(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) **1-0051645**

(51)[2020.01] **G10L 21/038**; G10L 21/0316 (13) **B**

(21) 1-2021-04356 (22) 12/12/2019
(86) PCT/EP2019/084974 12/12/2019 (87) WO2020/126857 25/06/2020
(30) 18215691.9 21/12/2018 EP; 19166643.7 01/04/2019 EP
(45) 25/09/2025 450 (43) 25/11/2021 404A
(73) Fraunhofer-Gesellschaft zur Foerderung der angewandten Forschung e.V. (DE)
Hansastraße 27c, 80686 Muenchen, Germany (DE)
(72) DISCH, Sascha (DE); STURM, Michael (DE).
(74) CÔNG TY LUẬT TRÁCH NHIỆM HỮU HẠN AMBYS HÀ NỘI (AMBYS HANOI)

(54) BỘ XỬ LÝ ÂM THANH, PHƯƠNG PHÁP TẠO TÍN HIỆU ÂM THANH ĐƯỢC NÂNG CAO TẦN SỐ SỬ DỤNG XỬ LÝ XUNG

(21) 1-2021-04356

(57) Sáng chế đề cập đến bộ xử lý âm thanh, phương pháp tạo ra tín hiệu âm thanh được nâng cao tần số sử dụng xử lý xung. Bộ xử lý âm thanh để tạo ra tín hiệu âm thanh được nâng cao tần số (420) từ tín hiệu âm thanh nguồn (30), bộ xử lý âm thanh bao gồm: bộ xác định đường bao (100) để xác định đường bao theo thời gian của ít nhất một phần tín hiệu âm thanh nguồn; bộ phân tích (200) để phân tích đường bao theo thời gian để xác định các giá trị theo thời gian của các đặc điểm nhất định của đường bao theo thời gian; bộ tổng hợp tín hiệu (300) để tạo ra tín hiệu tổng hợp, việc tạo ra bao gồm đặt các xung liên quan đến các giá trị theo thời gian được xác định, trong đó các xung được gán trọng số sử dụng các trọng số được suy ra từ các biên độ của đường bao theo thời gian liên quan đến các giá trị theo thời gian, trong đó các xung được đặt; và bộ tổ hợp (400) để tổ hợp ít nhất băng của tín hiệu tổng hợp không có trong tín hiệu âm thanh nguồn và tín hiệu âm thanh nguồn để thu được tín hiệu âm thanh được nâng cao tần số (420).

