số kích thước bước nhảy. Để đơn giản, cũng có thể viết là $\mathcal{X} = \text{STFT}(x)$. Từ $\mathcal{X}$, ảnh phổ cường độ $\mathcal{A}$ và ảnh phổ pha φ được suy ra như:

$$\mathcal{A}(m, k) := |\mathcal{X}(m, k)|, \quad (3')$$

$$\varphi(m, k) := \angle \mathcal{X}(m, k) \quad (4')$$

với $\varphi(m, k) \in [0, 2\pi)$. Giả định rằng, qua một số quy trình tách nguồn phù hợp, ước lượng STFT được cải biên (modified STFT - MSTFT) $\mathcal{X}^t$ là khả thi, mà biểu diễn tín hiệu thành phần chuyển tiếp. Cụ thể hơn, được thiết lập $\mathcal{X}^t := \mathcal{A}^t \odot \exp(i\varphi^t)$, trong đó $\mathcal{A}^t$ và φ^t là các ước lượng của cường độ, lần lượt ảnh phổ pha, và toán tử $\odot$ biểu thị phép nhân theo phần tử. Sự khôi phục miền thời gian của $\mathcal{X}^t$ đạt được bằng cách đầu tiên áp dụng phép biến đổi Fourier rời rạc (Discrete Fourier Transform - DFT) cho mỗi khung

phổ, tạo ra tập hợp các tín hiệu thời gian trung gian $y_m, m \in \mathbb{Z}$, được định nghĩa bởi

$$y_m(n) := \frac{1}{N} \sum_{k=0}^{N-1} \mathcal{A}^t(m, k) \exp(2\pi i k n / N). \quad (5')$$

đối với $n \in [0 : N - 1]$ và $y_m(n) := 0$ đối với $n \in \mathbb{Z} \setminus [0 : N - 1]$. Thứ hai, phương pháp khôi phục lỗi bình phương nhỏ nhất như

$$x^t(n) := \frac{\sum_{m \in \mathbb{Z}} y_m(n - mH) w(n - mH)}{\sum_{m \in \mathbb{Z}} w(n - mH)^2}, \quad (6')$$

$n \in \mathbb{Z}$ được áp dụng, trong đó cửa sổ phân tích w được tái sử dụng như cửa sổ tổng hợp. Để đơn giản, quy trình này được biểu thị là $x^t := \text{iSTFT}(\mathcal{A}^t)$ (được đề cập đến như LSEE-MSTFT trong [8]).

Vì sự ước lượng $\mathcal{A}^t$ cho thu được trong miền thời gian - tần số (time-frequency - TF), không thể giả định rằng x^t là tín hiệu đồng nhất. Trong thực tế, có thể bắt gặp sự nhòe chuyển tiếp và âm vang trước trong x^t . Điều này đặc biệt đúng đối với N lớn. Để sửa chữa vấn đề này, đề xuất việc tinh chỉnh lặp đi lặp lại $\mathcal{A}^t$ bằng quy trình sau, trong đó chỉ số lặp $\ell = 0, 1, 2, \dots, L \in \mathbb{N}$ được đưa ra và vị trí chuyển tiếp n_0 đã cho được sử dụng. $\mathcal{A}^t$ đã cho và $\varphi^{(0)}$, ước lượng M ban đầu của thành phần tín hiệu chuyển tiếp được đưa ra như $(\mathcal{A}^t)^{(0)} := \mathcal{A}^t \odot \exp(i\varphi^{(0)})$ và các bước sau đây được lặp lại cho $\ell = 0, 1, 2, \dots, L$

1. $(x^t)^{(\ell+1)} := \text{iSTFT}((\mathcal{A}^t)^{(0)})$ qua (5') và (6')
2. Bắt buộc $(x^t)^{(\ell+1)}(n) := 0$ for $n \in \mathbb{Z}, n < n_0$
3. $\varphi^{(\ell+1)} := \angle \text{STFT}((x^t)^{(\ell+1)})$ qua (2') và (4')
4. $(\mathcal{A}^t)^{(\ell+1)} := \mathcal{A}^t \odot \exp(i\varphi^{(\ell+1)})$

Phương án của Fig.5 có thể được mô tả nhiều hơn bình thường, sử dụng các tín hiệu thành phần được biểu thị với $\mathcal{A}_c$ thay vì các tín hiệu chuyển tiếp được mô tả với $\mathcal{A}^t$. Nhìn chung, đối với tất cả các phương án được mô tả, các tín hiệu được biểu thị bởi chỉ số dưới c có thể được thay thế bằng tín hiệu tín hiệu tương ứng được biểu thị bởi chỉ số dưới t và ngược lại. Chỉ số dưới c biểu thị tín hiệu thành phần trong đó chỉ số dưới t biểu thị tín hiệu chuyển tiếp, mà có thể là tín hiệu thành phần. Tuy nhiên, tín hiệu có chỉ

số dưới t có thể cũng được thay thế bởi tín hiệu (chung hơn) có chỉ số dưới c . Các phương án được mô tả đối với các tín hiệu chuyển tiếp không được giới hạn đến tín hiệu chuyển tiếp và nhờ đó có thể được áp dụng cho bất kỳ tín hiệu thành phần khác. Ví dụ, $\mathcal{A}^t$ có thể được thay thế bởi $\mathcal{A}_c$ và ngược lại.

Do đó, tín hiệu miền thời gian rời rạc, giá trị thực $x : \mathbb{Z} \rightarrow \mathbb{R}$ được xem xét để là hỗn hợp tuyến tính $x := \sum_{c=1}^C x_c$ of $C \in \mathbb{N}$ các tín hiệu thành phần x_c tương ứng với các nguồn riêng lẻ (ví dụ, các nhạc cụ). Như được thể hiện trên Fig.10a, mỗi tín hiệu thành phần chứa ít nhất một sự kiện âm thanh chuyển tiếp được tạo ra bởi nhạc cụ tương ứng (trong trường hợp minh họa hiện thời, bằng cách đánh trống). Hơn nữa, giả định rằng việc phiên âm biểu tượng là có sẵn mà cụ thể hóa thời gian bắt đầu (tức là, vị trí chuyển tiếp) và loại nhạc cụ cho mỗi sự kiện âm thanh. Từ sự phiên âm đó, tổng số lượng các sự kiện bắt đầu S được suy ra cũng như số các nhạc cụ đơn nhất C . Mục đích là để tách các tín hiệu thành phần riêng lẻ x_c từ hỗn hợp x như được thể hiện trên Fig.10. Đối với các mục đích đánh giá, giả định có các tín hiệu thành phần “oracle” (tức là, đúng) x_c có sẵn. x được khai triển trong miền TF, để đạt được mục đích này, STFT được sử dụng như sau. Để $\mathcal{X}(m, k)$ là hệ số TF giá trị phức tại khung thứ m và ngăn phổ thứ k . Hệ số được tính toán bởi

$$\mathcal{X}(m, k) := \sum_{n=0}^{N-1} x(n + mH)w(n) \exp(-2\pi i kn/N), \quad (1)$$

trong đó $w : [0 : N-1] \rightarrow \mathbb{R}$ là hàm cửa sổ thích hợp có kích thước khối $N \in \mathbb{N}$, và $H \in \mathbb{N}$ là tham số kích thước bước nhảy. Số lượng các ngăn tần số $K = N/2$ và số lượng các khung phổ $M \in [1 : M]$ được xác định bởi các mẫu tín hiệu có sẵn. Để đơn giản, có thể viết là $\mathcal{X} = \text{STFT}(x)$. Theo [2], $\mathcal{X}$ được gọi là STFT đồng nhất vì nó là tập hợp các số phức đã thu được từ tín hiệu miền thời gian thực x qua (1). Ngược lại, STFT không đồng nhất là tập hợp các số phức mà không thu được tín hiệu miền thời gian thực. Từ $\mathcal{X}$, ảnh phổ cường độ $\mathcal{A}$ và ảnh phổ pha φ được suy ra như

$$\mathcal{A}(m, k) := |\mathcal{X}(m, k)|, \quad (2)$$

$$\varphi(m, k) := \angle \mathcal{X}(m, k), \quad (3)$$

với $\varphi(m, k) \in [0, 2\pi)$.

Đề $V := \mathcal{A}^T \in \mathbb{R}_{\geq 0}^{K \times M}$ là ma trận không âm giữ phiên bản được chuyển vị của ảnh phổ cường độ $\mathcal{A}$ của hỗn hợp. Mục đích là để khai triển V thành các ảnh phổ cường độ thành phần V_c mà tương ứng với các nhạc cụ khác biệt như được thể hiện trên Fig.10b. Từ bây giờ, giả định rằng một số bộ ước lượng oracle trích xuất $\mathcal{A}_c := V_c^T$ mong muốn. Một phương pháp khả thi để ước lượng các cường độ thành phần sử dụng công nghệ khai triển tình trạng kỹ thuật sẽ được mô tả sau. Để khôi phục tín hiệu thành phần cụ thể x_c , các tác giả sáng chế thiết lập $\mathcal{X}_c := \mathcal{A}_c \odot \exp(i\varphi_c)$, trong đó $\mathcal{A}_c = V_c^T$ và φ_c là ước lượng của ảnh phổ pha thành phần. Thực tiễn chung là sử dụng thông tin pha hỗn hợp φ như ước lượng cho φ_c và để đảo ngược MSTFT thu được qua phương pháp khôi phục LSEE-MSTFT từ [1]. Phương pháp đầu tiên áp dụng phép biến đổi Fourier rời rạc (Discrete Fourier Transform - DFT) cho mỗi khung phổ trong $\mathcal{X}_c$, sử dụng tập hợp các tín hiệu thời gian trung gian y_m , với $m \in [0 : M - 1]$, được xác định bởi

$$y_m(n) := \frac{1}{N} \sum_{k=0}^{N-1} \mathcal{X}_c(m, k) \exp(2\pi i k n / N), \quad (4)$$

đối với $n \in [0 : N - 1]$ và $y_m(n) := 0$ for $n \in \mathbb{Z} \setminus [0 : N - 1]$. Thứ hai, phương pháp khôi phục lỗi bình phương nhỏ thu được bằng cách

$$x_c(n) := \frac{\sum_{m \in \mathbb{Z}} y_m(n - mH) w(n - mH)}{\sum_{m \in \mathbb{Z}} w(n - mH)^2}, \quad (5)$$

$n \in \mathbb{Z}$, trong đó cửa sổ phân tích w được tái sử dụng như cửa sổ tổng hợp. Để đơn giản, quy trình này được biểu thị là $x_c = \text{iSTFT}(\mathcal{X}_c)$ (được đề cập đến như LSEE-MSTFT trong [1]).

Vì MSTFT $\mathcal{X}_c$ được tạo ra trong miền TF, phải giả định rằng nó có thể là STFT, tức là, không tồn tại tín hiệu miền thời gian thực x_c điền đầy $\mathcal{X}_c = \text{STFT}(x_c)$. Nói theo trực giác, sự tác động qua lại lẫn nhau giữa cường độ và pha có thể sai lệch ngay khi cường độ trong các ngăn TF nhất định được cải biên. Trong thực tế, việc không đồng nhất có thể dẫn đến sự nhòe chuyển tiếp và các tiếng vang trước trong x_c , cụ thể đối với N lớn.

Để sửa chữa vấn đề này, các tác giả sáng chế đề xuất việc tối thiểu hóa lặp đi lặp lại sự không đồng nhất của $\mathcal{X}_c$ bằng phép mở rộng sau của quy trình LSEE-MSTFTM

[1]. Bây giờ, có thể giả định rằng $\mathcal{X}_c$ chứa chính xác một sự kiện bắt đầu chuyển tiếp, mà vị trí chính xác theo thời gian n_0 đã biết. Bây giờ, chỉ số phép lặp $\ell = 0, 1, 2, \dots, L \in \mathbb{N}$ được đưa ra. $\mathcal{A}_c$ đã cho và một số ước lượng pha ban đầu $(\varphi_c)^{(0)}$, ước lượng STFT ban đầu của tín hiệu thành phần đích $(\mathcal{X}_c)^{(0)} := \mathcal{A}_c \odot \exp(i(\varphi_c)^{(0)})$ được đưa ra và các bước sau được lặp lại đối với $\ell = 0, 1, 2, \dots, L$.

1. $(x_c)^{(\ell+1)} := \text{iSTFT} \left((\mathcal{X}_c)^{(\ell)} \right)$ qua (4) và (5)
2. Bắt buộc $(x_c)^{(\ell+1)}(n) := 0$ với $n \in \mathbb{Z}, n < n_0$
3. $(\varphi_c)^{(\ell+1)} := \angle \text{STFT} \left((x_c)^{(\ell+1)} \right)$ qua (1) và (3)
4. $(\mathcal{X}_c)^{(\ell+1)} := \mathcal{A}_c \odot \exp \left(i(\varphi_c)^{(\ell+1)} \right)$

Theo các phương án, điểm có lợi của các phương án được mô tả, bộ mã hóa hoặc giải mã là bước trung gian 2, mà bắt buộc các ràng buộc chuyển tiếp trong quy trình LSEE-MSTFTM.

Các hình vẽ từ Fig.6a đến Fig.6d thể hiện sơ đồ giản lược của sự khôi phục chuyển tiếp theo phương án biểu thị tín hiệu miền thời gian 46, đường bao tín hiệu phân tích 48 và vị trí chuyển tiếp 50. Fig.6 minh họa phương pháp được đề xuất hoặc thiết bị với tín hiệu thành phần đích 46, được phủ bằng đường bao của tín hiệu phân tích 48 của nó trên Fig.6a. Tín hiệu minh họa thể hiện diễn biến chuyển tiếp hoặc thành phần tín hiệu chuyển tiếp quanh n_0 50 khi các chuyển tiếp dạng sóng từ im lặng đến đường hình sin suy giảm theo hàm mũ hoặc sóng sin. Fig.6b thể hiện sự khôi phục miền thời gian thu được từ iSTFT với $(\varphi_c)^{(0)} = 0$ (tức là, pha 0 cho tất cả các ngăn TF). Qua sự trở ngại phá hủy của các khung chồng lấp, sự chuyển tiếp hoàn toàn bị phá hủy, biên độ của đường hình sin tăng mạnh mẽ và đường bao gần như phẳng. Fig.6c thể hiện sự khôi phục với độ nhòe chuyển tiếp được phát âm sau $L = 200$ phép lặp LSEE-MSTFTM. Fig.6d thể hiện sự chuyển tiếp được khôi phục sau $L = 200$ phép lặp của phương pháp được đề xuất gần với tín hiệu ban đầu hơn. Các gợn sóng nhỏ là có thể nhìn thấy trong đường bao phía trước n_0 , nhưng toàn thể sự khôi phục gần với tín hiệu ban đầu hơn. Trong các bản ghi thực tế, thường xuất hiện sự kiện bắt đầu đa chuyển tiếp xuyên suốt tín hiệu. Trong trường hợp này, một người có thể áp dụng phương pháp được đề xuất để tạo tín hiệu phân trích được định vị giữa các chuyển tiếp (tương ứng với các phần bắt đầu) như được

thể hiện trên Fig.9.

Fig.7 thể hiện sơ đồ khối dạng giản lược của thiết bị 2 theo phương án khác. Tương tự với Fig.4, bộ tính toán pha thực hiện sự tính toán pha trong miền tần số. Việc xử lý miền tần số có thể bằng với sự xử lý miền thời gian được mô tả với phương án được thể hiện trên Fig.5. Một lần nữa, tín hiệu miền thời gian 4 có thể được biến đổi tần số-thời gian sử dụng (bộ thực hiện) STFT 42 để suy ra chuỗi các khung miền tần số 12. Từ đó, bộ ước lượng cường độ được cải biên 44' có thể suy ra cường độ được cải biên 44 từ chuỗi các khung miền tần số 12. Bộ ước lượng pha ban đầu 18' có thể suy ra ước lượng pha ban đầu 18 từ chuỗi các khung miền tần số hoặc có thể cung cấp, ví dụ, ước lượng pha ban đầu bất kì. Sử dụng ước lượng cường độ được cải biên và ước lượng pha ban đầu, MSTFT 12' tính toán hoặc xác định chuỗi ban đầu của các khung miền tần số 12'', mà nhận các giá trị pha được cập nhật sau mỗi bước lặp. Khác với các phương án của Fig.5 là chuỗi (ban đầu) gồm các khung miền tần số 12'' trong bộ tính toán pha 8. Dựa trên các cửa sổ tổng hợp và phân tích miền thời gian, ví dụ, cửa sổ tổng hợp và phân tích được sử dụng trong ISTFT 22 hoặc STFT 26 trên Fig.5, bộ tính toán hạch nhân chập 52' có thể tính toán hạch nhân chập 52 sử dụng sự biểu diễn miền tần số của các cửa sổ tổng hợp và phân tích. Hạch nhân chập cắt (cắt ra hoặc sử dụng) các phần gồm các khung lân cận hoặc liền kề của khung miền tần số hiện thời mà sẽ chồng lấp khung hiện thời sử dụng sự chồng lấp và cộng trong ISTFT 22. Bộ tính toán dịch chuyển hạch 54' có thể tính toán hạch dịch chuyển 52 và áp dụng hạch dịch chuyển 52 cho các phần của các khung miền tần số liền kề để dịch chuyển các phần này đến vị trí chồng lấp đúng của khung miền tần số hiện thời. Điều này có thể mô phỏng toán tử chồng lấp của quy trình chồng lấp và cộng của ISTFT 22. Hơn nữa, khối 56 thực hiện sự bổ sung quy trình chồng lấp và cộng và thêm các phần chồng lấp của các khung liền kề với chu kỳ khung ở giữa. Phép tính toán và áp dụng hạch nhân chập, phép tính toán và áp dụng hạch dịch chuyển, và sự bổ sung trong khối 56 có thể được thực hiện trong bộ xử lý nhân chập 40. Đầu ra của bộ xử lý nhân chập 40 có thể là sự khôi phục miền tần số trung gian 28' của chuỗi gồm các khung miền tần số 12 hoặc chuỗi ban đầu gồm các khung miền tần số 12''. Sự khôi phục miền tần số trung gian 28' có thể được nhân chập (theo khung) với sự biểu diễn miền tần số của đường bao đích 14 sử dụng phép nhân chập 34. Đầu ra của phép nhân chập 34 có thể là chuỗi khác gồm các khung miền tần số 32' có các giá trị pha 10.

Các giá trị pha 10 thay thế ước lượng pha ban đầu 18 trong MSTFT 12' trong bước lặp. Phép lặp có thể được thực hiện L lần sử dụng bộ xử lý phép lặp 15. Sau khi quy trình lặp được dừng, hoặc tại điểm thời gian nhất định trong quy trình lặp, sự khôi phục miền tần số cuối cùng 28''' có thể được suy ra từ bộ xử lý nhân chập 40. Sự khôi phục miền tần số cuối cùng 28''' có thể là phép lặp khôi phục miền tần số trung gian 28' của bước lặp gần nhất. Sử dụng bộ chuyển đổi tần số thành thời gian 38, ví dụ ISTFT, sự khôi phục miền thời gian 28'' có thể thu được, mà có thể là tín hiệu âm thanh được xử lý 6.

Nói cách khác, có lợi khi áp dụng bước trung gian trong phép lặp LSEE-MSTFTM. Nó sẽ bắt tất cả các mẫu trước sự chuyển tiếp là 0 trước khi tính toán STFT lần nữa để thu được ước lượng được cập nhật của các pha $\varphi^{(l+1)}$. Hạn chế này có thể là bắt buộc trực tiếp trong miền TF. Do đó, thiết lập một số việc cần thiết trước có thể có lợi. Đầu tiên, việc tiêu chuẩn hóa cho tổng các hàm tạo cửa sổ dịch chuyển thời gian và bình phương trong mẫu số của (6) có thể được loại bỏ bằng cách áp đặt những ràng buộc nhất định trên w và H (ví dụ, sử dụng cửa sổ Hann đối xứng và yêu cầu phần dư $Q = N/H$ là cơ số 4 [2]). Số lượng các ngăn phổ (tùy thuộc vào phép liên hợp) đơn nhất mỗi khung là $K = N/2$, và đôi số tần số được đánh giá đối với $k \in [-K : K]$. Tập trung cho thời điểm mà khung phổ đơn, thuật toán áp dụng liên tục iSTFT và STFT một lần nữa có thể được biểu thị trong miền TF như sự xếp chồng các phần góp phổ được gán trọng số từ các khung đứng trước và tiếp theo. Chỉ các khung mà chồng lấp với khung chính giữa cần được xem xét. Điều này được thể hiện bởi chỉ số khung lân cận $q \in [-(Q-1) : (Q-1)]$. Hai hạch TF được dựng, hạch thứ nhất là hạch nhân chập.

$$\alpha(q, k) := \frac{1}{N} \sum_{n=0}^{N-1} w(n)w(n+qH) \exp(-2\pi i k n / N). \quad (7')$$

đã thu lại DFT của sản phẩm theo phần tử của cửa sổ tổng hợp với phiên bản được cắt và dịch chuyển thời gian của cửa sổ phân tích. Hạch thứ hai là hạch nhân

$$\beta(q, k) := \exp(2\pi i k (-q/Q)), \quad (8')$$

mà cần để dịch chuyển sự đóng góp từ các khung lân cận đến vị trí đúng bên trong khung chính giữa. Các hạch được áp dụng cho mỗi ngăn TF trong chuỗi.

$$(\mathcal{X}^*(m, k))^{(\ell+1)} := \sum_{q=-(Q-1)}^{Q-1} \beta(q, k) \sum_{p=-K}^K \alpha(q, p) (\mathcal{X}^*(m, k+p))^{(\ell)} \quad (9')$$

Bây giờ, sự khôi phục chuyển tiếp được đề xuất có thể được bao gồm theo cách đơn giản bởi phép toán nhân chập thứ hai mà chỉ cần được áp dụng cho các khung mà n_0 được định vị. Các hạch nhân chập tương ứng có thể được lấy theo khung từ STFT của hàm Heavyside được dịch chuyển phù hợp.

$$\mathcal{H}_{n_0}(n) := \begin{cases} 0, & n < n_0, \\ 1, & n \geq n_0, \end{cases} \quad (10')$$

Lưu ý rằng, ngoài việc sử dụng hàm được tạo dạng của bước này, các tác giả sáng chế đề xuất sử dụng STFT của các tín hiệu đường bao biên độ miền thời gian đường bao được tạo dạng tùy ý. Các tác giả sáng chế nêu ra rằng phạm vi rộng các hạn chế khôi phục có thể bị áp đặt qua phép điều biến tín hiệu phù hợp trong miền thời gian, phép nhân chập tương ứng trong miền TF.