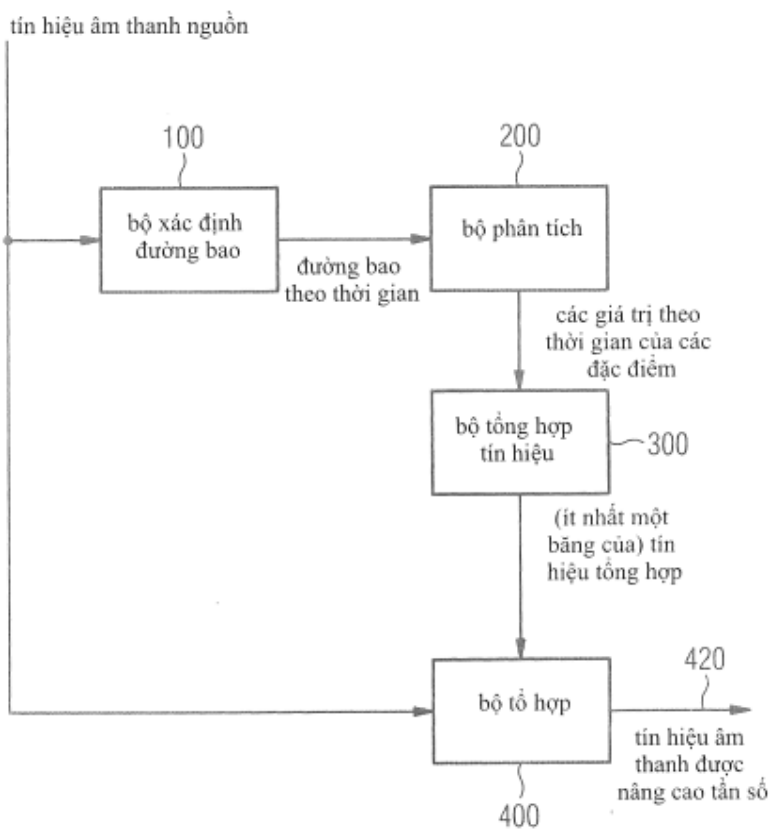

Fig. 1

Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến việc xử lý tín hiệu âm thanh và cụ thể là các giải pháp để tạo ra tín hiệu âm thanh được nâng cao tần số từ tín hiệu âm thanh nguồn.

Tình trạng kỹ thuật của sáng chế

Việc lưu trữ và truyền các tín hiệu âm thanh thường lệ thuộc vào giới hạn tốc độ truyền bit nghiêm ngặt. Trong quá khứ, bộ mã hóa bị buộc phải cắt giảm mạnh băng thông âm thanh được truyền khi chỉ có sẵn tốc độ truyền bit rất thấp. Các bộ mã hóa-giải mã âm thanh điều giải ngày nay có thể mã hóa nên các tín hiệu băng thông rộng bằng cách sử dụng các phương pháp mở rộng băng thông (bandwidth extention - BWE) [1-2].

Các thuật toán này phụ thuộc vào sự biểu diễn tham số của nội dung tần số cao (high frequency – HF), được tạo ra từ phần tần số thấp (low frequency – LF) được mã hóa dạng sóng của tín hiệu được giải mã bằng các công cụ chuyển vị thành vùng phổ HF ("sự ráp nối"). Trong khi thực hiện điều này, đầu tiên, bản vá "thô" được tạo ra, và tiếp theo, sự xử lý sau theo tham số được áp dụng lên bản vá "thô".

Thông thường, sự xử lý sau này được áp dụng để điều chỉnh các đặc tính cảm quan quan trọng chưa được xét đến trong khi tạo ra tần số cao thông qua sự chuyển vị và do đó cần được điều chỉnh trên bản vá "thô" kết quả thu được sau đó.

Tuy nhiên, nếu, ví dụ, cấu trúc mịn phổ trong bản vá được sao chép vào vùng đích nào đó có sự khác biệt lớn so với cấu trúc mịn phổ của nội dung ban đầu, các thành phần lạ không mong muốn có thể làm ảnh hưởng và suy giảm chất lượng cảm quan của tín hiệu âm thanh được giải mã. Trong các trường hợp này, sự xử lý sau được áp dụng thường không thể hiệu chỉnh hoàn toàn các đặc tính sai của bản vá "thô".

Tài liệu tham khảo

1. Dietz, M., Liljeryd, L., Kjörling, K., and Kunz, 0., "Spectral Band Replication, a Novel Approach in Audio Coding," in Audio Engineering Society Convention 112, 2002.

2. Disch, S., Niedermeier, A., Helmrich, C. R., Neukam, C., Schmidt, K., Geiger, R., Lecomte, J., Ghido, F., Nagel, F., and Edler, B., "Intelligent Gap Filling in Perceptual Transform Coding of Audio," in Audio Engineering Society Convention 141, 2016.

3. Laitinen M-V., Disch S., Oates C., Pulkki V. "Phase derivative correction of bandwidth extended signals for perceptual audio codecs." In 140th Audio Engineering Society International Convention 2016, AES 2016. Audio Engineering Society. 2016.

4. Atti, Venkatraman, Venkatesh Krishnan, Duminda A. Dewasurendra, Venkata Chebiyyam, Shaminda Subasingha, Daniel J. Sinder, Vivek Rajendran, Imre Varga, Jon Gibbs, Lei Miao, Volodya Grancharov and Harald Pobloth. "Super-wideband bandwidth extension for speech in the 3GPP EVS codec." 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (2015): 5927-5931.