Như được thể hiện trên [4], tải tính toán gồm áp dụng các toán tử miền tần số có thể được giảm bằng cách cắt hạch nhân chập α thành số lượng nhỏ hơn gồm các hệ số chính giữa. Điều này có thể được thúc đẩy bằng cách đánh giá qua quan sát, mà các hệ số rõ rệt nhất được định vị quanh $k = 0$. Các thử nghiệm đã thể hiện rằng sự khôi phục TF đã rất gần với sự khôi phục miền thời gian nếu α được cắt theo hướng tần số đến $k \in [-3 : +3]$. Ngoài ra, α là Hermitian, nếu các hàm cửa sổ được chọn một cách phù hợp. Dựa trên các phép đối xứng phức liên hợp, các phép nhân phức và do đó năng lượng xử lý, có thể dư. Hơn nữa, không cần thiết xem xét cập nhật pha của mỗi ngăn tần số. Thay vào đó, người ta có thể lựa chọn phân số của các ngăn mà thể hiện cường độ cao nhất, và áp dụng (9') chỉ với các ngăn này, vì chúng sẽ chi phối sự khôi phục. Như sẽ được thể hiện sau đây, ước đoán thứ nhất hợp lý sẽ giúp tăng tốc sự hội tụ của sự khôi phục.

Để đánh giá, sự khôi phục thông thường (được biểu thị như GL) LSEE-MSTFTM được so sánh với phương pháp được đề xuất (được biểu thị như TR) dưới hai chiến lược khởi tạo khác nhau đối với $(\mathcal{X}^*)^{(0)}$. Dưới đây, tập hợp dữ liệu được sử dụng, sự tạo ra mục kiểm tra, và các metric được đánh giá được sử dụng sẽ được mô tả.

Trong tất cả các thử nghiệm, tập hợp dữ liệu “IDMT-SMT-Trống” có sẵn công khai được sử dụng. Trong tập hợp con “WaveDrum02”, có 60 vòng lặp trống, mỗi vòng được cho như bản ghi rãnh đơn được tách biệt một cách hoàn hảo (tức là, các tín hiệu thành phần oracle) của ba nhạc cụ, trống gỗ, trống lười, và chũm chọe. Tất cả 3x60 bản ghi là ở định dạng PCM WAV không được nén với tốc độ lấy mẫu 44:1kHz, 16 Bit, đơn âm. Trộn tất cả ba rãnh đơn cùng nhau, 60 tín hiệu hỗn hợp thu được. Ngoài ra, các thời gian bắt đầu và sau đó, n_0 gần đúng trong số tất cả các lúc bắt đầu là có sẵn mỗi nhạc cụ riêng lẻ. Sử dụng thông tin này, tập hợp thử nghiệm của sự kiện bắt đầu được khôi phục bằng cách lấy các phần trích từ các hỗn hợp, mỗi phần được định vị giữa các sự bắt đầu liên tiếp của nhạc cụ đích. Làm như vậy, N mẫu phía trước mỗi phần trích được đệm 0. Lý do cơ bản là để thêm vào trước một cách thận trọng phần của sự im lặng trước vị trí chuyển tiếp cục bộ. Trong phần đó, sự ảnh hưởng suy giảm của lúc bắt đầu nốt trước có thể được tạo quy tắc và có khả năng xảy ra các tiếng vang trước có thể được đo. Lần lượt, điều này dẫn đến sự dịch chuyển ảo của vị trí chuyển tiếp cục bộ đến $n_0 + N$ (mà được biểu thị một lần nữa là n_0 cho sự thuận tiện về ký hiệu).

Fig.8 thể hiện sơ đồ miền thời gian dạng giản lược minh họa một phân đoạn hoặc khung của tín hiệu âm thanh hoặc vật thử nghiệm. Fig.8 thể hiện tín hiệu hỗn hợp 61a, tín hiệu chũm chọe đích 61b, sự khôi phục sử dụng LSEE-MSTFTM 61c được so với sự khôi phục chuyển tiếp 61d, cả hai thu được sau 200 phép lặp được áp dụng mỗi phần trích bắt đầu 60, mà là, ví dụ, phần giữa các đường nét đứt 60' và 60''. Tín hiệu trộn 61a thể hiện rõ sự ảnh hưởng của trống gỗ và trống lười đối với tín hiệu chũm chọe đích 61b.

Các hình vẽ từ Fig.9a đến Fig.9c minh họa các sơ đồ giản lược của các tín hiệu thành phần chũm chọe khác nhau của ví dụ vòng lặp trống. Vị trí chuyển tiếp n_0 62 được biểu thị bởi đường nét liền, trong đó các đường biên phần trích 60' và 60'' được biểu thị bởi các đường nét đứt. Fig.9a thể hiện tín hiệu hỗn hợp phía trên so với tín hiệu chũm chọe oracle phía dưới. Fig.9b thể hiện tín hiệu chũm chọe thu được từ sự khởi tạo với cường độ oracle và chu kỳ pha 0. Sự khôi phục sau L bằng 200 phép lặp của GL được thể hiện tại phía trên của Fig.9b so với TR tại phía dưới của Fig.9b. Fig.9c thể hiện tín hiệu chũm chọe thu được từ sự khởi tạo với cường độ trên cơ sở NMFD trong sự xử lý trên cơ sở NMFD pha 0 sẽ được mô tả viện dẫn đến (phần mô tả của) các hình vẽ từ

Fig.12 đến Fig.14. Sự khôi phục sau L bằng 200 phép lặp của GL được thể hiện tại phía trên của Fig.9c và TR tại phía dưới của Fig.9c. Vì phép khai triển thực hiện rất tốt đối với ví dụ vòng lặp trống, hầu hết không có sự chênh lệch về thị giác có thể nhận ra giữa Fig.9b và Fig.9c.

Fig.10 thể hiện sự minh họa dưới dạng giản đồ của tín hiệu. Fig.10a biểu thị tín hiệu trộn x 64a như tổng của $c = 3$ tín hiệu thành phần x_c , mỗi tín hiệu chứa chuỗi các mẫu âm trống tổng hợp, ví dụ, từ trống điện Roland TR808. x_1 64a''' biểu thị trống gõ, x_2 64a'' biểu thị trống lầy, và x_3 64a' biểu thị chũm chọe. Fig.10b thể hiện sự biểu diễn thời gian - tần số của ảnh phổ cường độ V của hỗn hợp và $c = 3$ ảnh phổ cường độ thành phần V_c . Để rõ ràng hơn, trục tần số được lấy mẫu lại đối với việc đặt cách theo lôga và các cường độ đã được nén theo lôga. Hơn nữa, các sự biểu diễn thời gian - tần số của các tín hiệu 64a được biểu thị với tín hiệu tham chiếu 64b. Hơn nữa, trên Fig.9, các đường biên phần trích được điều chỉnh được hình dung bởi các đường nét đứt và được dịch chuyển thực tế n_0 bởi nét liền. Vì các vòng lặp trống là các nhịp thực tế, các phần trích thể hiện mức xếp chồng khác nhau với các nhạc cụ trống còn lại được chơi đồng thời. Trên Fig.9a, hỗn hợp (phía trên) thể hiện sự ảnh hưởng rõ ràng của trống gõ so với tín hiệu chũm chọe tách biệt (phía dưới). Để so sánh, hai biểu đồ phía trên trên Fig.10a thể hiện phiên bản được phóng lớn của hỗn hợp x và thành phần chũm chọe x_3 của tín hiệu ví dụ được sử dụng. Trên biểu đồ phía dưới, có thể thấy trống gõ x_1 tách biệt. Được lấy mẫu từ, ví dụ, trống máy tính Roland TR 808 và giống đường hình sin suy giảm.

Dưới đây, các ảnh đánh giá sẽ được thể hiện cho các cảnh thử nghiệm khác nhau, trong đó hai trường hợp thử nghiệm để khởi tạo MSTFT được sử dụng. Trường hợp 1 sử dụng ước lượng pha ban đầu $(\varphi_c)^{(0)} := \varphi^{\text{Mix}}$ và ước lượng cường độ cố định $\mathcal{A}_c := \mathcal{A}_c^{\text{Oracle}}$. Theo ký hiệu chuyển tiếp, trường hợp 1 sử dụng ước lượng pha ban đầu $(\varphi)^{(0)} := \varphi_{\text{Mix}}$, và ước lượng cường độ cố định $\mathcal{A}^{\text{r}} := \mathcal{A}_{\text{Oracle}}^{\text{r}}$. Nói cách khác, thông tin pha của tín hiệu được tách hoặc tín hiệu riêng phần được lấy từ pha của tín hiệu âm thanh hỗn hợp, thay vào đó, ví dụ, pha của tín hiệu được tách hoặc tín hiệu riêng phần. Hơn nữa, trường hợp 2 sử dụng ước lượng pha ban đầu $(\varphi_c)^{(0)} := 0$ và ước lượng cường độ cố định $\mathcal{A}_c := \mathcal{A}_c^{\text{Oracle}}$. Theo ký hiệu chuyển tiếp, trường hợp 2 như ước lượng pha ban đầu $(\varphi)^{(0)} := 0$ và ước lượng cường độ cố định $\mathcal{A}^{\text{r}} := \mathcal{A}_{\text{Oracle}}^{\text{r}}$. Ở đây, ước lượng pha

ban đầu được khởi tạo sử dụng giá trị 0 (tùy ý), thậm chí kết quả thể hiện trên Fig.6b có thể thu được. Hơn nữa, cả hai trường hợp thử nghiệm sử dụng các giá trị biên độ của tín hiệu được tách hoặc riêng phần của tín hiệu âm thanh. Một lần nữa, có thể thấy rằng ký hiệu có thể được áp dụng chung.

$G((\mathcal{X}_c)^{(\ell)}) := \text{STFT}(\text{iSTFT}((\mathcal{X}_c)^{(\ell)}))$ được đưa ra để biểu thị ứng dụng liên tiếp iSTFT và STFT (lỗi đến thuật toán LSEE-MSTFTM) trên $(\mathcal{X}_c)^{(\ell)}$. Theo sau [10], tại mỗi phép lặp ℓ ước số đồng nhất được tiêu chuẩn hóa (normalized consistency measure - NCM) được tính toán là

$$C((\mathcal{X}_c)^{(\ell)}, \mathcal{X}_c^{\text{Oracle}}) := 10 \log_{10} \frac{\|G((\mathcal{X}_c)^{(\ell)}) - \mathcal{X}_c^{\text{Oracle}}\|^2}{\|\mathcal{X}_c^{\text{Oracle}}\|^2}, \quad (6)$$

đối với cả hai trường hợp thử nghiệm. Như ước số chuyên dụng hơn cho sự khôi phục chuyển tiếp, năng lượng tiếng vang trước được tính toán như

$$E((x_c)^{(\ell)}) := \sum_{n=n_0-N}^{n_1} |(x_c)^{(\ell)}(n)|^2, \quad (7)$$

từ phần giữa lúc bắt đầu phân trích và vị trí chuyển tiếp trong các sự khôi phục tín hiệu thành phần miền thời gian,, trung gian $(x_c)^{(\ell)} := \text{iSTFT}((\mathcal{X}_c)^{(\ell)})$ đối với cả hai trường hợp.

Fig.11a thể hiện phép khai căn ước số đồng nhất được tiêu chuẩn hóa so với số các phép lặp. Fig.11b thể hiện phép khai căn năng lượng tiếng vang trước so với số các phép lặp. Các đường cong thể hiện các phân trích thử nghiệm tổng thể trung bình. Hơn nữa, các kết quả được suy ra từ việc sử dụng thuật toán GL được biểu thị bởi các đường nét đứt, trong đó các kết quả được suy ra từ thuật toán TR được biểu thị sử dụng các đường nét liền. Hơn nữa, sự khởi tạo của trường hợp 1 được biểu thị với số tham chiếu 66a, 66a', trong đó các đường cong được suy ra sử dụng sự khởi tạo của trường hợp 2 được biểu thị với ký hiệu tham chiếu 66b, 66b'. Các đường cong của Fig.11 được suy ra bằng cách tính toán STFT của mỗi phân trích hỗn hợp thông qua (1) với $h = 1024$ và $n = 4096$ và biểu thị chúng như $\mathcal{X}_{\text{Mix}}$. Như đích tham chiếu, phân trích tương tự được lấy, và đệm 0 tương tự được áp dụng, tại thời điểm này từ rãnh đơn của mỗi nhạc cụ riêng, biểu thị STFT thu được là $\mathcal{X}_{\text{Orig}}^t$. Tín hiệu thành phần tương ứng là $\mathcal{X}_c^{\text{Oracle}}$. $L = 200$ phép

lập của cả hai LSEE-MSTFTM (GL) và phương pháp được đề xuất hoặc thiết bị (TR) được sử dụng.

Phép khai căn cả hai ước số chất lượng từ (11) và (12) so với $\mathcal{L}$ được thể hiện trên Fig.11. Sơ đồ (a) biểu thị rằng, trung bình, phương pháp được đề xuất (TR) thực hiện bằng với LSEE-MSTFTM (GL) về mặt giảm tính đồng nhất. Trong cả hai trường hợp, hành vi tương ứng tương tự của các số đo cho TR (đường nét liền) và GL (đường nét đứt) có thể được quan sát. Như được mong đợi, các đường cong 66a, 66a' (trường hợp 1) bắt đầu tại độ không đồng nhất ban đầu thấp hơn các đường cong 66b, 66b' (trường hợp 2), mà rõ ràng do sự khởi tạo với pha hỗn hợp φ_{Mix} . Sơ đồ 11b thể hiện lợi ích của TR cho việc giảm tiếng vang trước. Trong cả hai trường hợp thử nghiệm, các ước số TR 66a 66b (các đường nét liền) thể hiện năng lượng tiếng vang trước khoảng 20dB thấp hơn so với các ước số GL (đường nét đứt). Một lần nữa, $(x^{\text{r}})^{(0)}$ ban đầu đồng nhất hơn của trường hợp 1 66a, 66a' có thể thể hiện sự bắt đầu phân đầu đáng kể về sự giảm tiếng vang trước so với trường hợp 2 66b, 66b'. Bất ngờ là, sự xử lý TR được đề xuất được áp dụng cho trường hợp 2 làm tốt hơn GL được áp dụng cho trường hợp 1 về sự giảm tiếng vang trước cho $L > 100$. Từ các kết quả này, có thể được suy ra mà đủ để áp dụng chỉ một vài phép lặp (ví dụ, $L < 20$) của phương pháp được đề xuất trong các cảnh mà pha ban đầu hợp lý và ước lượng cường độ có sẵn. Tuy nhiên, có thể được áp dụng nhiều phép lặp (ví dụ, $L < 200$) trong trường hợp ước lượng cường độ tốt kết hợp với ước lượng pha yếu và ngược lại là có sẵn. Trên Fig.8, các phiên bản khác nhau của phân đoạn từ một vật thử nghiệm của trường hợp thử nghiệm 2 được thể hiện. Sự khôi phục TR 61d thể hiện rõ ràng các tiếng vang trước giảm so sánh với sự khôi phục với LSEE-MSTFTM 61c. Tín hiệu chũm chọe tham chiếu mà tín hiệu hỗn hợp 61a được thể hiện ở trên.

Tuy nhiên, các hình sau đây được suy ra sử dụng kích thước bước nhảy khác nhau và chiều dài cửa sổ khác nhau như được mô tả dưới đây.

Đối với phân tích hỗn hợp, STFT được tính toán qua (1) với $H = 512$ và $N = 2048$ và được biểu thị như $\mathcal{X}^{\text{Mix}}$. Vì tất cả các vật thử nghiệm có 44.1kHz, độ phân giải tần số là xấp xỉ 21,5Hz và độ phân giải theo thời gian là xấp xỉ 11,6ms. Cửa sổ Hann có kích thước N được sử dụng cho w . Như đích tham chiếu, phân tích tương tự được lấy, và

đệm 0 tương tự được áp dụng, nhưng tại thời điểm này từ rãnh đơn của mỗi nhạc cụ trông riêng, biểu thị STFT thu được là $\mathcal{X}_c^{\text{Oracle}}$. Tiếp theo, hai trường hợp khác nhau cho sự khởi tạo của $(\mathcal{X}_c)^{(0)}$ được định nghĩa chi tiết ở trên. Sử dụng các thiết lập này, sự không đồng nhất của $(\mathcal{X}_c)^{(0)}$ thu được được kì vọng là sẽ thấp hơn trong trường hợp 1 so với trường hợp 2. Biết rằng tồn tại $\mathcal{X}_c^{\text{Oracle}}$ đồng nhất, $L = 200$ phép lặp của cả hai LSEE-MSTFTM (GL) và phương pháp hoặc thiết bị (TR) được đề xuất đã đi qua.

Fig.12a thể hiện sơ đồ dạng giản lược của sự phát triển của ước số đồng nhất được chuẩn hóa so với số các lần lặp. Fig.12b thể hiện phép khai căn năng lượng tiếng vang trước so với số các phép lặp. Các đường cong thể hiện các phần trích thử nghiệm tổng thể trung bình. Nói cách khác, Fig.12 thể hiện phép khai căn cả hai ước số chất lượng từ (6) và (7) đối với ℓ . Fig.12a biểu thị rằng, trung bình, phương pháp được đề xuất (TR) thực hiện bằng với LSEE-MSTFTM (GL) về mặt giảm tính đồng nhất. Trong cả hai trường hợp thử nghiệm, các đường cong cho TR (nét liền) và GL (nét đứt) hầu như không phân biệt được, mà biểu thị rằng phương pháp mới, nghĩa là phương pháp hoặc thiết bị, thể hiện các đặc điểm hội tụ tương tự như phương pháp ban đầu. Như được mong đợi, các đường cong 66a, 66a' (Trường hợp 1) bắt đầu tại độ không đồng nhất ban đầu thấp hơn các đường cong 66b, 66b' (Trường hợp 2), mà rõ ràng do sự khởi tạo với pha hỗn hợp φ^{Mix} . Fig.12b thể hiện lợi ích của TR cho việc giảm tiếng vang trước. Trong cả hai trường hợp thử nghiệm, năng lượng tiếng vang trước cho TR (các đường nét liền) là khoảng 15dB thấp hơn và thể hiện sự sụt giảm mạnh hơn so với GL (đường nét đứt). Một lần nữa, $(\mathcal{X}_c)^{(0)}$ ban đầu đồng nhất hơn của Trường hợp 1 66a, 66a' thể hiện sự bắt đầu phần đầu đáng kể về sự giảm tiếng vang trước so với Trường hợp 2 66b, 66b'. Từ các kết quả này, suy ra rằng đủ để áp dụng chỉ một số phép lặp (ví dụ, $L < 20$) của phương pháp được đề xuất trong các ảnh trong đó pha ban đầu và ước lượng cường độ hợp lý là khả thi. Tuy nhiên, việc áp dụng nhiều phép lặp (ví dụ, $L < 200$) có thể có lợi trong trường hợp ước lượng cường độ tốt kết hợp với ước lượng pha yếu và ngược lại là hiện có.

Phần sau đây sẽ mô tả các phương án làm thế nào để áp dụng phương pháp hoặc thiết bị khôi phục chuyển tiếp được đề xuất trong cảnh khai triển âm thanh được thông báo điểm. Mục đích là sự trích các âm trống tách biệt khỏi các bản ghi trống phức điệu

với sự bảo tồn chuyển tiếp nâng cao. Ngược lại với các điều kiện phòng thí nghiệm lý tưởng được sử dụng trước đây, các ảnh phổ cường độ của các tín hiệu thành phần từ hỗn hợp được ước lượng. Để đạt được mục đích này, phép mở cuộn yếu tố ma trận không âm (Non-Negative Matrix Factor Deconvolution - NMFD) [3, 4] có thể được sử dụng như kỹ thuật khai triển. Các phương án mô tả chiến thuật để áp đặt các hạn chế được thông báo điểm trên NMFD. Cuối cùng, các thử nghiệm được lặp lại dưới các điều kiện thực tế hơn và các sự quan sát được thảo luận.

Sau đây, phương pháp NMFD được sử dụng để khai triển sự biểu diễn TF của x được mô tả ngắn gọn. Như được biểu thị, nhiều phương pháp tách thay thế tồn tại. Các công trình trước đó [3, 4] được áp dụng thành công NMFD, phiên bản nhân chập của NMF, đối với sự tách âm trống. Nói theo trực giác, mô hình nhân chập hoặc nằm dưới, chập mà tất cả các sự kiện âm thanh trong một tín hiệu trong số các tín hiệu thành phần có thể được giải thích bởi sự kiện nguyên mẫu mà đóng vai trò như phản hồi xung đến một số sự kích hoạt liên quan đến sự bắt đầu (ví dụ, đánh trống cụ thể). Trên Fig.10b, một người có thể thấy loại hành vi này trong thành phần chũm chọe V3. Do đó, tất cả các nấc của 8 sự kiện bắt đầu trông giống hoặc không giống các bản sao của nhau mà có thể được giải thích bằng cách chèn sự kiện nguyên mẫu tại mỗi vị trí bắt đầu.

NMF có thể được sử dụng để tính toán phép nhân tử hóa $V \approx W \cdot H$, trong đó các cột của $W \in \mathbb{R}_{\geq 0}^{K \times C}$ thể hiện các hàm cơ bản phổ (cũng được gọi là mẫu) và các hàng của $H \in \mathbb{R}_{\geq 0}^{C \times M}$ chứa các độ khuếch đại biến thiên theo thời gian (cũng được gọi là các sự kích hoạt). NMFD mở rộng mô hình này tới trường hợp nhân chập bằng cách sử dụng các mẫu hai chiều mà mỗi cơ số trong số C cơ số phổ có thể được hiểu là mẫu thông tin ảnh phổ cường độ gồm có $T \ll M$ khung phổ. Để đạt được mục đích này, phép làm gần đúng ảnh phổ nhân chập $V \approx \Lambda$ được tạo mô hình như

$$\Lambda := \sum_{\tau=0}^{T-1} W_{\tau} \cdot \overset{\tau \rightarrow}{H}, \quad (8)$$

trong đó

$(\cdot) \overset{\tau \rightarrow}{\cdot}$ biểu thị thuật toán dịch chuyển khung. Như trước đây, mỗi cột trong W_{τ}

$\in \mathbb{R}_{\geq 0}^{K \times C}$ thể hiện cơ số phổ của thành phần cụ thể, nhưng các phiên bản T thời gian khác của W_τ là có sẵn. Bằng cách nối chuỗi cột cụ thể từ tất cả các phiên bản của W_τ , có thể thu được ảnh phổ cường độ nguyên bản như được thể hiện trên Fig.13. NMFD thường bắt đầu với sự khởi tạo phù hợp của các ma trận $(W_\tau)^{(0)}$ và $(H)^{(0)}$. Tiếp theo, các ma trận này được cập nhật lặp đi lặp lại để giảm thiểu hóa ước số khoảng cách phù hợp giữa phép làm gần đúng nhân chập Λ và V .

Fig.13 thể hiện các mẫu NMFD và các sự kích hoạt được tính toán cho bản ghi trống ví dụ từ Fig.10. Ảnh phổ cường độ V được thể hiện trong sơ đồ bên phải phía dưới. Ba sơ đồ còn lại trên các biểu đồ này là các mẫu phổ trong W_τ mà được tách qua NMFD. Các kích hoạt tương ứng 78 của chúng và sự khởi tạo được thông báo điểm $70b$ $(H)^{(0)}$ được thể hiện trong ba sơ đồ phía trên.

Sự khởi tạo của $(W_\tau)^{(0)}$ và $(H)^{(0)}$ là các cách thức để hạn chế các mức độ tự do trong các phép lặp NMFD và áp đặt sự hội tụ cho giải pháp ý nghĩa về âm nhạc mong muốn. Một khả năng là áp đặt các hạn chế được thông báo điểm được suy ra từ sự phiên âm biểu tượng, được cân chỉnh thời gian. Để đạt được mục đích này, các hàng riêng lẻ của $(H)^{(0)}$ được khởi tạo như sau: Mỗi khung tương ứng với sự bắt đầu của nhạc cụ trống tương ứng được khởi tạo với xung của biên độ đơn vị, tất cả các khung còn lại với hằng số nhỏ. Sau đó, bộ lọc trung bình động hàm mũ không tuyến tính được áp dụng để tạo mô hình sự suy giảm ngắn đặc thù của sự kiện trống. Kết quả 70 của sự khởi tạo này ước thể hiện như đường cong 70b trong ba sơ đồ phía trên của Fig.13.

Các kết quả tách tốt nhất có thể thu được bởi sự khởi tạo được thông báo điểm của cả hai mẫu và các sự kích hoạt. Cho sự tách các nhạc cụ được tạo độ cao (ví dụ, piano), chuỗi họa âm cao nguyên mẫu có thể được dựng trong $(W_\tau)^{(0)}$. Đối với trống, khó khăn để tạo mô hình các cơ số phổ nguyên mẫu. Do đó, được đề xuất để khởi tạo các cơ số với ảnh phổ được tính trung bình hoặc tìm thừa số của các âm trống tách biệt [21, 22, 4]. Tuy nhiên, sự thay thế đơn giản được sử dụng mà đầu tiên tính toán NMF thông thường mà các sự kích hoạt của chúng H và các mẫu W được khởi tạo bởi $(H)^{(0)}$ được thông báo điểm và thiết lập $(W)^{(0)} := 1$.

Với các thiết lập này, các mẫu nhân tử hóa thu được thường là phép làm gần đúng

khá tốt của phổ trung bình của mỗi nhạc cụ trống liên quan. Phép tái tạo đơn giản các phổ này cho tất cả $\tau \in [0 : T - 1]$ đóng vai trò như sự khởi tạo tốt cho các ảnh phổ mẫu. Sau một số phép lặp NMFD, mỗi ảnh phổ mẫu thường tương ứng với ảnh phổ nguyên mẫu của nhạc cụ trống tương ứng và mỗi hàm kích hoạt tương ứng với sự kích hoạt khứ nhân chập của tất cả các sự kiện mà nhạc cụ trống cụ thể của nhạc cụ trống cụ thể qua bản ghi. Kết quả khai triển đặc thù được thể hiện Fig.13, trong đó người ta có thể thấy rằng các mẫu được trích (ba biểu đồ ngoài cùng bên trái) giống các phiên bản nguyên mẫu của các sự kiện bắt đầu trong V (sơ đồ bên phải phía dưới). Hơn nữa, vị trí của xung trong H được trích 70a (ba sơ đồ trên cùng) gần với cực đại của phép khởi tạo được thông báo điểm.

Sau đây, các tác giả sáng chế mô tả cách xử lý thêm các kết quả NMFD để trích các thành phần mong muốn. Để $H \in \mathbb{R}_{\geq 0}^{C \times M}$ là ma trận kích hoạt được nghiên cứu bởi NMFD. Sau đó, đối với mỗi $c \in [1 : C]$ ma trận $H_c \in \mathbb{R}_{\geq 0}^{C \times M}$ được định nghĩa bằng cách thiết lập tất cả các phần tử về 0 ngoại trừ hàng thứ c mà chứa các kích hoạt mong muốn được tìm thấy trước đó qua NMFD. Ảnh phổ cường độ thành phần thứ c được làm gần đúng bởi $\Lambda_c := \sum_{\tau=0}^{T-1} W_{\tau} \cdot \vec{H}_c^{\tau}$. Vì mô hình NMFD tạo ra chỉ phép làm gần đúng hạng thấp của V , các sắc thái phổ có thể được thu tốt. Để giải quyết vấn đề này, thực tiễn chung là tính toán các màn chắn mềm mà có được hiểu như ma trận gán trọng số phản ánh sự đóng góp của Λ_c cho hỗn hợp V . Màn chắn tương ứng với thành phần mong muốn có thể được tính toán như $M_c := \Lambda_c \oslash \left(\epsilon + \sum_{c=1}^C \Lambda_c \right)$, trong đó $\oslash$ biểu thị sự chia theo từng phần tử và ϵ là hằng số dương nhỏ để tránh sự chia bởi 0. Ước lượng trên cơ sở màn chắn của ảnh phổ cường độ thành phần thu được như $V_c := V \odot M_c$, với $\odot$ biểu thị phép nhân theo từng phần tử. Quy trình thường được đề cập đến như phép lọc Wiener.

Sau đây, thử nghiệm trước đó của Fig.12a, Fig.12b về cơ bản được lặp. Các tham số STFT và các đường biên phân trích tương tự được giữ như được sử dụng trong các ví dụ sớm hơn. Tuy nhiên lần này, các ảnh phổ cường độ thành phần không được suy ra từ các tín hiệu thành phần oracle, nhưng được trích từ hỗn hợp sử dụng 30 phép lặp NMFD. Tiếp theo, hai trường hợp thử nghiệm mới được giới thiệu. Trường hợp thử nghiệm 3 66c, 66c' sử dụng ước lượng pha ban đầu $(\varphi_c)^{(0)} := \varphi^{\text{Mix}}$ và ước lượng cường độ cố định $\mathcal{A}_c := V_c^T$, trong đó trường hợp thử nghiệm 4 66d sử dụng ước lượng pha ban đầu

$(\varphi_c)^{(0)} := 0$ và ước lượng cường độ cố định $\mathcal{A}_c := V_c^T$.

Fig.14a thể hiện phép khai căn ước số đồng nhất được tiêu chuẩn hóa so với số các phép lặp. Fig.14b thể hiện phép khai căn năng lượng tiếng vang trước so với số các phép lặp. Các đường cong thể hiện toàn bộ phần trích thử nghiệm trung bình, các giới hạn trực là giống với Fig.12. Hơn nữa, trên Fig.14a, sự giảm không đồng nhất được thu được sử dụng sự khôi phục TR 66c, 66d (các đường nét liền) không thể phân biệt được với phương pháp GL 66c', 66d' (các đường nét đứt). Các cải tiến là ít có nghĩa so với các số lượng mà có thể thu được khi sử dụng các ước lượng cường độ oracle (so với Fig.12a). Về trung bình, các sự khôi phục trong Trường hợp 3 66c, 66c' (được khởi tạo với φ^{Mix}) dường như nhanh chóng bị kẹt trong các điều kiện thuận lợi nhất. Đoán chừng, đây là do sự khai triển NMFD không hoàn hảo của các khung ảnh phổ liên quan đến sự bắt đầu, trong đó tất cả các nhạc cụ thể hiện sự phân bố cường độ phẳng nhiều hơn hoặc ít hơn và do đó thể hiện sự chồng lấp phổ tăng.

Trên Fig.14b, việc giảm tiếng vang trước với các ước lượng cường độ trên cơ sở NMFD $\mathcal{A}_c := V_c^T$ và pha 0 (Trường hợp 4, sơ đồ 66d, 66d') hoạt động hơi tệ hơn trong trường hợp 2 (so với Fig.12b). Điều này hỗ trợ các tìm kiếm sớm hơn, mà các ước lượng pha ban đầu yếu giúp ích nhiều nhất từ việc áp dụng nhiều phép lặp của phương pháp được đề xuất. Sự khôi phục GL sử dụng φ^{Mix} (Trường hợp 3, sơ đồ 66c, 66c') hơi tăng năng lượng tiếng vang trước qua các phép lặp. Ngược lại, việc áp dụng sự khôi phục TR tạo ra sự cải thiện đẹp.

Trên Fig.9, các sự khôi phục khác nhau của sự bắt đầu chùm chọi được lựa chọn từ vòng lặp trống ví dụ được thể hiện cụ thể. Bất chấp ước lượng cường độ được sử dụng (oracle trên Fig.9b hoặc dựa trên NMFD trên Fig.9c), sự khôi phục TR được đề xuất (phía dưới) thể hiện rõ các tiếng vang trước được giảm so với sự khôi phục GL thông thường (phía trên). Bằng các thử nghiệm nghe không chính thức (tốt hơn là sử dụng các ống nghe), một người có thể dễ dàng chỉ ra sự khác biệt mà có thể đạt được với các tổ hợp khác nhau của sự khởi tạo MSTFT. Thậm chí trong các trường, trong đó sự khôi phục cường độ không hoàn hảo dẫn đến các thành phần lạ xuyên âm không mong muốn trong các tín hiệu thành phần đơn, phương pháp TR theo các phương án tốt hơn là bảo toàn đặc điểm chuyển tiếp hơn là sự khôi phục GL thông thường. Hơn nữa, việc sử dụng

pha hỗn hợp cho sự khởi tạo MSTFT dường như là sự lựa chọn tốt vì người ta có thể thường nhận ra sự chênh lệch khó thấy trong sự khôi phục của pha suy giảm của các sự kiện trống so với các tín hiệu oracle. Tuy nhiên, sự chênh lệch âm sắc gây ra bởi sự khai triển cường độ không hoàn hảo được phiên âm nhiều hơn.

Các phương án thể hiện sự mở rộng hiệu quả đến quy trình LSEE-MSTFTM lặp đi lặp lại của Griffin và Lim cho sự khôi phục được phát triển của các thành phần tín hiệu chuyển tiếp trong việc tách nguồn âm nhạc. Thiết bị, bộ mã hóa, bộ giải mã hoặc phương pháp sử dụng thông tin phụ bổ sung về vị trí của các chuyển tiếp, mà có thể được cho trong cảnh tách nguồn được thông báo.

Theo các phương án khác, sự mở rộng hiệu quả đến quy trình LSEE-MSTFTM lặp đi lặp lại của Griffin và Lim cho sự khôi phục được phát triển của các thành phần tín hiệu chuyển tiếp trong việc tách nguồn âm nhạc được thể hiện. Phương pháp hoặc thiết bị sử dụng thông tin phụ bổ sung về vị trí của các chuyển tiếp, mà được giả định là được cho trong cảnh tách nguồn được thông báo. Hai thử nghiệm với tập hợp dữ liệu “IDMTSMT-Trống” có sẵn công khai đã thể hiện rằng phương pháp, bộ mã hóa, hoặc bộ giải mã theo phương án là có lợi để giảm các tiếng vang trước cả dưới các điều kiện phòng thí nghiệm cũng như đối với các tín hiệu thành phần thu được sử dụng kỹ thuật tách nguồn trong tình trạng kỹ thuật.

Theo các phương án, chất lượng cảm quan của các thành phần tín hiệu chuyển tiếp được trích trong ngữ cảnh của sự tách nguồn âm nhạc được cải thiện. Nhiều kỹ thuật trong tình trạng kỹ thuật được dựa trên việc áp dụng sự khai triển phù hợp thành phép biến đổi Fourier thời gian ngắn (Short-Time Fourier Transform - STFT) cường độ của tín hiệu hỗn hợp. Thông tin pha được sử dụng cho sự khôi phục các tín hiệu thành phần riêng lẻ thường được lấy từ hỗn hợp, dẫn đến STFT được cải biên (modified STFT - MSTFT) có giá trị phức. Các phương pháp khác nhau để khôi phục tín hiệu miền thời gian mà STFT của tín hiệu này làm gần đúng MSTFT đích. Do sự không nhất quán pha, các tín hiệu được khôi phục có thể chứa các thành phần lạ như các tiếng vang trước đúng trước các thành phần chuyển tiếp. Các phương án thể hiện sự mở rộng của quy trình khôi phục tín hiệu lặp đi lặp lại bởi Griffin và Lim để giải quyết vấn đề này. Thử nghiệm được làm thủ công cẩn thận sử dụng tập hợp thử nghiệm có sẵn thể hiện rằng phương

pháp và thiết bị giảm đáng kể các tiếng vang trước trong khi vẫn thể hiện các đặc tính hội tụ như phương pháp ban đầu.

Trong thử nghiệm khác, thể hiện rằng phương pháp hoặc thiết bị giảm đáng kể các tiếng vang trước trong khi vẫn thể hiện các đặc tính tương tự như phương pháp ban đầu bởi Griffin và Lin. Thử nghiệm thứ ba bao gồm sự khai triển âm thanh được thông báo điểm cũng thể hiện sự phát triển.

Các hình ảnh sau đây sẽ liên quan đến các phương án khác kết hợp với thiết bị 2.

Fig.15 thể hiện bộ mã hóa âm thanh 100 để mã hóa tín hiệu âm thanh 4. Bộ mã hóa âm thanh bao gồm bộ xử lý tín hiệu âm thanh và bộ xác định đường bao. Bộ xử lý tín hiệu âm thanh 102 được tạo cấu hình để mã hóa tín hiệu âm thanh miền thời gian sao cho tín hiệu âm thanh được mã hóa 108 bao gồm sự biểu diễn của chuỗi các khung miền tần số của tín hiệu âm thanh miền thời gian và sự biểu diễn của đường bao miền thời gian 106. Bộ xác định đường bao được tạo cấu hình để xác định đường bao từ tín hiệu âm thanh miền thời gian, trong đó bộ xác định đường bao còn được tạo cấu hình để so sánh đường bao được định trước để xác định sự biểu diễn của đường bao miền thời gian đích trên cơ sở so sánh. Đường bao có thể là đường bao miền thời gian của phần của tín hiệu âm thanh, ví dụ, và đường bao của khung hoặc phần khác của tín hiệu âm thanh. Hơn nữa, đường bao có thể được đề xuất cho bộ xử lý tín hiệu âm thanh mà có thể được tạo cấu hình để bao gồm đường bao trong tín hiệu âm thanh được mã hóa.

Nói cách khác, bộ mã hóa âm thanh (tiêu chuẩn) có thể được mở rộng đến bộ mã hóa âm thanh 100 bằng cách xác định đường bao, ví dụ, đường bao miền thời gian của phần, ví dụ, khung của tín hiệu âm thanh. Đường bao được suy ra có thể được so với tập hợp hoặc số lượng các đường bao miền thời gian được định trước trong bảng mã hoặc bảng tìm kiếm. Vị trí của đường bao được định trước phù hợp nhất có thể được mã hóa sử dụng, ví dụ, số lượng các bit. Do đó, có thể được sử dụng bốn bit để định vị, ví dụ, 16 đường bao miền thời gian được định trước khác, năm bit để định vị, ví dụ, 32 đường bao miền thời gian được định trước, hoặc bất kỳ số lượng các bit nào, phụ thuộc vào số lượng các đường bao miền thời gian được định trước khác.

Fig.16 thể hiện bộ giải mã âm thanh 110 bao gồm thiết bị 2 và giao diện đầu vào 112. Giao diện đầu vào 112 có thể nhận tín hiệu âm thanh được mã hóa. Tín hiệu âm

thanh được mã hóa có thể bao gồm sự biểu diễn của chuỗi các khung miền tần số và sự biểu diễn đường bao miền thời gian đích.

Nói cách khác, bộ giải mã 110 có thể nhận tín hiệu âm thanh được mã hóa cho ví dụ từ bộ mã hóa 100. Giao diện đầu vào 112 hoặc thiết bị 2, hoặc các phương tiện khác có thể trích đường bao miền thời gian đích 14 hoặc sự biểu diễn của nó, ví dụ, chuỗi các bit biểu thị vị trí của đường bao miền thời gian đích trong bảng tìm kiếm hoặc bảng mã. Hơn nữa, thiết bị 2 có thể giải mã tín hiệu âm thanh được mã hóa 108, ví dụ, bằng cách điều chỉnh các pha sai lệch của tín hiệu âm thanh được mã hóa vẫn có các giá trị cường độ, hoặc thiết bị có thể các giá trị pha đúng của tín hiệu âm thanh được giải mã, ví dụ, từ đơn vị giải mã mà mà giải mã đủ hoặc thậm chí hoàn hảo cường độ phổ của tín hiệu âm thanh, và thiết bị còn điều chỉnh pha của tín hiệu âm thanh được giải mã, mà có thể được tính toán bởi bộ phận giải mã.

Fig.17 thể hiện tín hiệu âm thanh 114 bao gồm sự biểu diễn của chuỗi các khung miền tần số 12 và sự biểu diễn đường bao miền thời gian đích 14. Sự biểu diễn chuỗi các khung miền tần số của tín hiệu âm thanh miền thời gian 12 có thể được tín hiệu âm thanh được mã hóa theo sơ đồ mã hóa âm thanh tiêu chuẩn. Hơn nữa, sự biểu diễn đường bao miền thời gian đích 14 có thể là sự biểu diễn bit của đường bao miền thời gian. Sự biểu diễn có thể được suy ra, ví dụ, sử dụng việc lấy mẫu và lượng tử hóa đường bao miền thời gian đích hoặc bởi phương pháp số hóa khác. Hơn nữa, sự biểu diễn đường bao miền thời gian đích 14 có thể là chỉ số của, ví dụ, bảng mã hoặc bảng tìm kiếm được biểu thị hoặc được mã hóa với số lượng các bit.

Fig.18 thể hiện sơ đồ khối dạng giản lược của bộ xử lý tách nguồn âm thanh 116 theo phương án. Bộ xử lý tách nguồn âm thanh bao gồm thiết bị 2 và bộ tạo màn chắn phổ 118. Bộ tạo màn chắn phổ có thể tạo màn chắn phổ của tín hiệu âm thanh ban đầu 4 để suy ra tín hiệu âm thanh được cải biên 120. So sánh với tín hiệu âm thanh ban đầu 4, tín hiệu âm thanh được cải biên 120 có thể bao gồm số lượng các băng tần hoặc các ngăn tần số thời gian. Hơn nữa, tín hiệu âm thanh được cải biên có thể bao gồm chỉ một nguồn hoặc một nhạc cụ hoặc một loa phát thanh (con người) của tín hiệu âm thanh 4, trong đó các đóng góp tần số của các nguồn, các loa phát thanh, hoặc các nhạc cụ khác bị ẩn hoặc bị chắn. Tuy nhiên, vì các giá trị cường độ của tín hiệu âm thanh được cải

biên 120 có thể khớp với các giá trị cường độ của tín hiệu âm thanh được xử lý 6, các giá trị pha của tín hiệu âm thanh được cải biên có thể sai lệch. Do đó, thiết bị 2 có thể hiệu chỉnh các giá trị pha của tín hiệu âm thanh được cải biên so với đường bao miền thời gian 14.

Fig.19 thể hiện sơ đồ khối dạng giản lược của bộ xử lý nâng cao băng thông 122 theo phương án. Bộ xử lý nâng cao băng thông 122 được tạo cấu hình để xử lý tín hiệu âm thanh được mã hóa 124. Hơn nữa, bộ xử lý nâng cao băng thông 122 bao gồm bộ xử lý nâng cao 126 và thiết bị 2. Bộ xử lý nâng cao 126 được tạo cấu hình để tạo ra tín hiệu nâng cao 127 từ băng tín hiệu âm thanh nằm trong tín hiệu âm thanh được mã hóa và trong đó bộ xử lý nâng cao 26 được tạo cấu hình để trích đường bao miền thời gian đích 14 từ sự biểu diễn được mã hóa nằm trong tín hiệu được mã hóa 122 hoặc từ băng tín hiệu âm thanh nằm trong tín hiệu được mã hóa. Hơn nữa, thiết bị 2 có thể xử lý tín hiệu nâng cao 126 sử dụng đường bao miền thời gian đích.

Nói cách khác, bộ xử lý nâng cao 126 có thể mã hóa lỗi băng tín hiệu âm thanh hoặc nhận băng tín hiệu âm thanh được mã hóa lỗi của tín hiệu âm thanh được mã hóa. Hơn nữa, bộ xử lý nâng cao 126 có thể tính toán các băng khác của tín hiệu âm thanh sử dụng, ví dụ, các tham số của tín hiệu âm thanh được mã hóa và phân băng cơ sở được mã hóa lỗi của tín hiệu âm thanh. Hơn nữa, đường bao miền thời gian đích 14 có thể có trong tín hiệu âm thanh được mã hóa 124, hoặc bộ xử lý nâng cao có thể được tạo cấu hình để tính toán đường bao miền thời gian đích từ phân băng cơ sở của tín hiệu âm thanh.

Fig.20 minh họa việc biểu diễn dưới dạng giản đồ của phổ. Phổ được chia nhỏ ra trong các băng hệ số tỉ lệ SCB tại đó có bảy băng hệ số tỉ lệ SCB1 tới SCB7 trong ví dụ minh họa của Fig.20. Các băng hệ số tỉ lệ có thể là băng hệ số tỉ lệ AAC mà được định rõ trong tiêu chuẩn AAC và có băng thông tăng dần tới tần số trên như đã minh họa trên Fig.20 dưới dạng giản đồ. Ưu tiên thực hiện điền đầy khoảng trống thông minh không từ phần đầu phổ, tức là, tại tần số thấp, nhưng để bắt đầu thuật toán IGF tại tần số bắt đầu IGF được minh họa tại 309. Do đó, băng tần số lõi mở rộng từ tần số thấp nhất đến tần số bắt đầu IGF. Phía trên tần số bắt đầu IGF, việc phân tích phổ được áp dụng để phân tách các thành phần phổ có độ phân giải cao 304, 305, 306, 307 (tập hợp thứ nhất

gồm các phần phổ thứ nhất) từ các thành phần có độ phân giải thấp được biểu diễn bởi tập hợp thứ hai gồm phần phổ thứ hai. Fig.20 minh họa phổ mà là được nhập một cách minh họa vào bộ xử lý nâng cao 126, tức là, bộ mã hóa lỗi có thể vận hành trong phạm vi đầy đủ, nhưng mã hóa một số lượng đáng kể giá trị phổ 0, tức là, các giá trị phổ 0 này được lượng tử hóa đến 0 hoặc là được thiết lập đến 0 trước khi lượng tử hóa hoặc sau khi lượng tử hóa. Dù thế nào đi nữa, bộ mã hóa lỗi hoạt động trong khoảng đầy đủ, tức là, nếu phổ sẽ được minh họa, tức là, bộ giải mã lỗi không nhất thiết phải nhận biết bất kỳ việc điền đầy khoảng trống thông minh hoặc mã hóa tập hợp thứ hai gồm các phần phổ thứ hai với độ phân giải phổ thấp hơn.