Bản chất kỹ thuật của sáng chế

Mục đích của sáng chế nhằm cải thiện chất lượng cảm quan thông qua việc tạo ra mang tính thích ứng tín hiệu mới có nội dung bản vá "thô" tín hiệu tần số cao điền đầy khoảng trống hoặc được ước lượng, mà được thích ứng về cảm quan với tín hiệu LF.

Bằng cách thu được sẵn tín hiệu "thô" được thích ứng về cảm quan, các ước số hiệu chỉnh sau cần thiết khác sẽ được giảm thiểu. Hơn nữa, tín hiệu thô được thích ứng về cảm quan có thể cho phép lựa chọn tần số giao chéo thấp hơn giữa LF và HF so với các phương pháp truyền thống [3].

Trong các sơ đồ BWE, việc tái tạo vùng phổ HF trên tần số được gọi là giao chéo đã cho thường dựa trên sự ráp nối phổ. Thông thường, vùng HF bao gồm nhiều bản vá được xếp chồng và mỗi bản vá trong số các bản vá này được lấy từ các vùng thông băng (band-pass - BP) của phổ LF dưới tần số giao chéo đã cho.

Các hệ thống theo tình trạng kỹ thuật thực hiện hiệu quả sự ráp nối trong giàn lọc hoặc sự biểu diễn biến đổi thời gian-tần số bằng cách sao chép tập hợp gồm các hệ số băng con liền kề từ phổ nguồn sang vùng phổ đích.

Trong bước tiếp theo, âm sắc, độ nhiễu âm và đường bao phổ được điều chỉnh sao cho gần giống với các đặc tính cảm quan và đường bao của tín hiệu HF ban đầu đã được đo trong bộ mã hóa và được truyền trong dòng bit như là thông tin phụ BWE.

Tái tạo băng phổ (Spectral Band Replication - SBR) là BWE được biết đến rộng rãi, được sử dụng trong các bộ mã hóa-giải mã âm thanh đương đại như mã hóa âm thanh tiên tiến hiệu quả cao (High Efficiency Advanced Audio Coding - HE-AAC) và sử dụng các kỹ thuật nêu trên [1].

Điền đầy khoảng trống thông minh (Intelligent Gap Filling - IGF) thể hiện kỹ thuật mã hóa bán tham số trong các bộ mã hóa-giải mã hiện đại như Âm thanh 3D MPEG-H hoặc bộ mã hóa-giải mã 3gpp EVS [2]. IGF có thể được áp dụng để điền đầy các lỗ phổ được đưa vào bởi quy trình lượng tử hóa trong bộ mã hóa do các ràng buộc về tốc độ bit thấp.

Thông thường, nếu kho bit hạn chế không cho phép mã hóa rõ ràng, các lỗ phổ xuất hiện trong vùng tần số cao (high-frequency - HF) của tín hiệu trước tiên và ngày càng ảnh hưởng đến toàn bộ dải phổ phía trên cho tốc độ bit thấp nhất.

Ở phía bộ giải mã, các lỗ phổ như vậy được thay thế thông qua IGF bằng cách sử dụng nội dung HF tổng hợp được tạo ra theo kiểu bán tham số từ nội dung tần số thấp (low-frequency - LF) và sự xử lý sau được điều khiển bởi thông tin phụ tham số bổ sung như sự điều chỉnh đường bao phổ và "mức độ làm trắng" phổ.

Tuy nhiên, sau sự xử lý sau nêu trên, vẫn có thể tồn tại sự không khớp mà có thể dẫn đến sự nhận biết thành phần lạ. Sự không khớp này thường bao gồm:

- Không khớp sóng hài: các thành phần lạ về nhịp điệu do các thành phần sóng hài đặt sai

- Không khớp pha: sự phân tán của các tín hiệu kích thích tương tự xung dẫn đến mất cảm nhận về độ ù trong tiếng nói hữu thanh hoặc các tín hiệu của kèn đồng

- Không khớp âm sắc: âm sắc bị phóng đại hoặc quá nhỏ

Do đó, các phương pháp hiệu chỉnh tần số và pha đã được đề xuất để hiệu chỉnh các dạng không khớp này [3] thông qua xử lý sau bổ sung. Theo sáng chế này, các tác giả sáng chế đề xuất tránh đưa những thành phần lạ này vào tín hiệu "thô", thay vì sửa chúng trong bước xử lý sau như được tìm thấy trong các phương pháp trong tình trạng kỹ thuật.