Tốt hơn là, độ phân giải cao được định rõ bằng việc mã hóa theo vạch của các vạch phổ như các vạch MDCT, trong khi độ phân giải thứ hai hoặc độ phân giải thấp được định rõ bằng, ví dụ, việc tính toán chỉ một giá trị phổ đơn trên băng thừa số tỉ lệ, tại đó băng thừa số tỉ lệ bao trùm một số vạch tần số. Do đó, độ phân giải thấp thứ hai là, đối với độ phân giải phổ của nó, thấp hơn nhiều so với độ phân giải thứ nhất hoặc cao được định rõ bởi việc mã hóa dòng thông minh được áp dụng bởi bộ mã hóa lỗi như AAC hoặc USAC.

Do thực tế rằng bộ mã hóa là bộ mã hóa lỗi vào thực tế rằng có thể, nhưng không nhất thiết phải là, các thành phần của tập hợp thứ nhất gồm các phần phổ trong mỗi băng, bộ mã hóa lỗi tính toán hệ số tỉ lệ cho mỗi băng không chỉ trong phạm vi lỗi dưới tần số IGF 309, nhưng cũng ở trên tần số bắt đầu IGF cho đến khi tần số tối đa mà nhỏ hơn hoặc bằng một nửa tần số lấy mẫu, tức là, $f_s/2$. Do đó, các phần âm được mã hóa 302, 304, 305, 306, 307 của Fig.20 và, trong phương án này cùng với các thừa số tỉ lệ SCB1 đến SCB7 tương ứng với dữ liệu phổ độ phân giải cao. Dữ liệu phổ độ phân giải thấp được tính toán bắt đầu từ tần số bắt đầu IGF và tương ứng với các giá trị thông tin năng lượng E_1, E_2, E_3, E_4 , mà được truyền dẫn cùng với các thừa số tỉ lệ SF4 đến SF7.

Cụ thể, khi bộ mã hóa lỗi dưới điều kiện tốc độ truyền dữ liệu thấp, thuật toán điền đầy nhiều âm bổ sung trong băng lỗi, tức là, thấp hơn trong tần số so với tần số bắt đầu IGF, tức là, trong các băng hệ số tỉ lệ SB1 tới SB3 có thể được áp dụng bổ sung. Trong việc điền đầy nhiều âm, có các dòng phổ liên kế khác nhau mà đã được lượng tử hóa tới 0. Trên phía bộ giải mã, các giá trị phổ được lượng tử hóa tới 0 được tổng hợp lại và các

giá trị phổ được tổng hợp lại được điều chỉnh trong cường độ của chúng sử dụng năng lượng điền đầy nhiễu âm. Năng lượng điền đầy nhiễu âm, có thể có được các giới hạn tuyệt đối hoặc các giới hạn liên quan đặc biệt đối với hệ số tỉ lệ như trong USAC tương ứng với năng lượng của tập hợp các giá trị phổ được lượng tử hóa tới 0. Các vạch phổ điền đầy nhiễu âm này có thể cũng được xem xét là tập hợp thứ ba gồm các phần phổ thứ ba mà được tái tạo bằng việc tổng hợp điền đầy nhiễu âm truyền thẳng mà không cần bất kỳ thuật toán IGF phụ thuộc vào việc tái tạo tần số sử dụng ô tần số từ các tần số khác cho ô tần số khôi phục sử dụng các giá trị phổ từ phạm vi nguồn và thông tin năng lượng E_1, E_2, E_3, E_4 .

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

Phần được mã hóa lõi hoặc băng tần được mã hóa lõi của tín hiệu âm thanh được mã hóa 124 có thể bao gồm sự biểu diễn độ phân giải cao của tín hiệu âm thanh lên đến tần số ngắt hoặc tần số bắt đầu IGF 309. Phía trên tần số bắt đầu IGF 309 này, tín hiệu âm thanh có thể bao gồm các băng thừa số tỉ lệ được mã hóa với độ phân giải thấp, ví dụ, sử dụng sự mã hóa theo tham số. Tuy nhiên, sử dụng phần băng cơ sở được mã hóa lõi và ví dụ, các tham số, tín hiệu âm thanh được mã hóa 124 có thể được giải mã. Điều này có thể được thực hiện một hoặc nhiều lần.

Điều này có thể đề xuất sự khôi phục tốt của các giá trị cường độ thậm chí phía trên tần số ngắt thứ nhất 130. Tuy nhiên, ít nhất quanh các tần số ngắt giữa các băng hệ số tỉ lệ liên tiếp, tần số trên cùng hoặc cao nhất của phần băng cơ sở được mã hóa lõi 128 có thể liên kết với tần số thấp nhất của phần băng cơ sở được mã hóa lõi do việc đệm phần băng cơ sở được lõi do các tần số cao hơn phía trên tần số bắt đầu IGF 309, các giá trị pha có thể sai lệch. Do đó, tín hiệu âm thanh được khôi phục băng cơ sở có thể

được nhập vào thiết bị 2 để khôi phục các pha của tín hiệu được mở rộng băng.

Hơn nữa, việc nâng cao băng thông hoạt động vì phần băng cơ sở được mã hóa lỗi bao gồm nhiều thông tin liên quan đến tín hiệu âm thanh ban đầu. Điều này dẫn đến kết luận rằng đường bao của phần băng cơ sở được mã hóa lỗi là ít nhất tương tự với đường bao của tín hiệu âm thanh ban đầu, thậm chí đường bao của tín hiệu âm thanh ban đầu có thể được nhấn mạnh do các thành phần tần số cao khác của tín hiệu âm thanh, mà không được thể hiện hoặc không có trong phần băng cơ sở được mã hóa lỗi.

Fig.21 thể hiện sự biểu diễn giản lược của sự khôi phục miền thời gian (trung gian) sau số lượng thứ nhất gồm các bước lặp phía trên, và sau số lượng thứ hai gồm các bước lặp tại phía dưới của Fig.21. Các gợn sóng tương đối cao 132 là kết quả của sự không đồng nhất của các khung liên kề của chuỗi các khung miền tần số. Thông thường, bắt đầu từ tín hiệu miền thời gian, STFT nghịch đảo của STFT tín hiệu miền thời gian một lần nữa dẫn đến tín hiệu miền thời gian. Ở đây, các khung miền tần số liên kề là đồng nhất sau STFT được áp dụng, sao cho quy trình chồng lắp và cộng của thuật toán STFT nghịch đảo cộng tổng hoặc biểu lộ tín hiệu ban đầu. Tuy nhiên, bắt đầu từ miền tần số với các giá trị pha sai lệch, các khung miền tần số không đồng nhất (tức là, không đồng nhất), trong đó STFT của ISTFT của tín hiệu miền tần số không dẫn đến tín hiệu âm thanh phù hợp hoặc đồng nhất như được biểu thị ở phía trên của Fig.21. Tuy nhiên, đã được chứng minh bằng toán học rằng thuật toán, nếu được áp dụng cho cường độ ban đầu, giảm các gợn sóng 132 trong mỗi bước lặp dẫn đến tín hiệu âm thanh được khôi phục (gần như hoàn hảo) được biểu thị tại phía dưới Fig.21. Ở đây, các gợn sóng 132 giảm. Nói cách khác, cường độ của tín hiệu miền thời gian trung gian chuyển đổi đến giá trị cường độ ban đầu của chuỗi các khung miền tần số sau mỗi bước lặp. Phải được lưu ý rằng kích thước bước nhảy là 0,5 giữa các cửa sổ tổng hợp liên tiếp 136 được lựa chọn để thuận tiện và có thể được thiết lập về bất kỳ giá trị phù hợp nào, như, ví dụ, 0,75.

Fig.22 thể hiện sơ đồ khối dạng giản lược của phương pháp 2200 xử lý tín hiệu âm thanh để thu được tín hiệu âm thanh được xử lý. Phương pháp 2200 bao gồm bước 2205 tính toán các giá trị pha cho các giá trị phổ của chuỗi các khung miền tần số biểu diễn các khung chồng lắp của tín hiệu âm thanh, trong đó các giá trị pha được tính toán trên

cơ sở thông tin trên đường bao miền thời gian đích liên quan đến tín hiệu âm thanh được xử lý, để tín hiệu âm thanh được xử lý có ít nhất trong phép làm gần úng đường bao miền thời gian đích và đường bao phổ được xác định bởi chuỗi các khung miền tần số.

Fig.23 thể hiện sơ đồ khối dạng giản lược của phương pháp 2300 giải mã âm thanh. Phương pháp 2300 bao gồm trong bước 2305 phương pháp 2200 và trong bước 2310, nhận tín hiệu ược mã hóa, tín hiệu bao gồm sự biểu diễn chuỗi các khung miền tần số, và sự biểu diễn của đường bao miền thời gian đích.

Fig.24 thể hiện sơ đồ khối dạng giản lược của phương pháp 2400 tách nguồn âm thanh. Phương pháp 2400 bao gồm bước 2405 thực hiện phương pháp 2200, và bước 2410 của màn chắn phổ của tín hiệu âm thanh ban đầu để thu được tín hiệu âm thanh được cải biên được nhập vào thiết bị để xử lý, trong đó tín hiệu âm thanh được xử lý là tín hiệu nguồn được tách liên quan đến đường bao miền thời gian đích.

Fig.25 thể hiện sơ đồ khối dạng giản lược của phương pháp nâng cao băng thông của tín hiệu âm thanh được mã hóa. Phương pháp 2500 bao gồm bước 2505 tạo ra tín hiệu nâng cao từ băng tín hiệu âm thanh nằm trong tín hiệu được mã hóa, bước 2510 thực hiện phương pháp 2200, và bước 2515, trong đó vận hành chung bao gồm tách đường bao miền thời gian đích từ sự biểu diễn được mã hóa nằm trong tín hiệu được mã hóa hoặc từ băng tín hiệu âm thanh nằm trong tín hiệu được mã hóa.

Fig.26 thể hiện sơ đồ khối dạng giản lược của phương pháp 2600 mã hóa âm thanh. Phương pháp 2600 bao gồm bước 2605 mã hóa tín hiệu âm thanh miền thời gian sao cho tín hiệu âm thanh được mã hóa bao gồm sự biểu diễn chuỗi các khung miền tần số của tín hiệu âm thanh miền thời gian và sự biểu diễn đường bao miền thời gian đích, và bước 2610 xác định đường bao từ tín hiệu âm thanh miền thời gian, trong đó bộ xác định đường bao còn được tạo cấu hình để so sánh đường bao với tập hợp các đường bao được định trước để xác định sự biểu diễn đường bao miền thời gian đích trên cơ sở so sánh.

Các phương án khác của sáng chế đề cập đến ví dụ sau. Điều này có thể là phương pháp, thiết bị, hoặc chương trình máy tính để

- 1) khôi phục lặp đi lặp lại tín hiệu miền thời gian từ sự biểu diễn miền thời gian - tần số,

2) tạo ra ước lượng ban đầu cho thông tin cường độ và pha và sự biểu diễn miền thời gian - tần số,

3) áp dụng sự mô phỏng tín hiệu trung gian cho các đặc tính tín hiệu nhất định trong suốt phép lặp,

4) biến đổi sự biểu diễn miền thời gian - tần số về miền thời gian,

5) điều biến tín hiệu miền thời gian với đường bao biên độ tùy ý,

6) biến đổi tín hiệu miền thời gian được điều biến về miền thời gian - tần số,

7) sử dụng thông tin pha thu được để cập nhật sự biểu diễn miền thời gian - tần số,

8) mô phỏng chuỗi phép biến đổi nghịch đảo và phép biến đổi tịnh tiến bởi quy trình miền thời gian - tần số mà thêm các đóng góp được nhân chấp và được dịch chuyển cụ thể từ các khung liền kề với khung chính giữa,

9) làm gần đúng quy trình phía trên bằng các hạch nhân chấp được cắt và khai thác các đặc tính đối xứng,

10) mô phỏng phép điều biến miền thời gian bằng phép nhân chấp của khung mong muốn với sự biểu diễn thời gian - tần số của đường bao đích,

11) áp dụng các phép mô phỏng miền thời gian - tần số bằng cách phụ thuộc miền thời gian, ví dụ áp dụng các thuật toán chỉ để lựa chọn các ngăn thời gian - tần số, hoặc

12) sử dụng các quy trình được mô tả trên đây cho việc mã hóa âm thanh cảm quan, tách nguồn âm thanh, và/hoặc nâng cao băng thông.

Nhiều loại đánh giá trong cảnh khai triển âm thanh được áp dụng cho thiết bị hoặc phương pháp theo phương án, trong đó mục đích là trích các âm trống tách biệt từ các bản ghi trống đa âm. Tập hợp thử nghiệm có sẵn công khai mà được làm giàu với tất cả thông tin phụ cần thiết, như các tín hiệu thành phần “oracle” chính xác và các vị trí chuyển tiếp chính xác của chúng. Trong một thử nghiệm, trong các điều kiện phòng thí nghiệm, việc sử dụng tất cả thông tin phụ được thực hiện để tập trung vào đánh giá lợi ích của phương pháp hoặc thiết bị để bảo toàn chuyển tiếp trong sự khôi phục tín hiệu. Trong các điều kiện được lý tưởng hóa này, phương pháp được đề xuất có thể giảm đáng

kể các tiếng vang trước trong khi vẫn thể hiện các đặc tính hội tụ tương tự như phương pháp hoặc thiết bị ban đầu. Trong thử nghiệm khác, kỹ thuật khai triển theo tình trạng kỹ thuật [3, 4] được sử dụng với các hạn chế được thông báo điểm để ước lượng STFTM của tín hiệu thành phần từ hỗn hợp. Trong các điều kiện (thực tế hơn) này, phương pháp được đề xuất vẫn tạo ra các cải thiện đáng kể.

Sẽ được hiểu rằng trong phần mô tả này, các tín hiệu trên các vạch đôi khi được đặt tên bởi các số tham chiếu cho các vạch hoặc đôi khi được biểu thị bởi bản thân các số tham chiếu của chúng, mà đã được quy cho các vạch. Do đó, ký hiệu sao cho vạch có tín hiệu nhất định biểu thị tín hiệu của vạch. Vạch có thể là vạch vật lý trong sự thực thi được gắn cứng. Tuy nhiên, trong sự thực thi được điện toán hóa, vạch vật lý không tồn tại, nhưng tín hiệu được biểu diễn bởi vạch được truyền dẫn từ một môđun tính đến môđun tính toán khác.

Mặc dù, sáng chế đã được mô tả trong ngữ cảnh của các sơ đồ khối trong đó các khối biểu diễn các thành phần phần cứng thực sự hoặc logic, sáng chế cũng có thể được thực hiện bởi phương pháp được thực hiện bằng máy tính. Trong trường hợp sau, các khối biểu diễn các bước phương pháp tương ứng trong đó các bước này có nghĩa là các chức năng được thực hiện bởi các khối phần cứng logic hoặc vật lý.

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

Các tín hiệu được truyền dẫn hoặc được mã hóa đổi mới có thể được lưu trữ trên vật ghi lưu trữ số hoặc có thể được truyền trong môi trường truyền như môi trường truyền không dây hoặc môi trường truyền có dây như Internet.

Phụ thuộc vào các yêu cầu thực thi nhất định, các phương án của sáng chế có thể

được thực thi trong phần cứng hoặc trong phần mềm. Phương án có thể được thực hiện sử dụng vật lưu trữ số, ví dụ, đĩa mềm, DVD, Blu-ray, CD, ROM, PROM, EPROM, EEPROM hoặc bộ nhớ FLASH, có các tín hiệu điều khiển có thể đọc theo điện tử được lưu trữ trên đó, mà kết hợp (hoặc có thể kết hợp) với hệ thống máy tính khả trình sao cho phương pháp tương ứng được thực hiện. Do đó, vật ghi lưu trữ kỹ thuật số có thể đọc được bằng máy tính.

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

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

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

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

Do đó, phương án khác của các phương pháp theo sáng chế là, vật mang dữ liệu (hoặc vật khi lưu trữ không tạm thời 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 số các phương pháp được mô tả ở đây. Vật mang dữ liệu, vật ghi lưu trữ số hoặc vật ghi được ghi lại dạng hữu hình điển hình và/hoặc không chuyển tiếp.

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

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

phương pháp được mô tả ở đây.

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

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

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

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

YÊU CẦU BẢO HỘ

1. Thiết bị xử lý tín hiệu âm thanh để thu được tín hiệu âm thanh được xử lý, thiết bị bao gồm:

bộ tính toán pha để tính toán các giá trị pha cho các giá trị phổ của chuỗi các khung miền tần số biểu diễn các khung chồng lấp của tín hiệu âm thanh,

trong đó bộ tính toán pha được tạo cấu hình để tính toán các giá trị pha dựa trên thông tin về đường bao miền thời gian đích liên quan đến tín hiệu âm thanh được xử lý, để tín hiệu âm thanh được xử lý bao gồm ít nhất trong phép tính gần đúng đường bao miền thời gian đích và đường bao phổ được xác định bởi chuỗi các khung miền tần số.

2. Thiết bị theo điểm 1,

trong đó bộ tính toán pha bao gồm:

bộ xử lý phép lặp để thực hiện thuật toán lặp để tính toán, bắt đầu từ các giá trị pha ban đầu, các giá trị pha cho các giá trị phổ sử dụng đích tối ưu hóa yêu cầu sự đồng nhất của các khối chồng lấp trong phạm vi chồng lấp,

trong đó bộ xử lý lặp được tạo cấu hình để sử dụng, trong bước lặp khác, ước lượng pha được cập nhật phụ thuộc vào đường bao miền thời gian đích.

3. Thiết bị theo điểm 1, trong đó bộ tính toán pha được tạo cấu hình để áp dụng phép điều biến biên độ cho sự khôi phục miền thời gian trung gian của tín hiệu âm thanh trên cơ sở đường bao miền thời gian đích.

4. Thiết bị theo điểm 1, trong đó bộ tính toán pha được tạo cấu hình để áp dụng phép nhân chập của sự biểu diễn phổ và ít nhất một đường bao miền thời gian đích và ít nhất một sự khôi phục miền tần số trung gian hoặc các phần hoặc các băng được lựa chọn hoặc chỉ phần thông cao hoặc chỉ một vài phần dải giữa của ít nhất một đường bao miền thời gian đích hoặc ít nhất một sự khôi phục miền tần số trung gian của tín hiệu âm thanh.

5. Thiết bị theo điểm 3, trong đó bộ tính toán pha bao gồm:

bộ chuyển đổi tần số thành thời gian tính toán sự khôi phục miền thời gian trung gian của tín hiệu âm thanh từ chuỗi các khung miền tần số và các ước lượng giá trị pha

ban đầu hoặc các ước lượng giá trị pha của bước lặp trước đó,

bộ điều biến biên độ điều biến sự khôi phục miền thời gian sử dụng đường bao miền thời gian đích để thu được tín hiệu âm thanh được điều biến biên độ, và

bộ chuyển đổi thời gian thành tần số để chuyển đổi tín hiệu được điều biến biên độ thành chuỗi các khung miền tần số khác bao gồm các giá trị pha, và

trong đó bộ tính toán pha được tạo cấu hình để sử dụng, đối với bước lặp tiếp theo, các giá trị pha và các giá trị phổ của chuỗi các khung miền tần số.

6. Thiết bị theo điểm 5,

trong đó bộ tính toán pha được tạo cấu hình để phát ra sự khôi phục miền thời gian trung gian như tín hiệu âm thanh được xử lý, khi điều kiện xác định phép lặp thỏa mãn.

7. Thiết bị theo điểm 4,

trong đó bộ tính toán pha bao gồm:

bộ xử lý nhân chập để áp dụng hạch nhân chập và để áp dụng hạch dịch chuyển và để bổ sung phần chồng lấp của khung liền kề của khung chính giữa cho khung chính giữa để thu được sự khôi phục miền tần số của tín hiệu âm thanh.

8. Thiết bị theo điểm 4,

trong đó bộ tính toán pha được tạo cấu hình để sử dụng các giá trị pha thu được bởi phép nhân chập như các ước lượng giá trị pha được cập nhật đối với bước lặp tiếp theo.

9. Thiết bị theo điểm 4,

thiết bị còn bao gồm bộ chuyển đổi đường bao đích để chuyển đổi đường bao miền thời gian đích thành miền phổ.

10. Thiết bị theo điểm 4,

thiết bị còn bao gồm bộ chuyển đổi tần số thành thời gian để tính toán sự khôi phục miền thời gian từ sự khôi phục miền tần số trung gian sử dụng các ước lượng giá trị pha thu được từ bước lặp gần nhất và chuỗi các khung miền tần số.

11. Thiết bị theo điểm 4,

trong đó bộ tính toán pha bao gồm bộ xử lý nhân chập để xử lý chuỗi các khung miền tần số, trong đó bộ xử lý nhân chập được tạo cấu hình để áp dụng quy trình chồng lấp và cộng miền tần số cho chuỗi các khung miền tần số trong miền tần số để xác định sự khôi phục miền tần số trung gian.

12. Thiết bị theo điểm 11,

trong đó bộ xử lý phép nhân chập được tạo cấu hình để xác định, dựa trên khung miền tần số hiện thời, phần của khung miền tần số liền kề mà đóng góp cho khung miền tần số hiện thời sau phép chồng lấp và cộng miền thời gian được thực hiện trong miền tần số,

bộ xử lý phép nhân chập còn được tạo cấu hình để xác định vị trí chồng lấp của phần của khung miền tần số liền kề nằm trong khung tần số hiện thời và để thực hiện sự bổ sung các vị trí của các khung miền tần số liền kề với khung miền tần số hiện thời tại vị trí chồng lấp.

13. Thiết bị theo điểm 11, trong đó bộ xử lý nhân chập được tạo cấu hình để biến đổi tần số thành thời gian, phép tổng hợp miền thời gian và cửa sổ phân tích miền thời gian để xác định phần của khung miền tần số liền kề mà đóng góp cho khung miền tần số sau phép chồng lấp và cộng miền thời gian được thực hiện trong miền tần số, trong đó bộ xử lý nhân chập còn được tạo cấu hình để dịch chuyển vị trí của khung miền tần số liền kề cho vị trí chồng lấp nằm trong khung miền tần số hiện thời và để áp dụng phần của khung miền tần số liền kề cho khung hiện thời tại vị trí chồng lấp.

14. Thiết bị theo điểm 1,

trong đó bộ tính toán pha được tạo cấu hình để thực hiện thuật toán lặp phù hợp với quy trình khôi phục tín hiệu lặp đi lặp lại bởi Griffin và Lim.

15. Bộ giải mã âm thanh bao gồm:

thiết bị theo điểm 1, và

giao diện đầu vào để nhận tín hiệu được mã hóa, tín hiệu được mã hóa bao gồm sự biểu diễn của chuỗi các khung miền tần số và sự biểu diễn đường bao miền thời gian đích.

16. Bộ xử lý tách nguồn âm thanh bao gồm:

thiết bị để xử lý theo điểm 1, màn chắn phổ để chắn phổ của tín hiệu âm thanh ban đầu để thu được tín hiệu âm thanh được cải biên được nhập vào thiết bị để xử lý,

trong đó tín hiệu âm thanh được xử lý là tín hiệu nguồn được tách liên quan đến đường bao miền thời gian đích.

17. Bộ xử lý nâng cao bằng thông để xử lý tín hiệu âm thanh được mã hóa, bộ xử lý nâng cao bằng thông bao gồm:

bộ xử lý nâng cao để tạo ra tín hiệu nâng cao từ băng tín hiệu âm thanh nằm trong tín hiệu được mã hóa, và

thiết bị để xử lý theo điểm 1,

trong đó bộ xử lý nâng cao được tạo cấu hình để trích đường bao miền thời gian đích từ sự biểu diễn được mã hóa nằm trong tín hiệu được mã hóa hoặc từ băng tín hiệu âm thanh nằm trong tín hiệu được mã hóa.

18. Phương pháp xử lý tín hiệu âm thanh để thu được tín hiệu âm thanh được xử lý, phương pháp bao gồm:

tính toán các giá trị pha cho các giá trị phổ của chuỗi các khung miền tần số biểu diễn các khung chồng lấp của tín hiệu âm thanh,

trong đó các giá trị pha được tính toán dựa trên thông tin về đường bao miền thời gian đích liên quan đến tín hiệu âm thanh được xử lý, để tín hiệu âm thanh được xử lý bao gồm ít nhất trong phép tính gần đúng đường bao miền thời gian đích và đường bao phổ được xác định bởi chuỗi các khung miền tần số.

19. Phương pháp giải mã âm thanh, phương pháp bao gồm:

phương pháp xử lý tín hiệu âm thanh để thu được tín hiệu âm thanh được xử lý, phương pháp bao gồm:

tính toán các giá trị pha cho các giá trị phổ của chuỗi các khung miền tần số biểu diễn các khung chồng lấp của tín hiệu âm thanh,

trong đó các giá trị pha được tính toán dựa trên thông tin về đường bao miền thời gian đích liên quan đến tín hiệu âm thanh được xử lý, để tín hiệu âm thanh được xử lý

bao gồm ít nhất trong phép tính gần đúng đường bao miền thời gian đích và đường bao phổ được xác định bởi chuỗi các khung miền tần số;

nhận tín hiệu được mã hóa, tín hiệu được mã hóa bao gồm sự biểu diễn của chuỗi các khung miền tần số, và sự biểu diễn của đường bao miền thời gian đích.

20. Phương pháp tách nguồn âm thanh, phương pháp bao gồm:

phương pháp xử lý tín hiệu âm thanh để thu được tín hiệu âm thanh được xử lý, phương pháp bao gồm:

tính toán các giá trị pha cho các giá trị phổ của chuỗi các khung miền tần số biểu diễn các khung chồng lấp của tín hiệu âm thanh,

trong đó các giá trị pha được tính toán dựa trên thông tin về đường bao miền thời gian đích liên quan đến tín hiệu âm thanh được xử lý, để tín hiệu âm thanh được xử lý bao gồm ít nhất trong phép tính gần đúng đường bao miền thời gian đích và đường bao phổ được xác định bởi chuỗi các khung miền tần số, và

tạo màn chắn phổ của tín hiệu âm thanh ban đầu để thu được tín hiệu âm thanh được cải biên được nhập vào thiết bị để xử lý;

trong đó tín hiệu âm thanh được xử lý là tín hiệu nguồn được tách liên quan đến đường bao miền thời gian đích.

21. Phương pháp nâng cao băng thông của tín hiệu âm thanh được mã hóa, phương pháp bao gồm:

tạo ra tín hiệu nâng cao từ băng tín hiệu âm thanh nằm trong tín hiệu được mã hóa;

phương pháp xử lý tín hiệu âm thanh để thu được tín hiệu âm thanh được xử lý, phương pháp bao gồm:

tính toán các giá trị pha cho các giá trị phổ của chuỗi các khung miền tần số biểu diễn các khung chồng lấp của tín hiệu âm thanh,

trong đó các giá trị pha được tính toán dựa trên thông tin về đường bao miền thời gian đích liên quan đến tín hiệu âm thanh được xử lý, để tín hiệu âm thanh được xử lý bao gồm ít nhất trong phép tính gần đúng đường bao miền thời gian đích và đường bao phổ được xác định bởi chuỗi các khung miền tần số;

trong đó tạo ra bao gồm trích đường bao miền thời gian đích từ sự biểu diễn được mã hóa nằm trong tín hiệu được mã hóa hoặc từ băng tín hiệu âm thanh nằm trong tín hiệu được mã hóa.

22. Vật ghi lưu trữ số không tạm thời bao gồm chương trình máy tính được chứa trên đó để thực hiện phương pháp xử lý tín hiệu âm thanh để thu được tín hiệu âm thanh được xử lý, phương pháp bao gồm:

tính toán các giá trị pha cho các giá trị phổ của chuỗi các khung miền tần số biểu diễn các khung chồng lấp của tín hiệu âm thanh,

trong đó các giá trị pha được tính toán dựa trên thông tin về đường bao miền thời gian đích liên quan đến tín hiệu âm thanh được xử lý, để tín hiệu âm thanh được xử lý bao gồm ít nhất trong phép tính gần đúng đường bao miền thời gian đích và đường bao phổ được xác định bởi chuỗi các khung miền tần số,

khi chương trình máy tính nêu trên được chạy bởi máy tính.

23. Vật ghi lưu trữ số không tạm thời bao gồm chương trình máy tính được chứa trên đó để thực hiện phương pháp giải mã âm thanh, phương pháp bao gồm:

phương pháp xử lý tín hiệu âm thanh để thu được tín hiệu âm thanh được xử lý, phương pháp bao gồm:

tính toán các giá trị pha cho các giá trị phổ của chuỗi các khung miền tần số biểu diễn các khung chồng lấp của tín hiệu âm thanh,

trong đó các giá trị pha được tính toán dựa trên thông tin về đường bao miền thời gian đích liên quan đến tín hiệu âm thanh được xử lý, để tín hiệu âm thanh được xử lý bao gồm ít nhất trong phép tính gần đúng đường bao miền thời gian đích và đường bao phổ được xác định bởi chuỗi các khung miền tần số;

nhận tín hiệu được mã hóa, tín hiệu được mã hóa bao gồm sự biểu diễn của chuỗi các khung miền tần số, và sự biểu diễn của đường bao miền thời gian đích,

khi chương trình máy tính nêu trên được chạy bởi máy tính.

24. Vật ghi lưu trữ số không tạm thời bao gồm chương trình máy tính được chứa trên đó để thực hiện phương pháp tách nguồn âm thanh, phương pháp bao gồm:

phương pháp xử lý tín hiệu âm thanh để thu được tín hiệu âm thanh được xử lý, phương pháp bao gồm:

tính toán các giá trị pha cho các giá trị phổ của chuỗi các khung miền tần số biểu diễn các khung chồng lấp của tín hiệu âm thanh,

trong đó các giá trị pha được tính toán dựa trên thông tin về đường bao miền thời gian đích liên quan đến tín hiệu âm thanh được xử lý, để tín hiệu âm thanh được xử lý bao gồm ít nhất trong phép tính gần đúng đường bao miền thời gian đích và đường bao phổ được xác định bởi chuỗi các khung miền tần số, và

tạo màn chắn phổ của tín hiệu âm thanh ban đầu để thu được tín hiệu âm thanh được cải biên được nhập vào thiết bị để xử lý;

trong đó tín hiệu âm thanh được xử lý là tín hiệu nguồn được tách liên quan đến đường bao miền thời gian đích,

khi chương trình máy tính nêu trên được chạy bởi máy tính.

25. Vật ghi lưu trữ số không tạm thời bao gồm chương trình máy tính được chứa trên đó để thực hiện phương pháp nâng cao băng thông của tín hiệu âm thanh được mã hóa, phương pháp bao gồm:

tạo ra tín hiệu nâng cao từ băng tín hiệu âm thanh nằm trong tín hiệu được mã hóa;

phương pháp xử lý tín hiệu âm thanh để thu được tín hiệu âm thanh được xử lý, phương pháp bao gồm:

tính toán các giá trị pha cho các giá trị phổ của chuỗi các khung miền tần số biểu diễn các khung chồng lấp của tín hiệu âm thanh,

trong đó các giá trị pha được tính toán dựa trên thông tin về đường bao miền thời gian đích liên quan đến tín hiệu âm thanh được xử lý, để tín hiệu âm thanh được xử lý bao gồm ít nhất trong phép tính gần đúng đường bao miền thời gian đích và đường bao phổ được xác định bởi chuỗi các khung miền tần số;

trong đó tạo ra bao gồm trích đường bao miền thời gian đích từ sự biểu diễn được mã hóa nằm trong tín hiệu được mã hóa hoặc từ băng tín hiệu âm thanh nằm trong tín hiệu được mã hóa,

khi chương trình máy tính nêu trên được chạy bởi máy tính.