Các cách triển khai BWE khác dựa trên các kỹ thuật miền thời gian để ước lượng tín hiệu HF [4], điển hình là thông qua việc áp dụng hàm phi tuyến tính trên dạng sóng LF miền thời gian như nắn dòng, bình phương hoặc hàm lũy thừa. Theo cách này,

bằng cách làm biến dạng LF, hỗn hợp phong phú của các âm bội hòa hợp và không hòa hợp được tạo ra có thể được sử dụng như tín hiệu "thô" để khôi phục nội dung HF.

Sau đây, đặc biệt là sự không khớp sóng hài là vấn đề, vì về nội dung đa âm, các kỹ thuật này tạo ra hỗn hợp dày đặc của các âm bội sóng hài mong muốn mà chắc chắn được trộn với các thành phần không hài hòa không mong muốn.

Trong khi xử lý sau có thể dễ dàng làm tăng nhiễu âm, nó hoàn toàn thất bại trong việc loại bỏ các thành phần âm sắc không hài hòa không mong muốn khi chúng được đưa vào HF được ước lượng "thô".

Mục đích của sáng chế nhằm đề xuất giải pháp cải tiến để tạo ra tín hiệu âm thanh được nâng cao tần số từ tín hiệu âm thanh nguồn.

Mục đích này đạt được nhờ bộ xử lý âm thanh theo điểm 1, phương pháp xử lý tín hiệu âm thanh theo điểm 17 hoặc chương trình máy tính theo điểm 18.

Sáng chế dựa trên phát hiện rằng chất lượng cảm quan được cải thiện của sự mở rộng băng thông âm thanh hoặc điền đầy khoảng trống, hoặc nói chung, sự nâng cao tần số thu được nhờ sự tạo ra được thích ứng tín hiệu mới có nội dung bản vá "thô" tín hiệu tần số cao (high-frequency - HF) điền đầy khoảng trống hoặc được ước lượng. Bằng cách thu được tín hiệu "thô" được thích ứng về cảm quan, các ước số hiệu chỉnh sau cần thiết khác có thể được tối thiểu hóa hoặc thậm chí loại bỏ.

Các phương án của sáng chế được chỉ ra là kích thích xung được đồng bộ hóa đường bao dạng sóng (Waveform Envelope Synchronized Pulse Excitation - WESPE) được dựa trên sự tạo ra tín hiệu tương tự chuỗi xung trong miền thời gian, trong đó sự đặt xung thực tế được đồng bộ hóa với đường bao miền thời gian. Đường bao miền thời gian được suy ra từ tín hiệu tần số thấp (low-frequency - LF) có sẵn ở đầu ra của, ví dụ, bộ mã hóa lõi, hoặc có sẵn từ bất kỳ nguồn nào khác của tín hiệu âm thanh nguồn. Do đó, thu được tín hiệu "thô" được thích ứng về cảm quan.

Bộ xử lý âm thanh tương ứng với khía cạnh của sáng chế được tạo cấu hình để tạo ra âm thanh được nâng cao tần số từ tín hiệu âm thanh nguồn và bao gồm bộ xác định đường bao để xác định đường bao theo thời gian của ít nhất một phần tín hiệu âm thanh nguồn. Bộ phân tích được tạo cấu hình để phân tích đường bao theo thời gian nhằm xác định các giá trị có các đặc điểm nhất định của đường bao theo thời gian. Các giá trị này có thể là các giá trị hoặc năng lượng theo thời gian hoặc các giá trị khác liên

quan đến các đặc điểm. Bộ tổng hợp tín hiệu được đặt để tạo ra tín hiệu tổng hợp, trong đó sự tạo ra tín hiệu tổng hợp bao gồm đặt các xung liên quan đến các giá trị thời gian được xác định, trong đó các xung được gán trọng số sử dụng các trọng số suy ra từ các biên độ của đường bao theo thời gian, trong đó các biên độ liên quan đến các giá trị theo thời gian, tại đó các xung được đặt. Bộ tổ hợp ở đó để tổ hợp ít nhất băng của tín hiệu tổng hợp không nằm trong tín hiệu âm thanh nguồn và tín hiệu âm thanh nguồn để thu được tín hiệu âm thanh được nâng cao tần số.