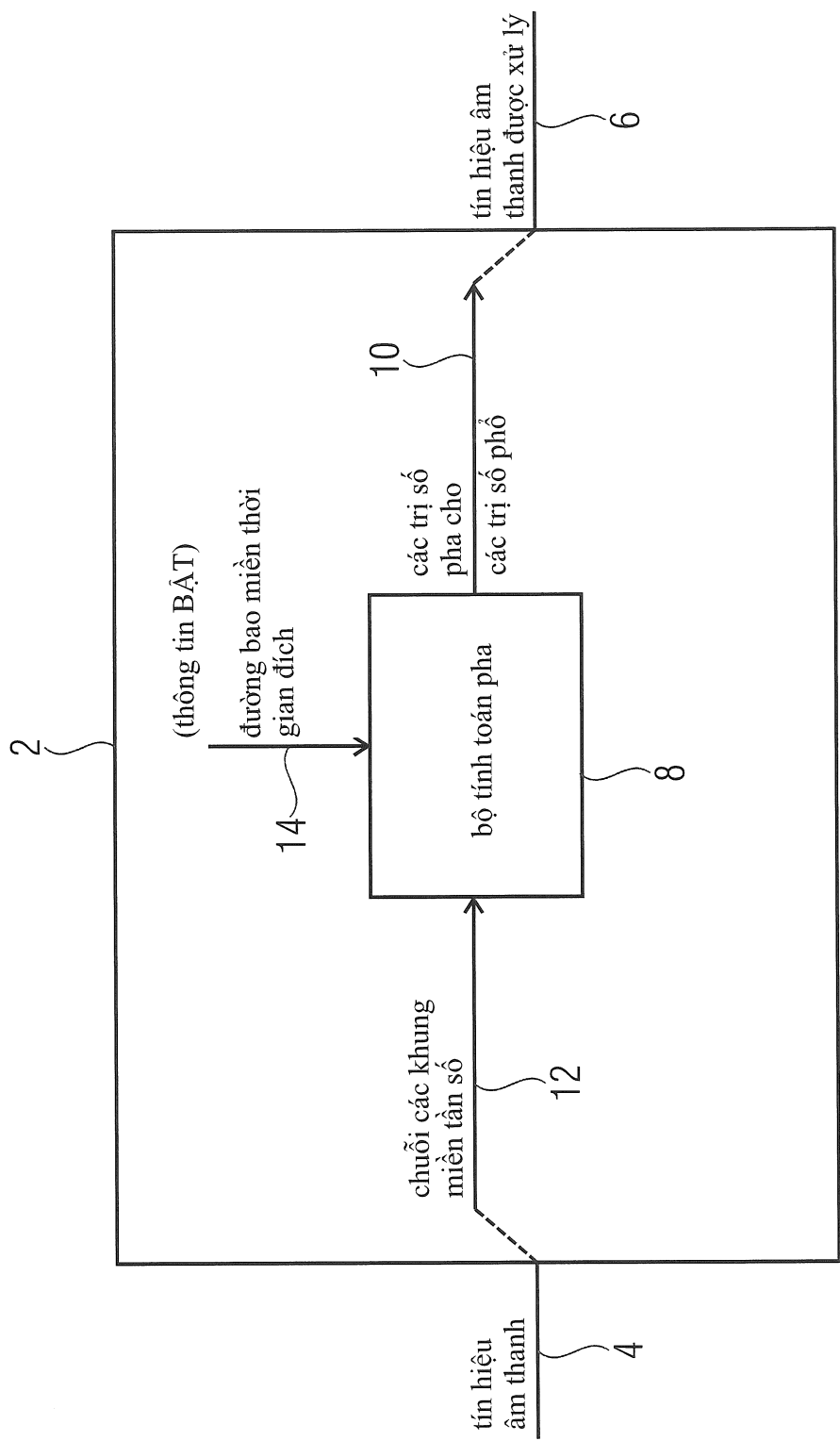


FIG 1

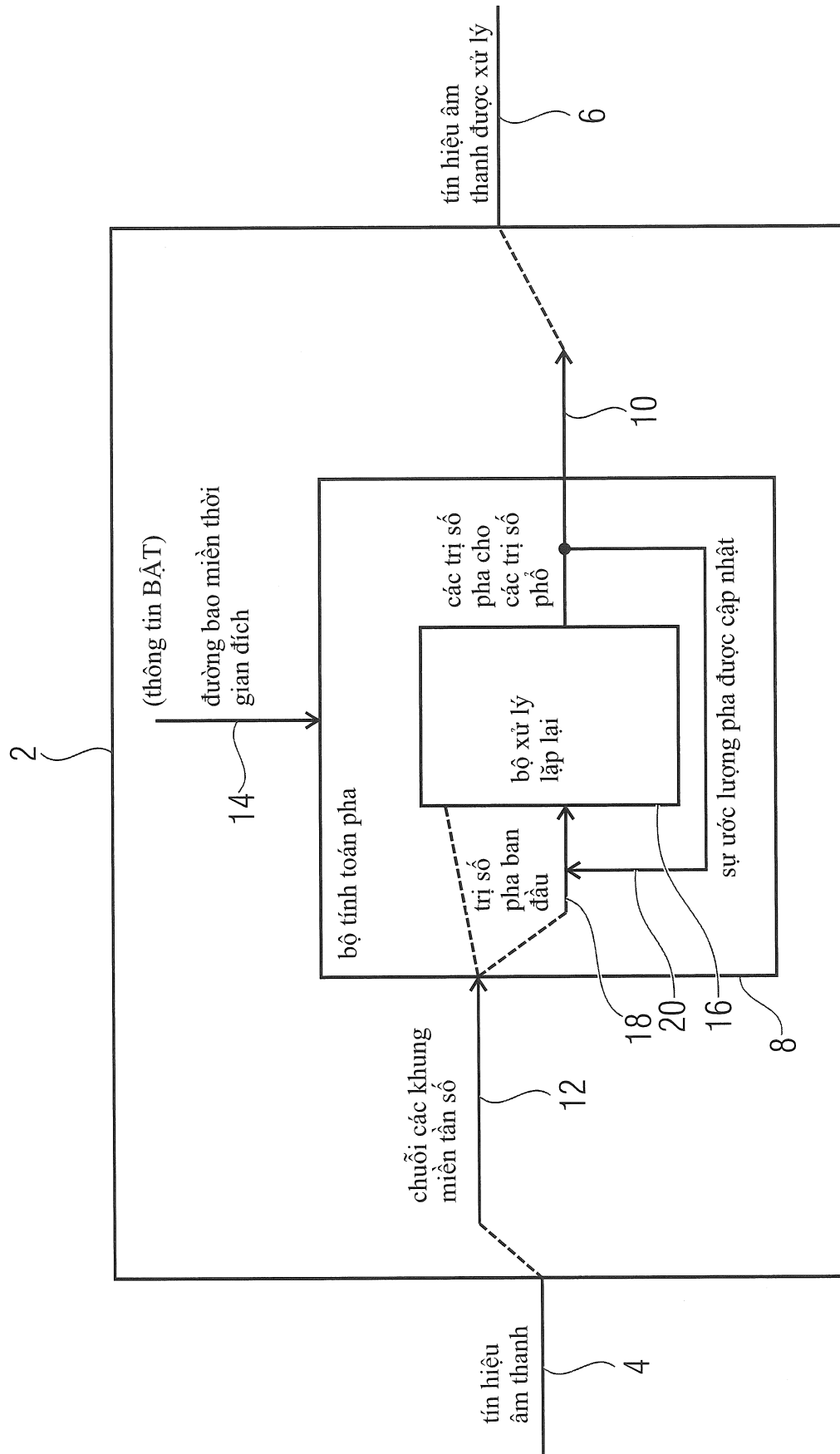


FIG 2

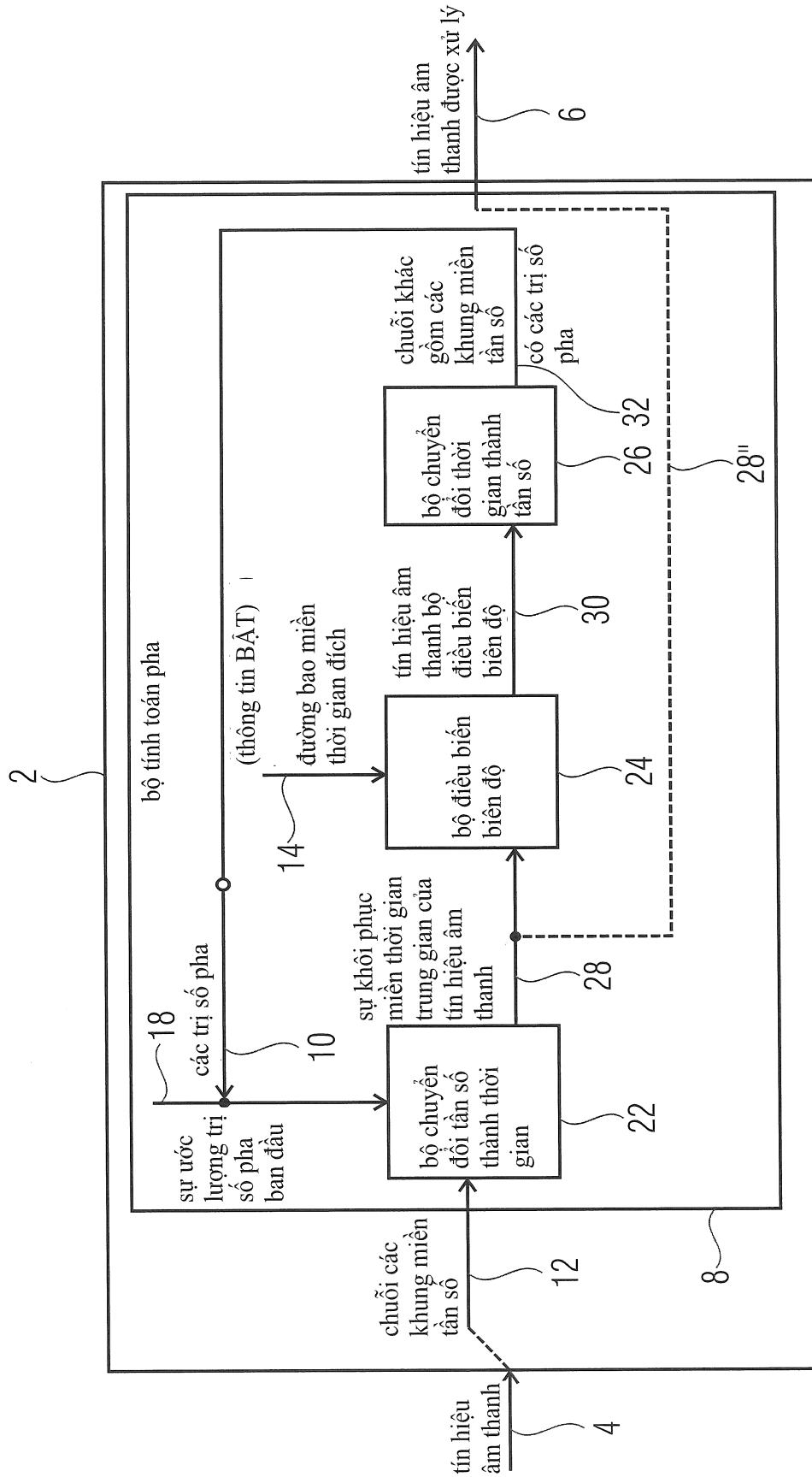


FIG 3

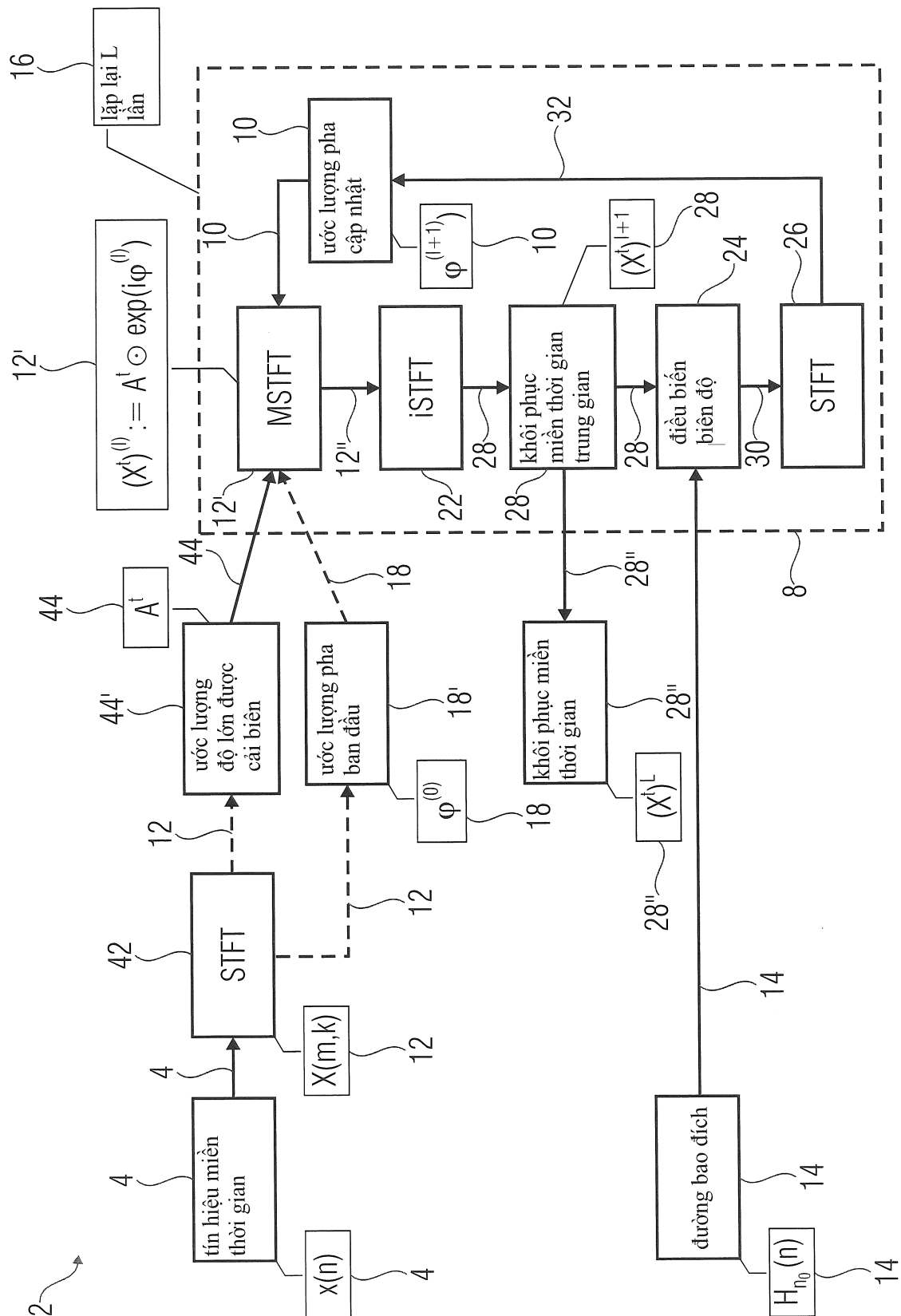


FIG 5

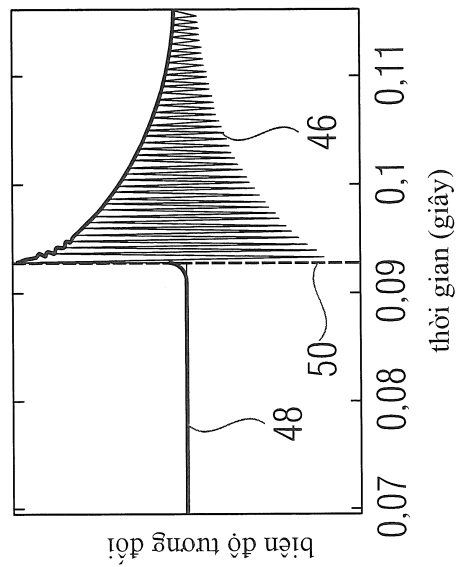


FIG 6A

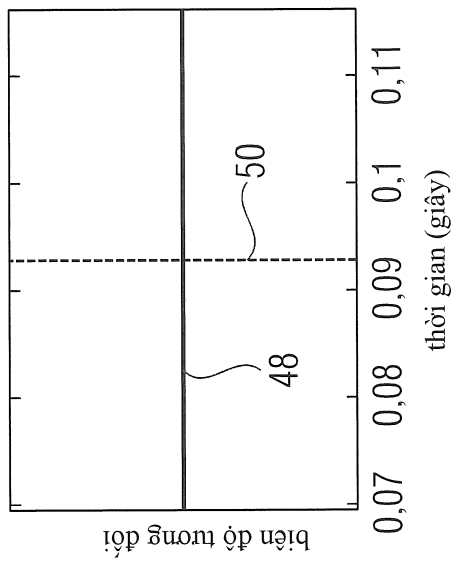


FIG 6B

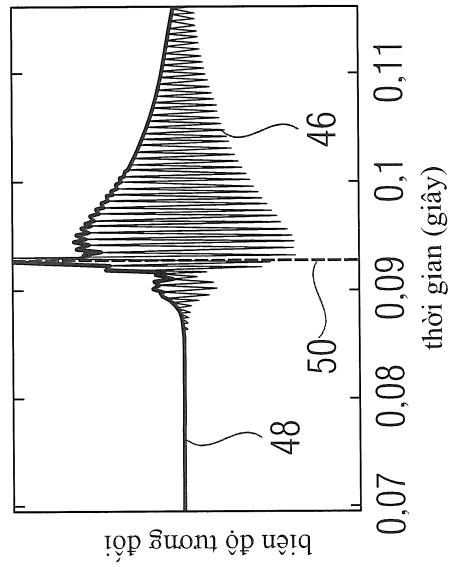


FIG 6C

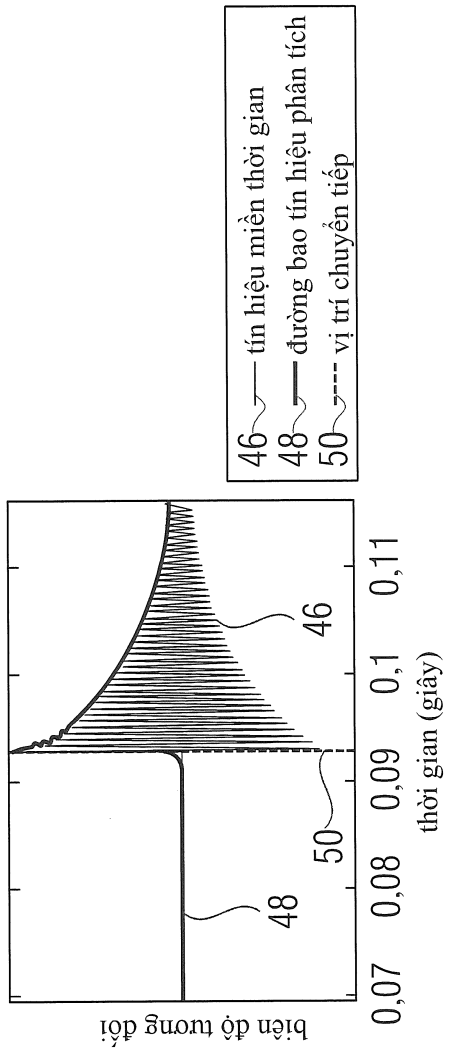


FIG 6D

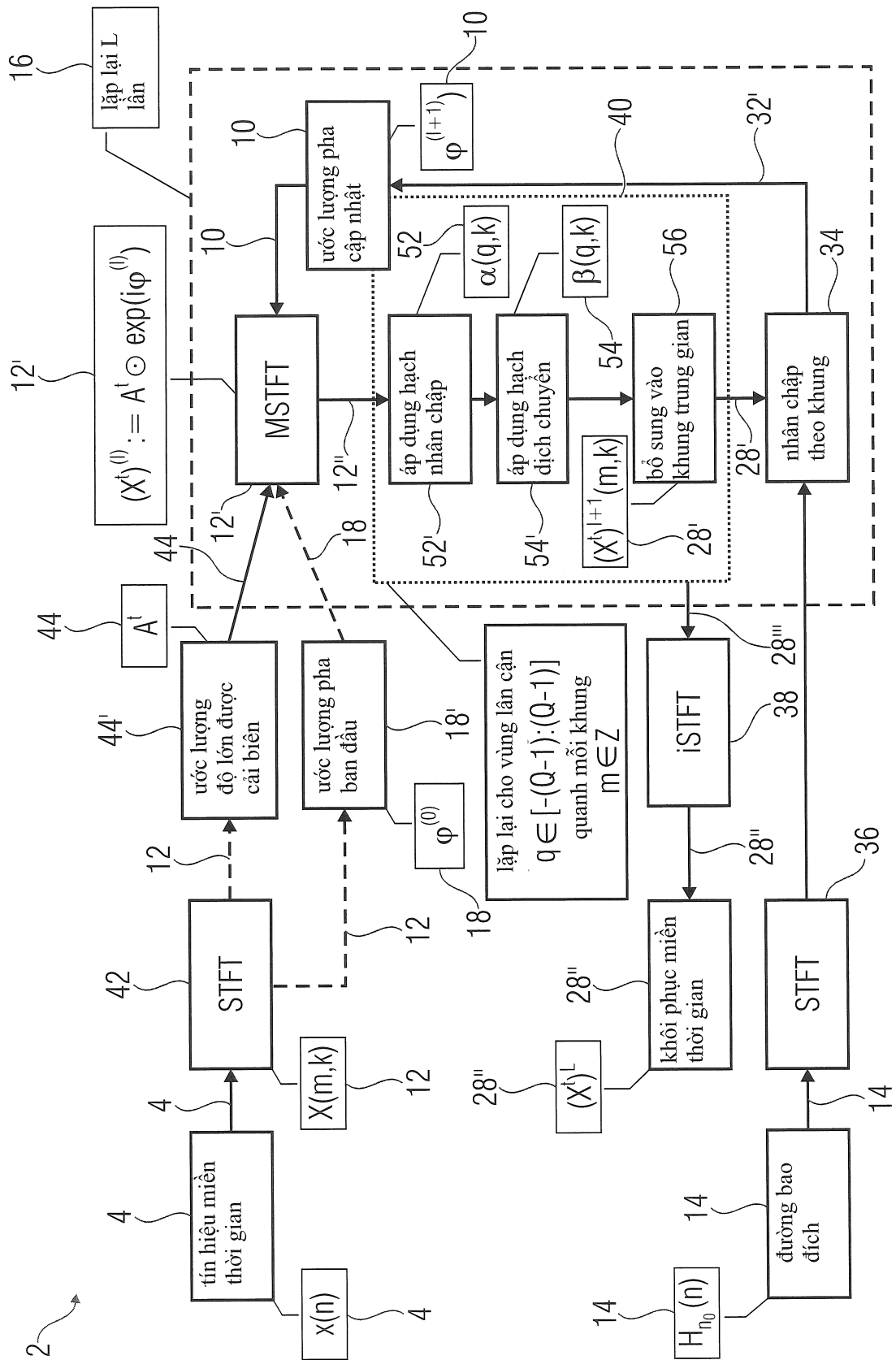


FIG 7

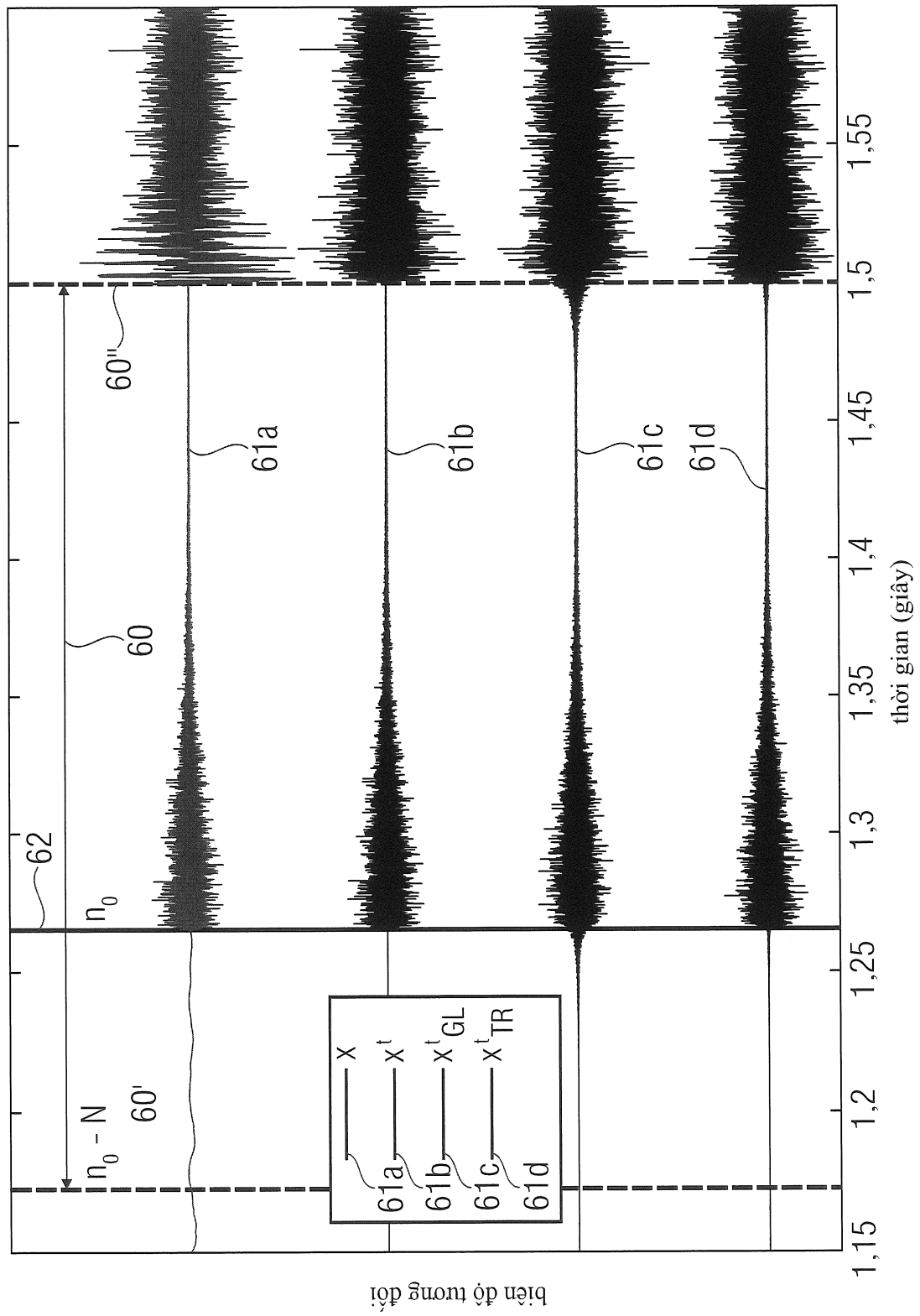


FIG 8

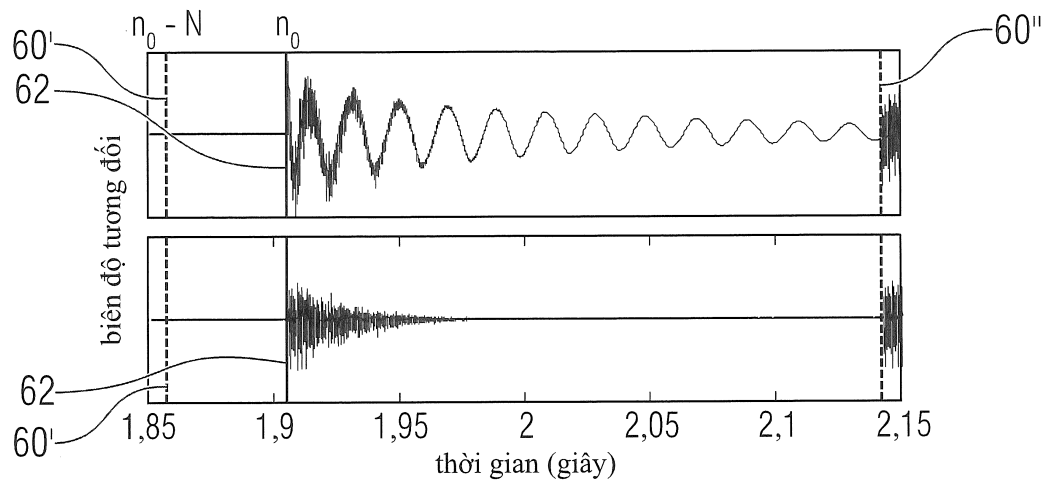


FIG 9A

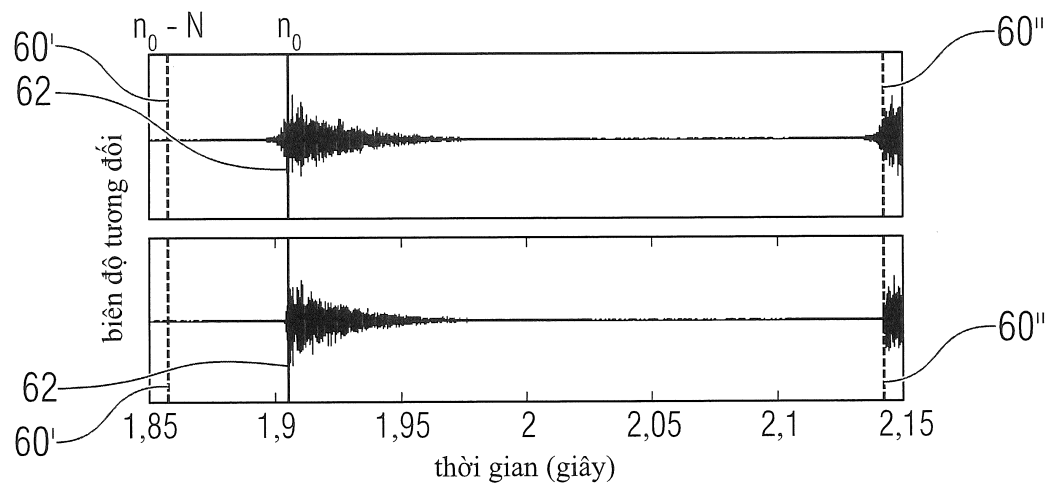


FIG 9B

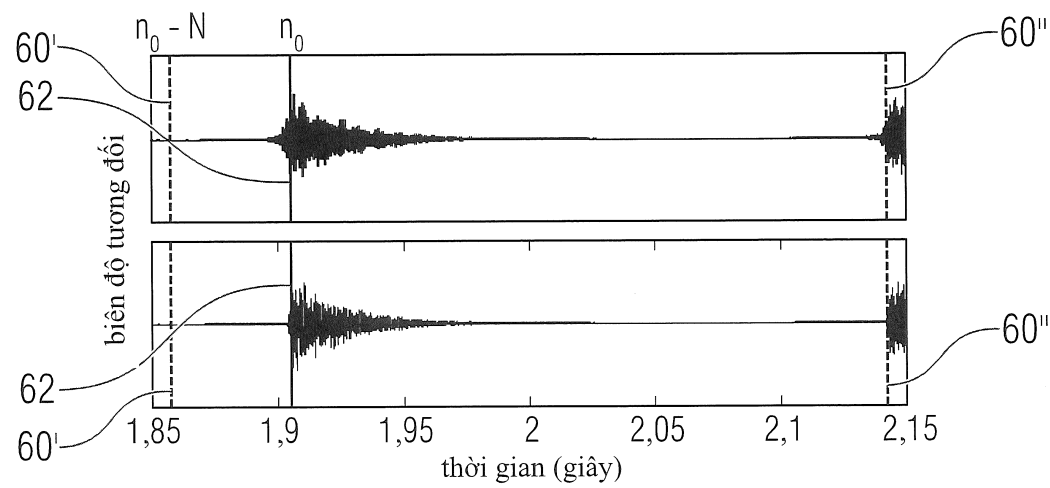


FIG 9C

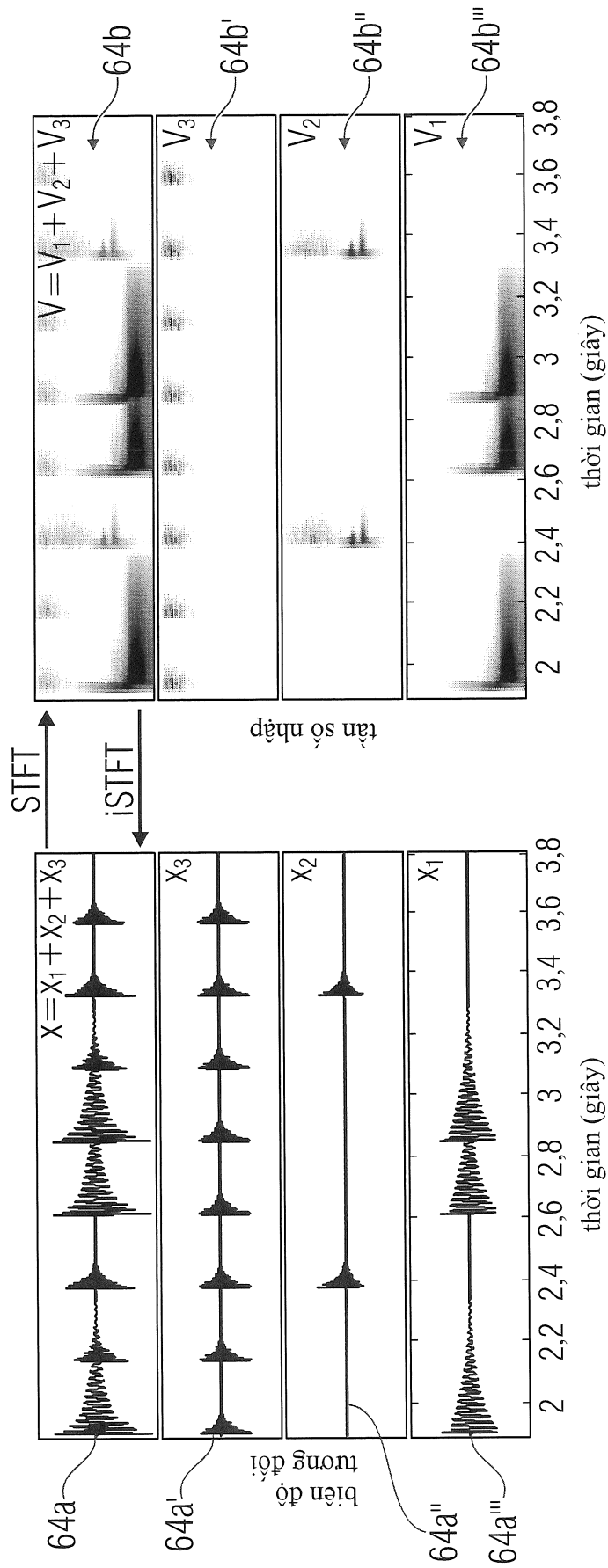


FIG 10B

FIG 10A

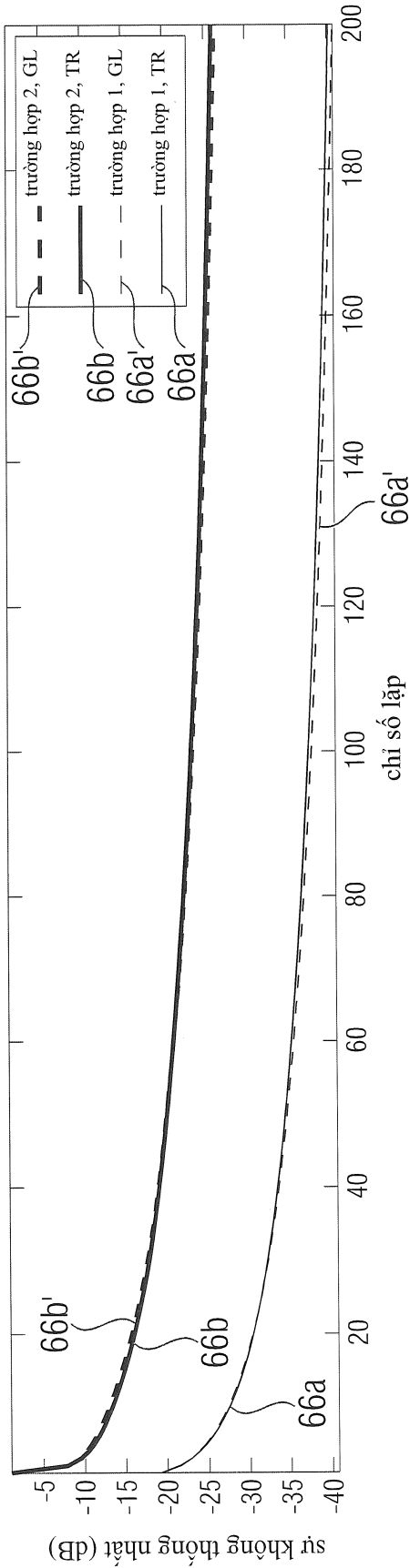


FIG 11A

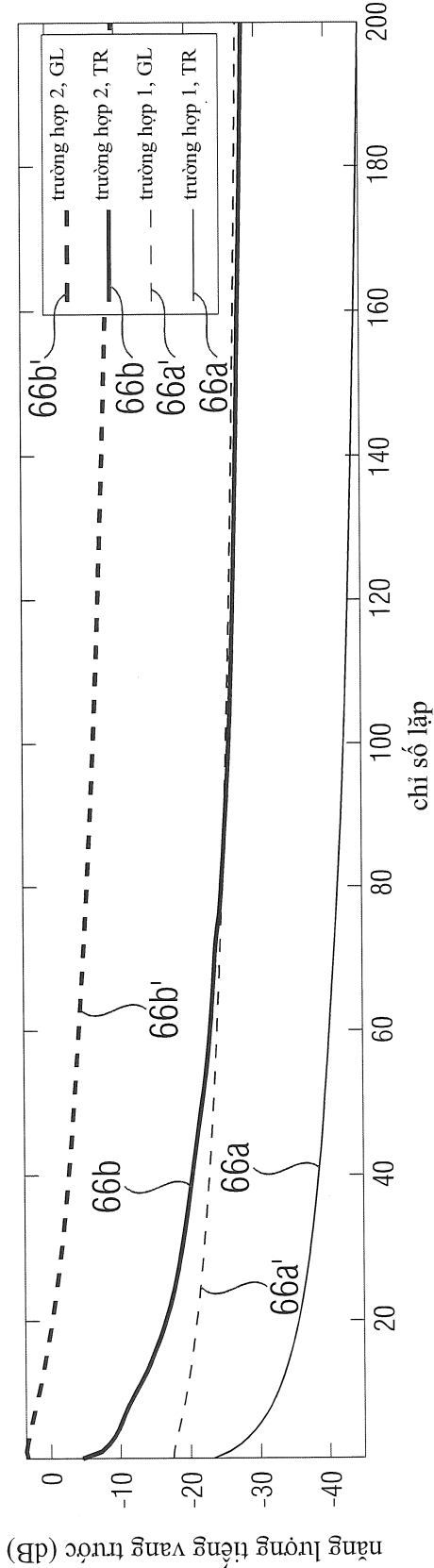


FIG 11B

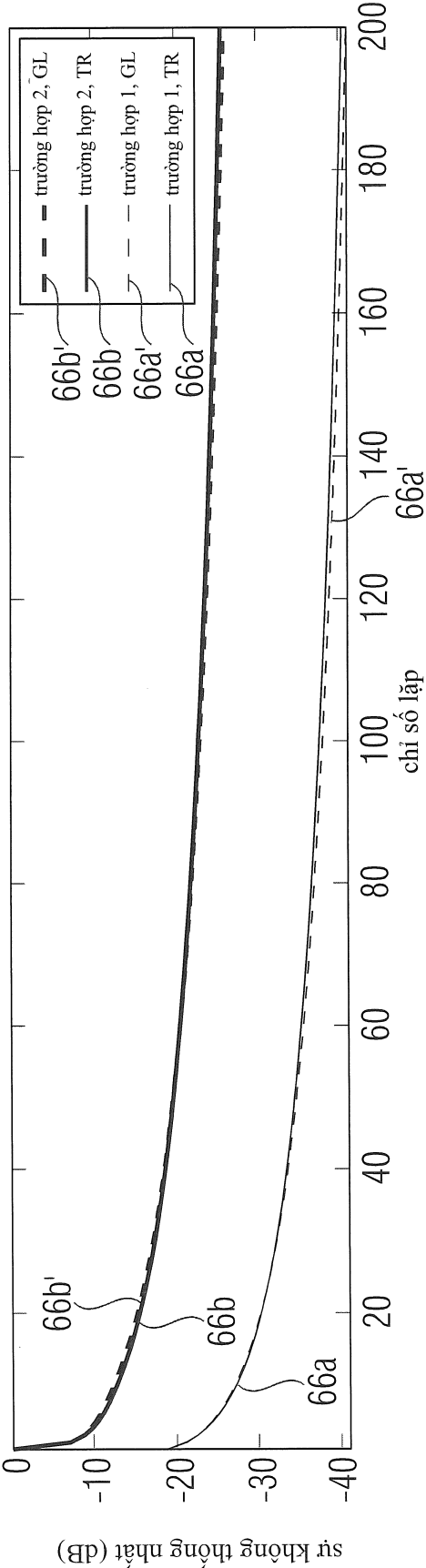


FIG 12A

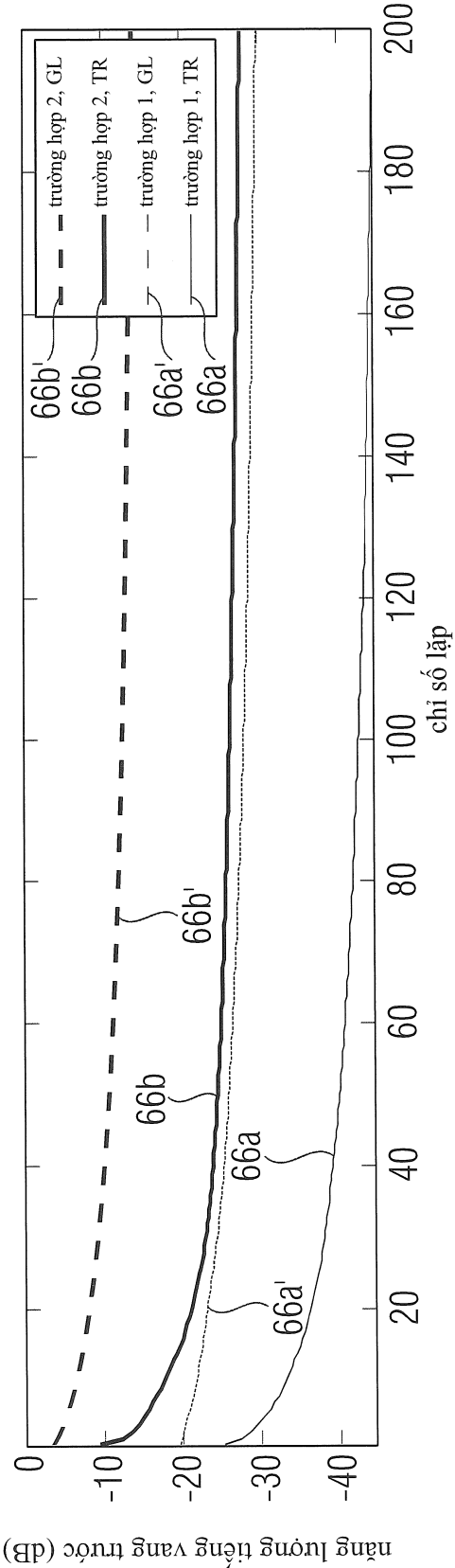


FIG 12B

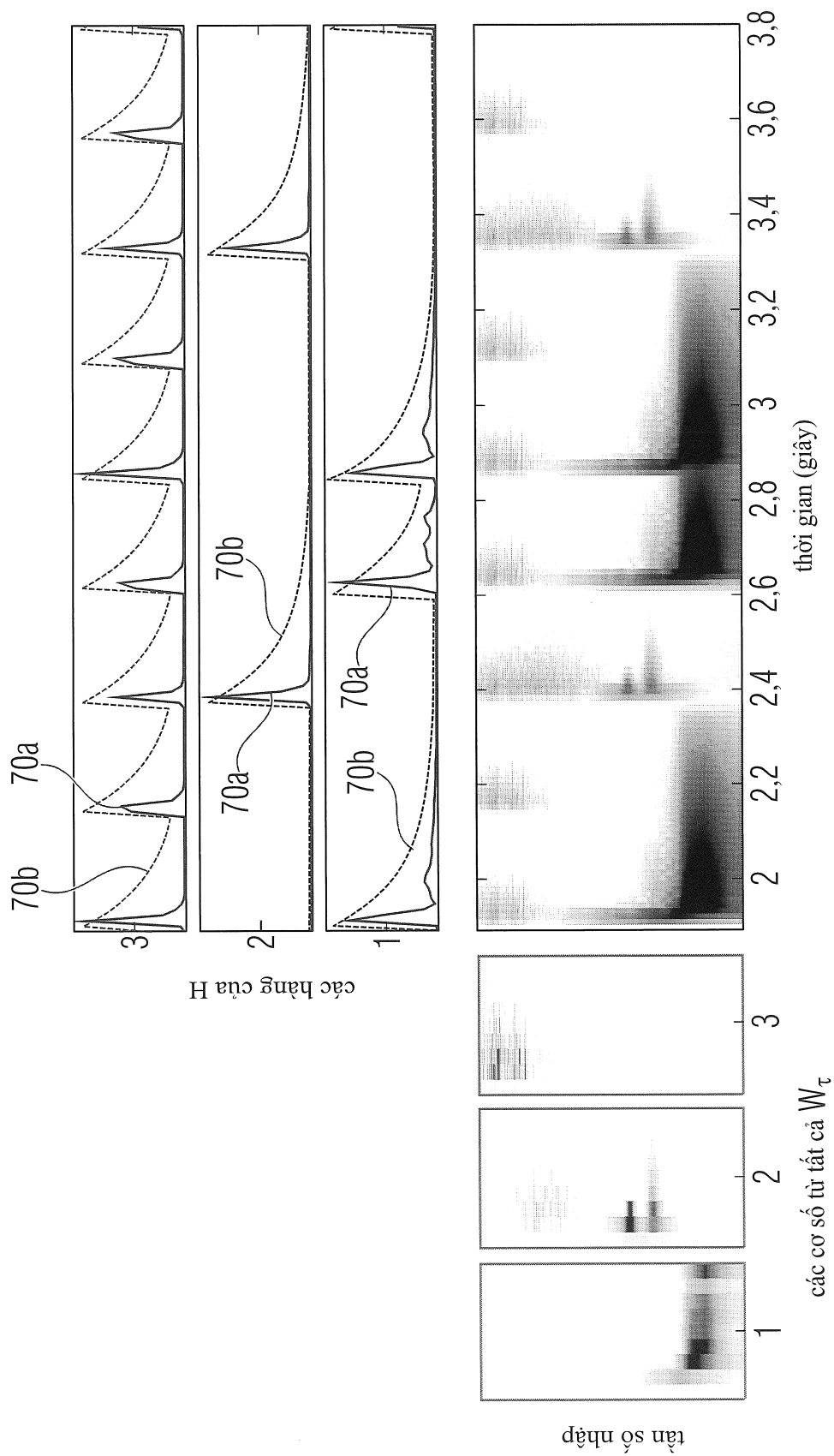


FIG 13

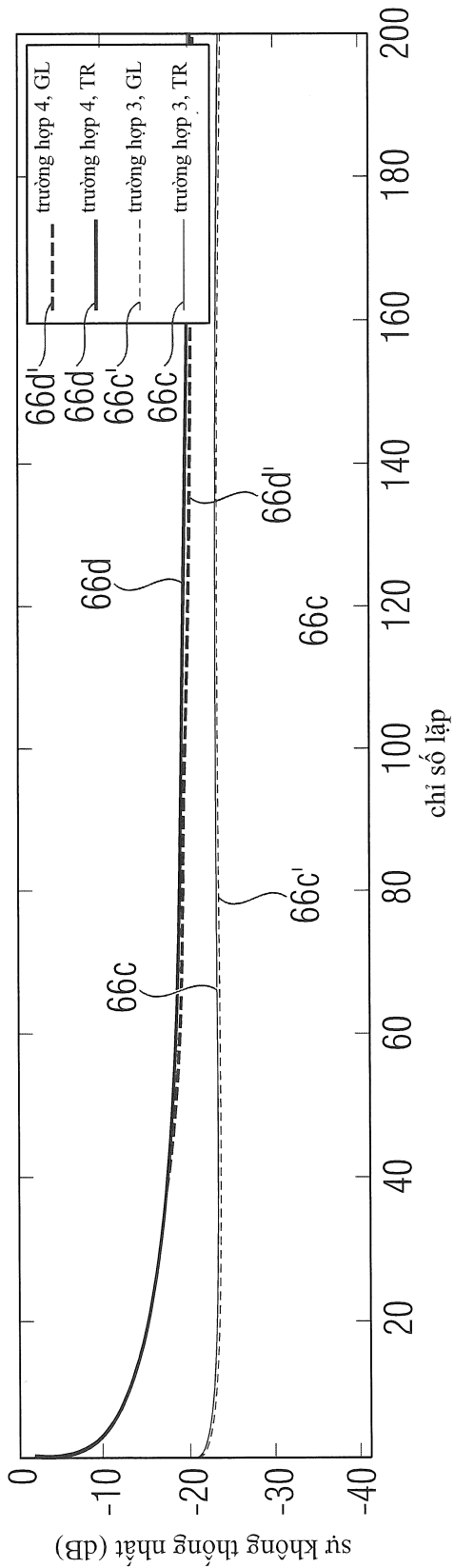


FIG 14A

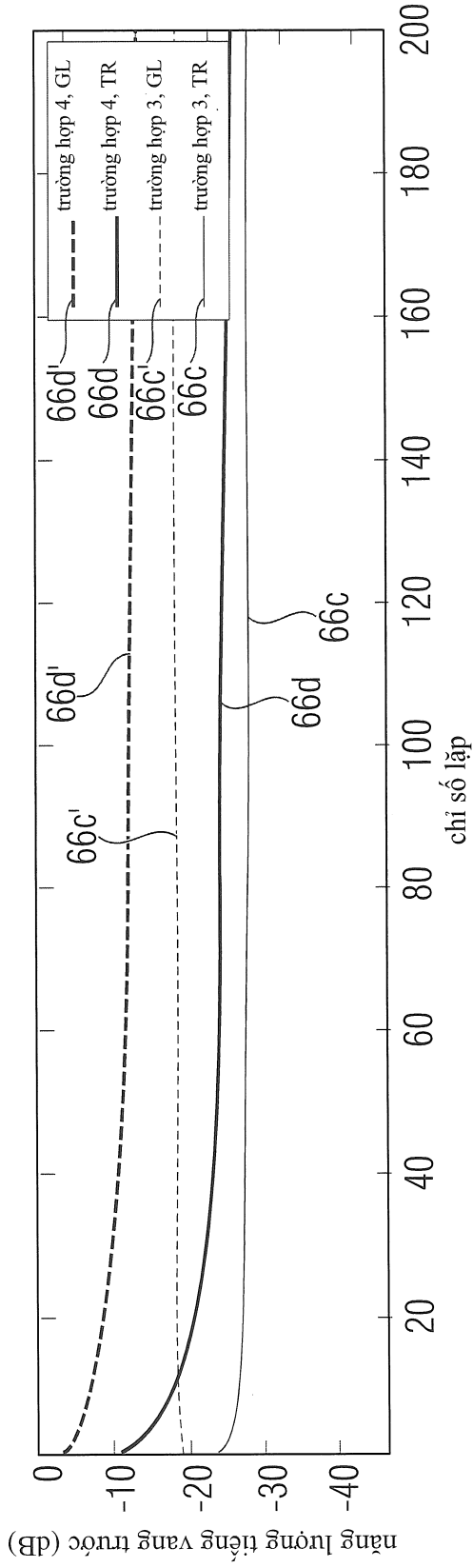


FIG 14B

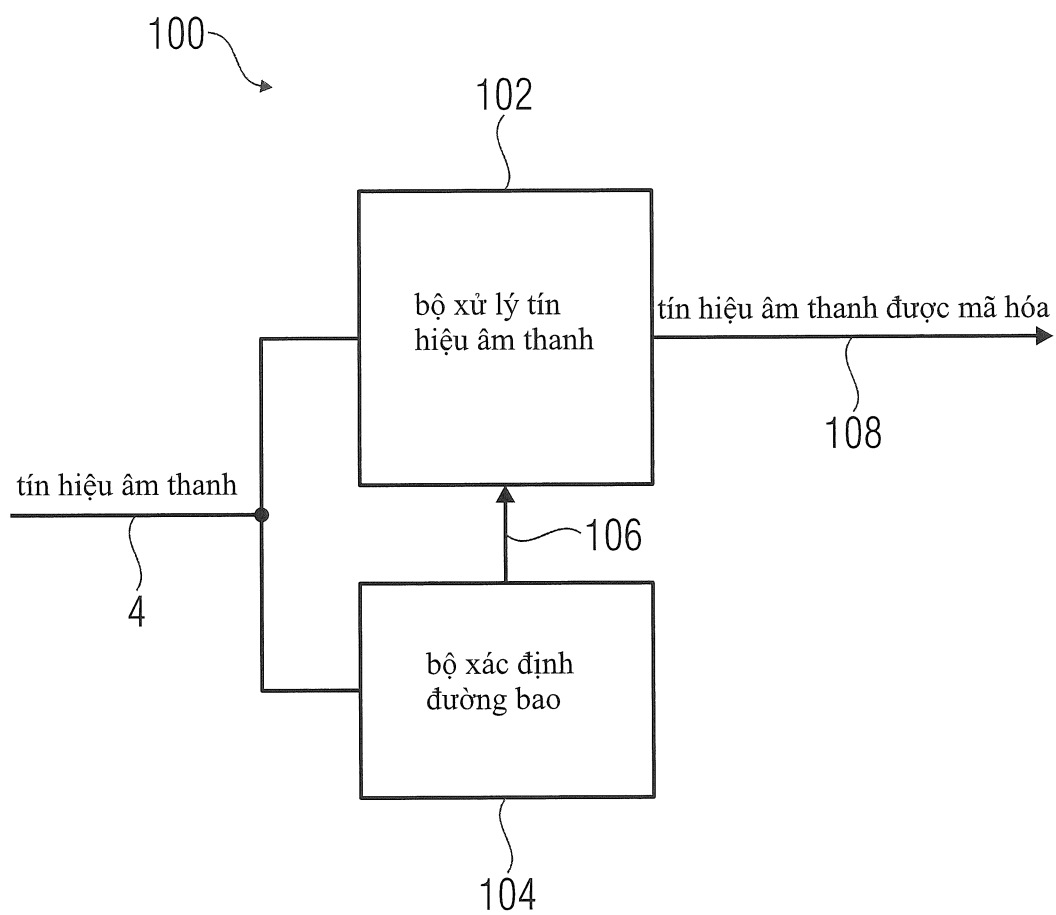


FIG 15

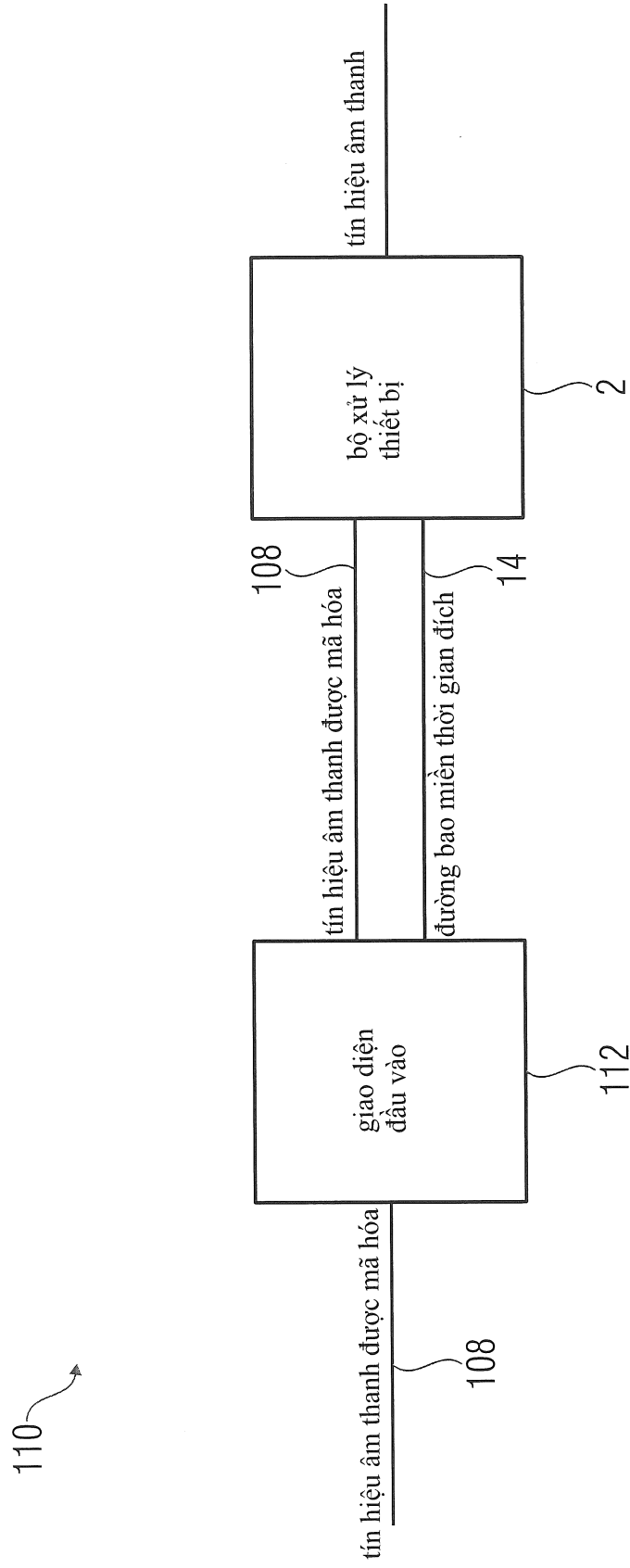


FIG 16

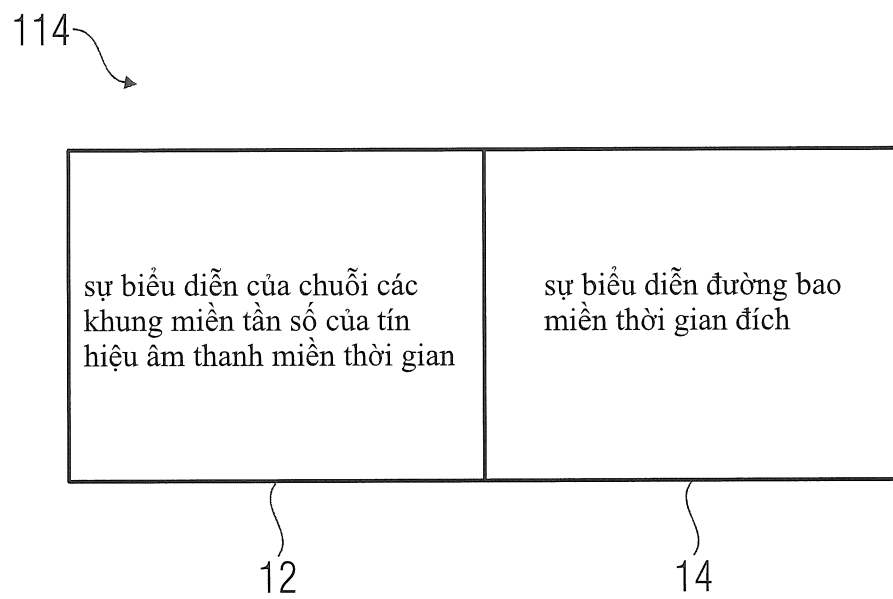


FIG 17

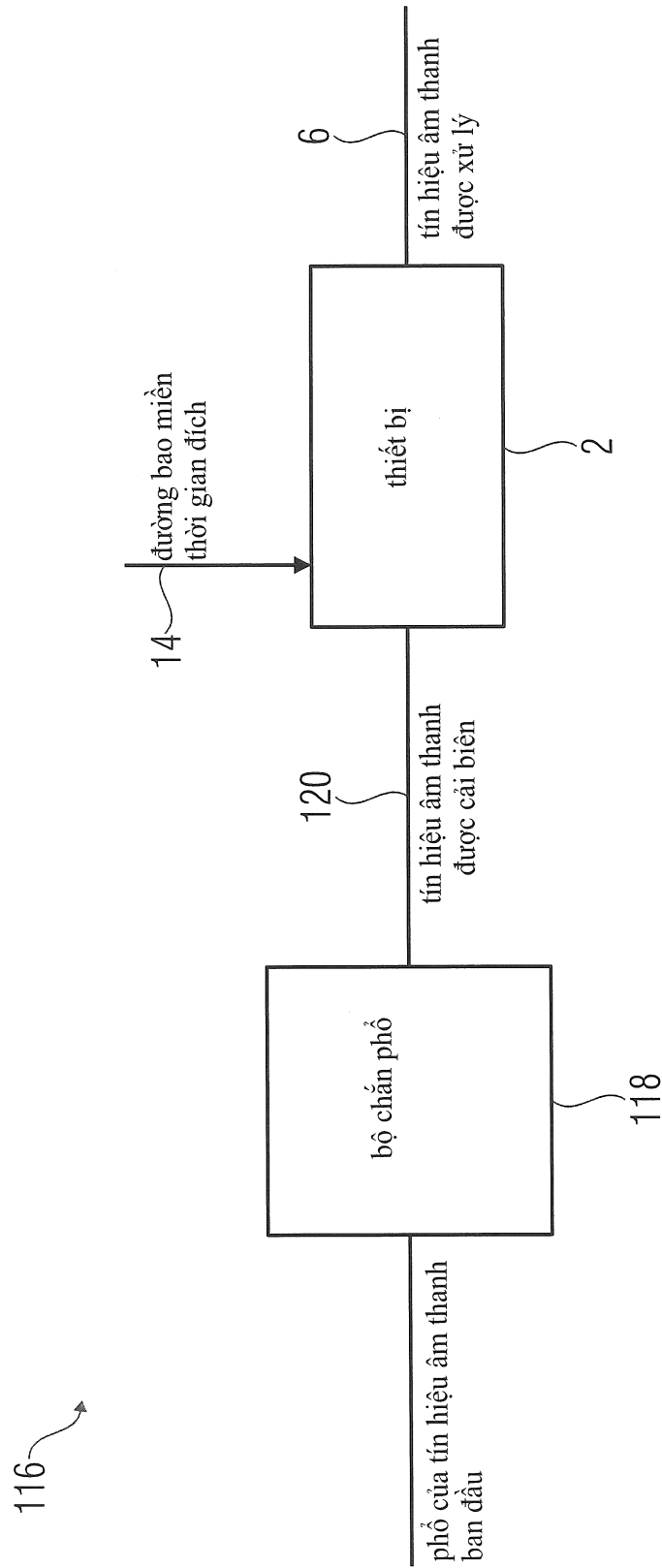


FIG 18

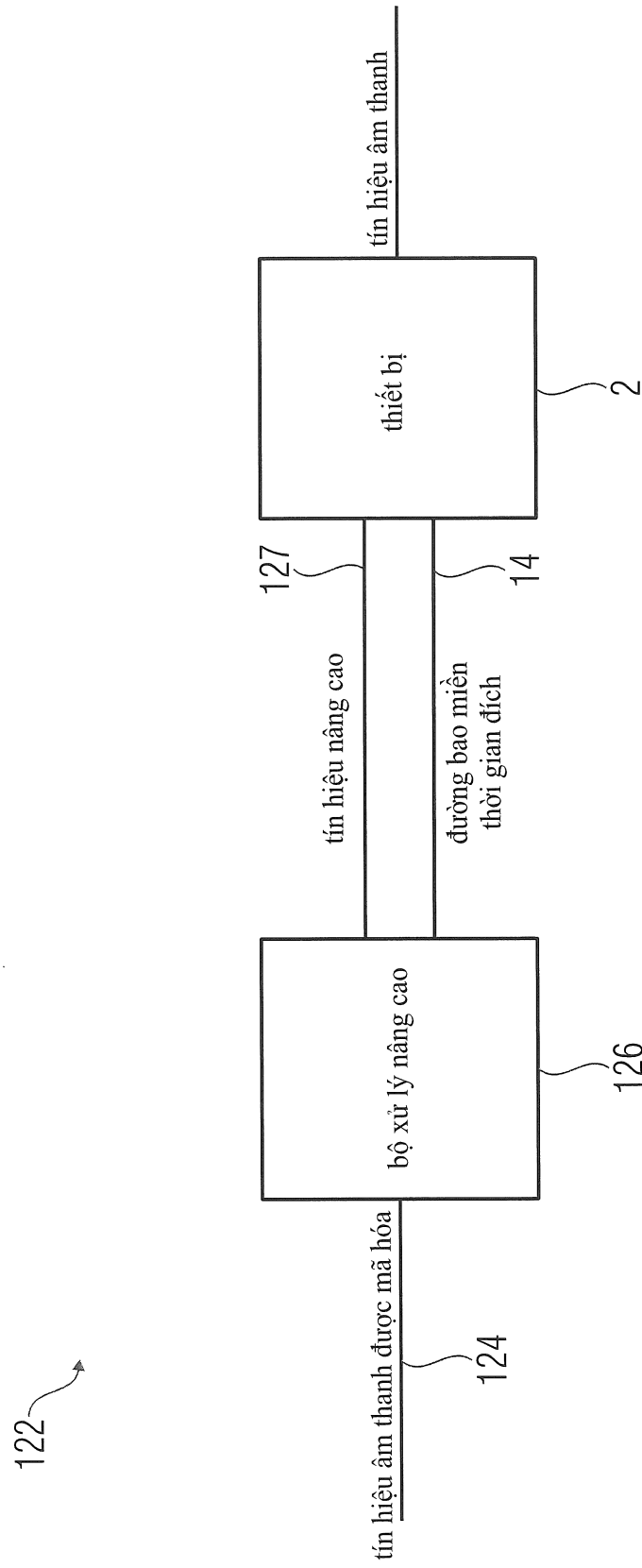


FIG 19

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

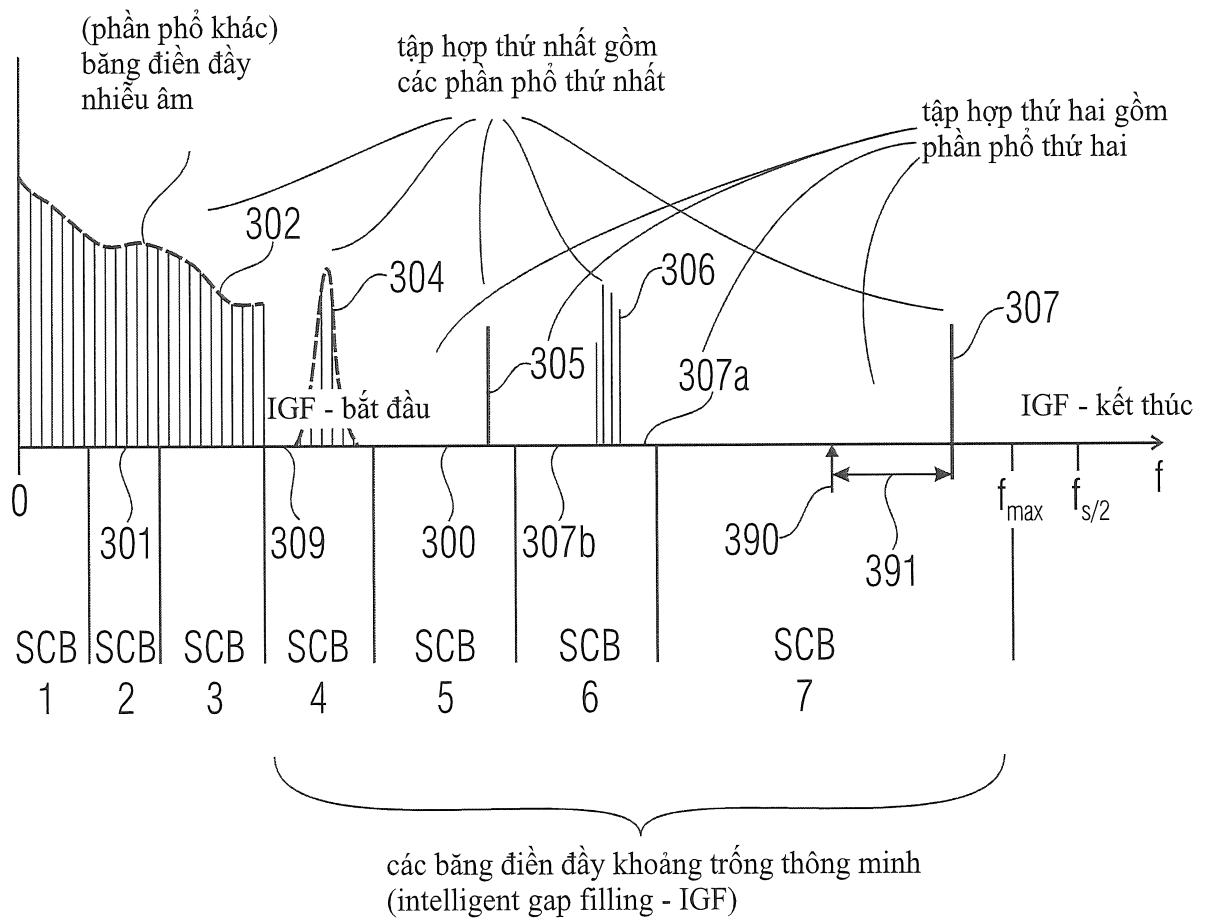


FIG 20

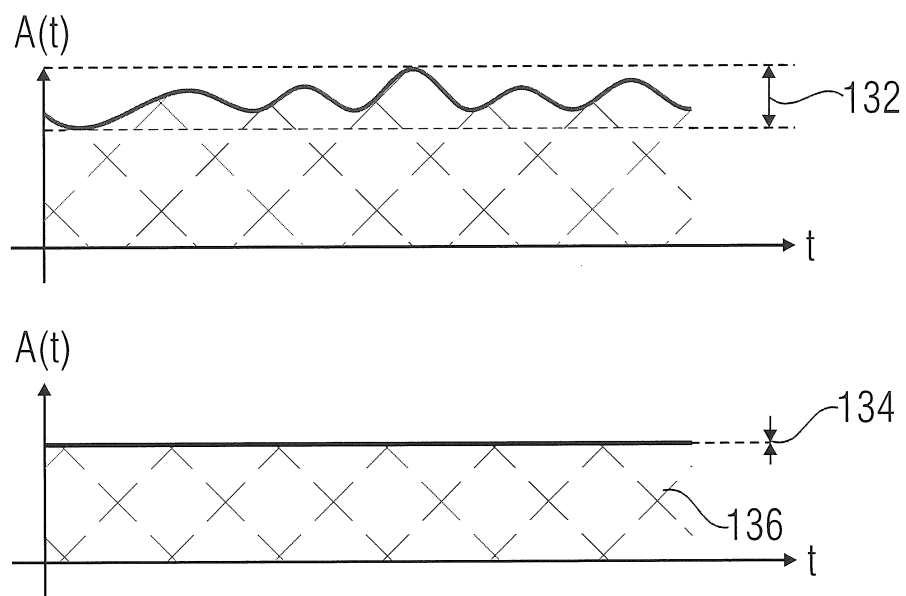


FIG 21

2200