Sáng chế có lợi thế ở chỗ, trái ngược với sự tạo ra có phần "mù" các tần số cao hơn từ tín hiệu âm thanh nguồn, ví dụ, bằng cách sử dụng phép xử lý phi tuyến tính hoặc tương tự, sáng chế đề xuất quy trình được điều khiển dễ dàng bằng cách xác định đường bao theo thời gian của tín hiệu nguồn và bằng cách đặt các xung tại các đặc điểm nhất định của đường bao theo thời gian như các cực đại cục bộ của đường bao theo thời gian hoặc các cực tiểu cục bộ của đường bao theo thời gian hoặc bằng cách đặt các xung luôn nằm giữa hai cực tiểu cục bộ của đường bao theo thời gian hoặc liên quan theo bất cứ cách nào khác với các đặc điểm nhất định của đường bao theo thời gian. Xung có nội dung tần số nói chung là phẳng trên toàn bộ dải tần số đang xem xét. Do đó, ngay cả khi các xung được sử dụng không phải là lý tưởng về mặt lý thuyết nhưng, ví dụ, gần được như thế, nội dung tần số của các xung không lý tưởng này, tức là, các xung không phù hợp với hình dạng Dirac lý tưởng, tuy nhiên tương đối phẳng trong dải tần số quan tâm, ví dụ, trong khoảng từ 0 đến 20kHz trong ngữ cảnh của việc điền đầy khoảng trống thông minh (intelligent gap filling - IGF) hoặc trong dải tần số, ví dụ, trong khoảng từ 8kHz đến 16 hoặc 20kHz trong ngữ cảnh mở rộng băng thông âm thanh, tại đó tín hiệu nguồn bị giới hạn băng thông.

Do đó, tín hiệu tổng hợp bao gồm các xung này cung cấp nội dung tần số cao dày đặc và được điều khiển dễ dàng. Bằng cách đặt một số xung trên mỗi đường bao theo thời gian, ví dụ, được trích xuất từ khung của tín hiệu âm thanh nguồn, thu được sự định hình trong miền phổ, vì các nội dung tần số của các xung khác nhau được đặt viện dẫn đến các đặc điểm nhất định chồng lên nhau trong miền phổ để phù hợp với ít nhất các đặc điểm nổi trội hoặc, nói chung là, các đặc điểm nhất định của đường bao theo thời gian của tín hiệu âm thanh nguồn. Do thực tế là các pha của các giá trị phổ được thể hiện bởi các xung bị khóa với nhau và do thực tế là, tốt hơn là các xung

dương hoặc các xung âm được đặt bởi bộ tổng hợp tín hiệu, các pha của các giá trị phổ được thể hiện bởi xung riêng lẻ trong số các xung khác nhau được khóa với nhau. Do đó, thu được tín hiệu tổng hợp được điều khiển có các đặc trưng miền tần số rất hữu ích. Thông thường, tín hiệu tổng hợp là tín hiệu băng rộng mở rộng trên toàn bộ dải tần số âm thanh hiện có, tức là, cũng mở rộng sang dải LF. Để thực sự tạo ra tín hiệu cuối cùng mà sau cùng được tổ hợp với tín hiệu âm thanh nguồn nhằm mục đích nâng cao tần số, ít nhất băng của tín hiệu tổng hợp, như băng cao hoặc tín hiệu được xác định bởi thông băng được trích xuất và được thêm vào tín hiệu âm thanh nguồn.

Giải pháp sáng chế có tiềm năng được thực hiện đầy đủ trong miền thời gian, tức là, không có bất kỳ biến đổi cụ thể nào. Miền thời gian là miền thời gian điển hình hoặc miền thời gian được lọc mã hóa dự báo tuyến tính (linear prediction coding - LPC), tức là tín hiệu miền thời gian đã được làm trắng phổ và cuối cùng cần được xử lý bằng bộ lọc tổng hợp LPC để đưa lại hình dạng phổ ban đầu để hữu ích cho việc kết xuất tín hiệu âm thanh. Do đó, sự xác định đường bao, phân tích, tổng hợp tín hiệu, trích xuất băng tín hiệu tổng hợp và tổ hợp cuối cùng đều có thể được thực hiện trong miền thời gian để có thể tránh được bất kỳ biến đổi phổ-thời gian hay biến đổi thời gian-phổ xảy ra độ trễ điển hình. Tuy nhiên, ngữ cảnh theo sáng chế cũng linh hoạt ở chỗ một số quy trình như xác định đường bao, tổng hợp tín hiệu và tổ hợp cũng có thể được thực hiện một phần hoặc toàn bộ trong miền phổ. Do đó, việc triển khai sáng chế, tức là, cho dù các quy trình nhất định theo yêu cầu của sáng chế được triển khai trong miền thời gian hoặc miền phổ, luôn có thể được thích ứng toàn bộ với khuôn khổ tương ứng của thiết kế bộ giải mã điển hình cần thiết cho ứng dụng nhất định. Ngữ cảnh sáng chế thậm chí còn linh hoạt trong ngữ cảnh của bộ mã hóa tiếng nói LPC, trong đó, ví dụ, sự nâng cao tần số của tín hiệu kích thích LPC (ví dụ: tín hiệu kích thích được mã hóa biến đổi (Transform Coded eXcitation - TCX) được thực hiện. Sự tổ hợp tín hiệu tổng hợp và tín hiệu âm thanh nguồn được thực hiện trong miền thời gian LPC và sự chuyển đổi cuối cùng từ miền thời gian LPC sang miền thời gian bình thường được thực hiện với bộ lọc tổng hợp LPC, trong đó, cụ thể, sự điều chỉnh đường bao được ưu tiên điển hình của tín hiệu tổng hợp được thực hiện trong tầng lọc tổng hợp LPC cho phần phổ tương ứng được biểu diễn bằng ít nhất băng của tín hiệu tổng hợp. Do đó, các thuật toán xử lý sau cần thiết điển hình được tổ hợp với sự điều chỉnh

đường bao trong một tầng lọc. Các thuật toán xử lý sau này có thể liên quan đến việc lọc tổng hợp LPC, lọc chỉnh giảm được biết từ bộ giải mã tiếng nói, các thuật toán lọc sau khác như các thuật toán lọc sau âm trầm hoặc các quy trình lọc sau nâng cao âm thanh khác dựa trên dự báo dài hạn (Long Term Prediction - LTP) như được tìm thấy trong các bộ giải mã TCX hoặc các bộ giải mã khác.

Mô tả vắn tắt các hình vẽ

Các phương án ưu tiên của sáng chế được thảo luận sau đây viện dẫn đến các hình vẽ kèm theo, trong đó:

Fig.1 là sơ đồ khối của phương án về bộ xử lý âm thanh theo sáng chế;

Fig.2 là sự mô tả chi tiết hơn cho phương án ưu tiên về bộ xác định đường bao của Fig.1;

Fig.3a minh họa phương án cho việc tính toán đường bao theo thời gian của tín hiệu âm thanh băng con hoặc băng đầy đủ;

Fig.3b minh họa cách triển khai khác cho sự tạo ra đường bao theo thời gian;

Fig.3c minh họa lưu đồ cho cách triển khai sự xác định tín hiệu phân tích ở Fig.3a sử dụng biến đổi Hilbert;

Fig.4 minh họa cách triển khai ưu tiên cho bộ phân tích của Fig.1;

Fig.5 minh họa cách triển khai ưu tiên cho bộ tổng hợp của Fig.1;

Fig.6 minh họa phương án ưu tiên cho bộ xử lý âm thanh như thiết bị hoặc phương pháp được sử dụng trong ngữ cảnh bộ giải mã lõi;