Tính toán các trị số pha cho các trị số phổ của chuỗi các khung miền tần số biểu diễn các khung chồng lấp của tín hiệu âm thanh, trong đó các trị số pha được tính toán dựa trên thông tin về đường bao miền thời gian đích liên quan đến tín hiệu âm thanh được xử lý, để tín hiệu âm thanh được xử lý có ít nhất trong xấp xỉ đường bao miền thời gian đích và đường bao phổ được xác định bằng chuỗi các khung miền tần số

2205

FIG 22

2300

Phương pháp 2200

2305

Nhận tín hiệu được mã hóa, tín hiệu được mã hóa bao gồm sự biểu diễn chuỗi các khung miền tần số, và sự biểu diễn của đường bao miền thời gian

2310

FIG 23

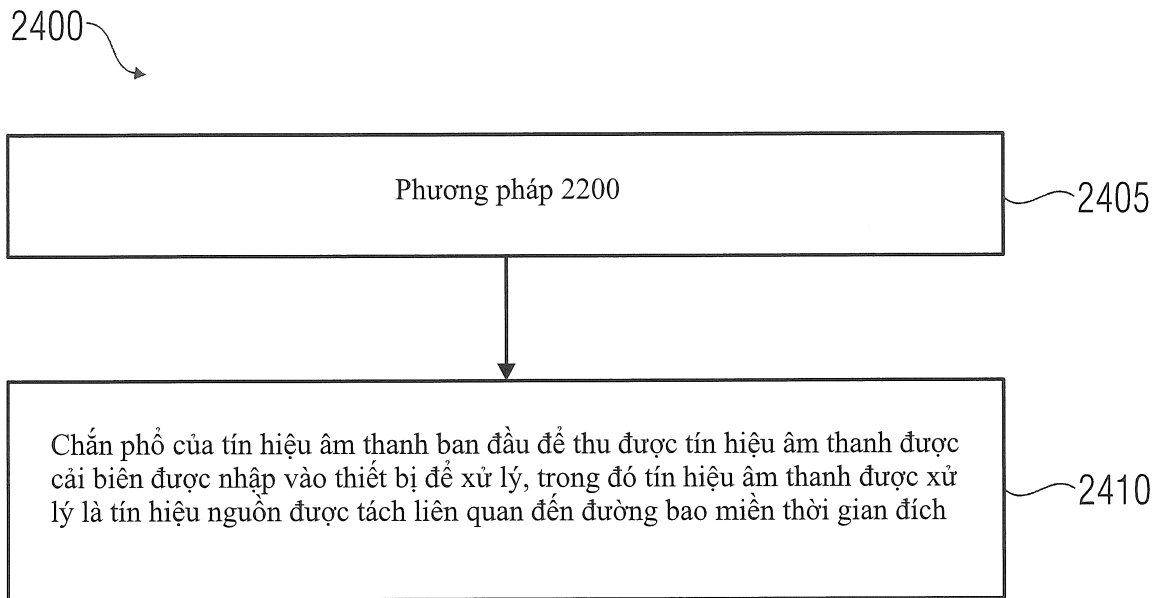


FIG 24

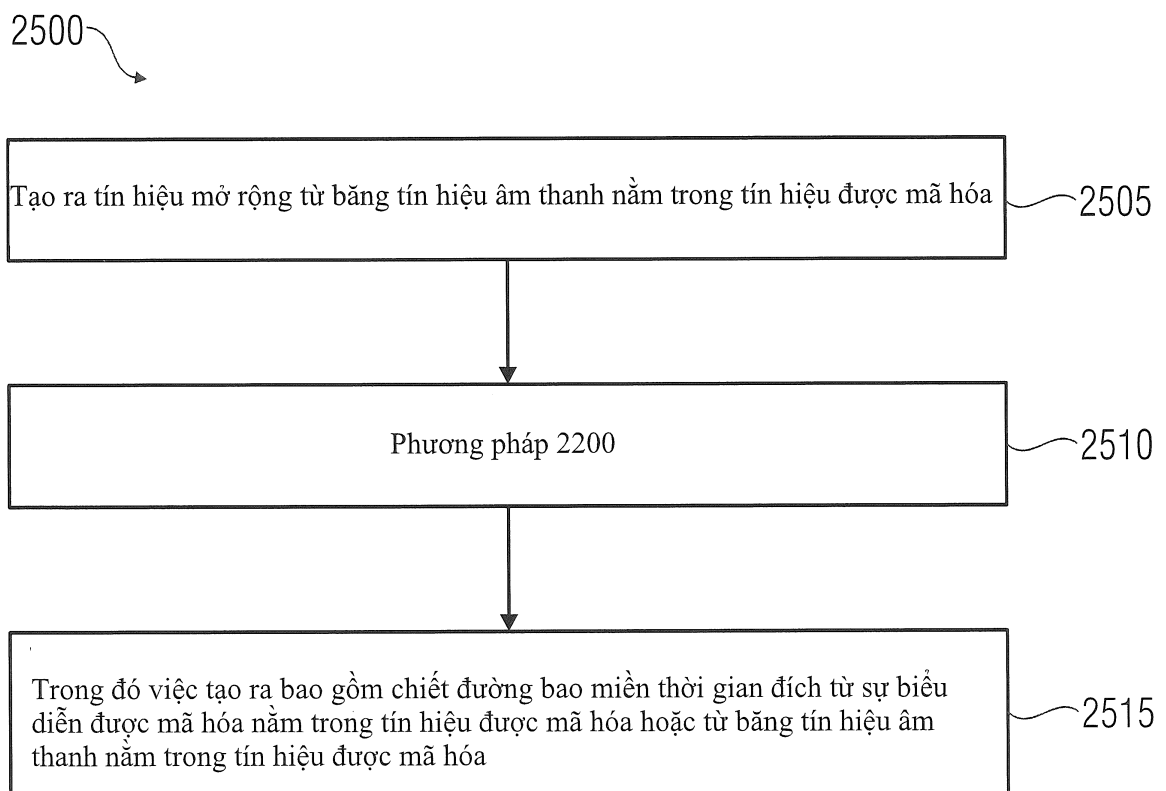


FIG 25

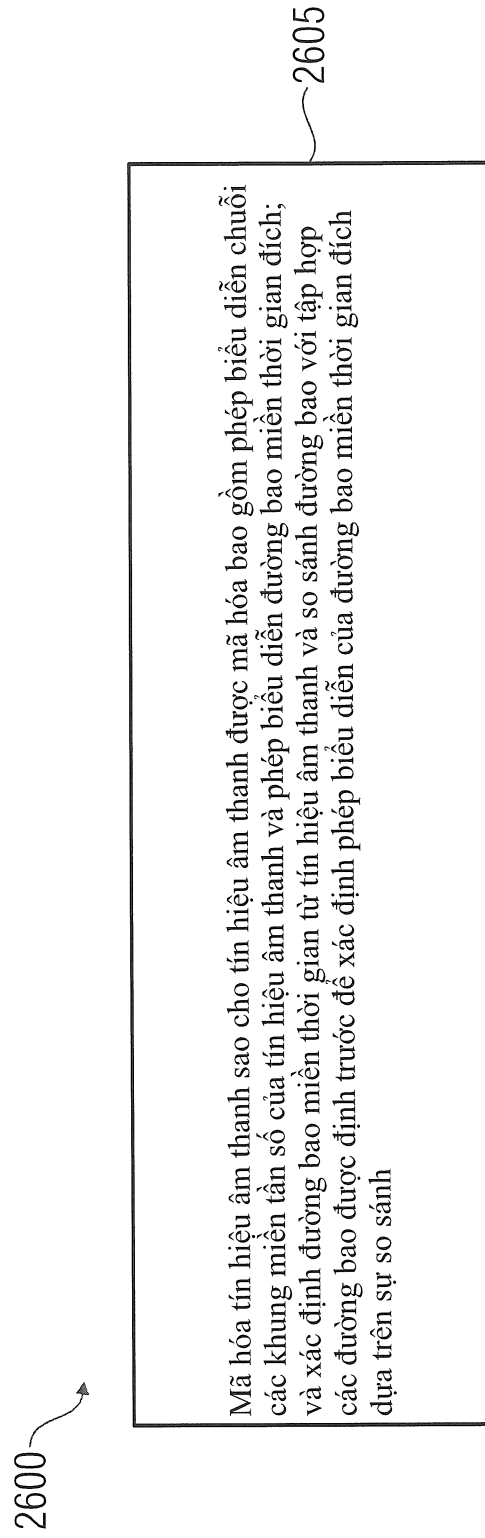


FIG 26