Fig.7 minh họa cách triển khai ưu tiên, trong đó sự tổ hợp tín hiệu tổng hợp và tín hiệu âm thanh nguồn được thực hiện trong miền LPC;

Fig.8 minh họa phương án khác của sáng chế, trong đó bộ lọc thông cao hoặc thông băng, sự điều chỉnh đường bao và sự tổ hợp tín hiệu âm thanh nguồn và tín hiệu tổng hợp được thực hiện trong miền phổ;

Fig.9a minh họa một vài tín hiệu trong quá trình nâng cao tần số đối với mục âm thanh "tiếng nói nam giới người Đức";

Fig.9b minh họa biểu đồ phổ cho mục âm thanh "tiếng nói nam giới người Đức";

Fig.10a minh họa một vài tín hiệu trong quá trình nâng cao tần số đối với mục âm "ống chỉnh tần số cơ bản";

Fig.10a minh họa biểu đồ phổ cho mục âm "ống chỉnh tần số cơ bản";

Fig.11a minh họa một vài tín hiệu trong quá trình nâng cao tần số đối với mục âm thanh "Madonna Vogue"; và

Fig.11b minh họa một vài tín hiệu trong quá trình nâng cao tần số đối với mục âm thanh "Madonna Vogue".

Mô tả chi tiết sáng chế

Fig.1 minh họa bộ xử lý âm thanh để tạo ra tín hiệu âm thanh được nâng cao tần số 420 tại đầu ra của bộ tổ hợp 400 từ tín hiệu âm thanh nguồn một mặt được nhập vào bộ xác định đường bao 100 và mặt khác được nhập vào bộ tổ hợp 400.

Bộ xác định đường bao 100 được tạo cấu hình để xác định đường bao theo thời gian của ít nhất một phần của tín hiệu âm thanh nguồn. Bộ xác định đường bao có thể sử dụng tín hiệu âm thanh nguồn băng đầy đủ hoặc, ví dụ, chỉ băng hoặc một phần của tín hiệu âm thanh nguồn có tần số biên thấp hơn nhất định, chẳng hạn như tần số, ví dụ, 100, 200 hoặc 500Hz. Đường bao theo thời gian được chuyển tiếp từ bộ xác định đường bao 100 đến bộ phân tích 200 để phân tích đường bao theo thời gian nhằm xác định các giá trị có các đặc điểm nhất định của đường bao theo thời gian. Các giá trị này có thể là giá trị hoặc năng lượng theo thời gian hoặc các giá trị khác liên quan đến các đặc điểm. Các đặc điểm nhất định có thể, ví dụ, là các cực đại cục bộ của đường bao theo thời gian, các cực tiểu cục bộ của đường bao theo thời gian, các điểm về 0 của đường bao theo thời gian hoặc các điểm nằm giữa hai cực tiểu cục bộ hoặc các điểm nằm giữa hai cực đại cục bộ mà trong đó, ví dụ, các điểm giữa các đặc điểm đó là các giá trị có cùng khoảng cách theo thời gian với các đặc điểm lân cận. Do đó, các đặc điểm nhất định này cũng có thể là các điểm nằm nửa chừng giữa hai cực tiểu cục bộ hoặc hai cực đại cục bộ. Tuy nhiên, trong các phương án ưu tiên, sự xác định các cực đại cục bộ của đường bao theo thời gian sử dụng, ví dụ, xử lý tính toán đường cong được ưu tiên hơn. Các giá trị theo thời gian có các đặc điểm nhất định của đường bao theo thời gian được chuyển tiếp đến bộ tổng hợp tín hiệu 300 để tạo ra tín hiệu tổng hợp. Sự tạo ra tín hiệu tổng hợp bao gồm việc đặt các xung liên quan đến các đường bao theo thời gian được xác định, trong đó các xung được gán trọng số, trước khi đặt hoặc sau khi được đặt bằng cách sử dụng các trọng số được suy ra từ các biên độ của đường bao theo thời gian, các biên độ liên quan đến các giá trị theo thời gian

nhận được từ bộ phân tích hoặc liên quan đến các giá trị theo thời gian, nơi các xung được đặt.

Ít nhất băng của tín hiệu tổng hợp hoặc băng cao đầy đủ của tín hiệu tổng hợp hoặc một số băng riêng lẻ và riêng biệt của tín hiệu tổng hợp hoặc thậm chí toàn bộ tín hiệu tổng hợp được chuyển tiếp đến bộ tổ hợp 400 để tổ hợp ít nhất một băng của tín hiệu tổng hợp không nằm trong tín hiệu âm thanh nguồn và tín hiệu âm thanh nguồn để thu được tín hiệu âm thanh được nâng cao tần số.

Trong phương án ưu tiên, bộ xác định đường bao được tạo cấu hình như được minh họa trên Fig.2. Trong phương án này, tín hiệu âm thanh nguồn hoặc ít nhất một phần của tín hiệu âm thanh nguồn được phân tách thành nhiều tín hiệu băng con như được minh họa ở 105. Một hoặc nhiều hoặc thậm chí tất cả các băng con được chọn hoặc sử dụng như minh họa ở 110 để xác định các đường bao theo thời gian riêng lẻ cho mỗi băng con (được chọn) như minh họa ở 120. Như minh họa ở 125, các đường bao theo thời gian được chuẩn hóa hoặc lọc và, các đường bao theo thời gian riêng lẻ được tổ hợp với nhau như được chỉ ra ở 130 để thu được đường bao theo thời gian cuối cùng ở đầu ra của bộ xác định đường bao. Đường bao theo thời gian cuối cùng này có thể là đường bao được tổ hợp như được xác định bằng quy trình được minh họa trên Fig.2. Tùy thuộc vào cách triển khai, tầng lọc bổ sung 115 có thể được thực hiện để chuẩn hóa hoặc lọc các băng con được chọn riêng lẻ. Nếu tất cả các băng con được sử dụng, tất cả các băng con đó được chuẩn hóa hoặc lọc như được chỉ ra trong khối 115. Quy trình chuẩn hóa được chỉ ra ở 125 có thể được bỏ qua và quy trình không thực hiện chuẩn hóa hoặc lọc các đường bao theo thời gian được xác định này là hữu ích, khi các băng con, từ đó mà các đường bao theo thời gian được xác định trong khối 120, đã được chuẩn hóa hoặc lọc tương ứng. Tất nhiên, cả hai quy trình 115, 125 có thể cũng được thực hiện hoặc, thậm chí là theo cách khác, chỉ các quy trình xác định đường bao theo thời gian cho mỗi băng con (được chọn) 120 và sự tổ hợp sau đó của các đường bao theo thời gian 130 có thể được thực hiện mà không cần bất kỳ quy trình nào được minh họa bằng khối 115 hoặc 125.

Trong cách triển khai khác, quá trình phân tách trong khối 105 hoàn toàn không thể được thực hiện, nhưng có thể được thay thế bằng việc lọc thông cao với tần số giao chéo thấp như tần số giao chéo 20, 50, 100 hoặc, ví dụ tần số dưới 500Hz và chỉ một

đường bao theo thời gian được xác định từ kết quả của sự lọc thông cao này. Tất nhiên, cũng có thể tránh được việc lọc thông cao và chỉ có một đường bao theo thời gian từ tín hiệu âm thanh nguồn và điển hình, khung của tín hiệu âm thanh nguồn được suy ra, trong đó tín hiệu âm thanh nguồn được ưu tiên xử lý trong các khung thường chồng lấp, nhưng thậm chí các khung không chồng lấp cũng có thể được sử dụng. Sự lựa chọn được chỉ ra trong khối 110 được triển khai trong ngữ cảnh nhất định, ví dụ, khi người ta xác định rằng tín hiệu băng con nhất định không đáp ứng các tiêu chí cụ thể liên quan đến các đặc điểm của tín hiệu băng con hoặc bị loại khỏi việc xác định đường bao theo thời gian cuối cùng vì bất kỳ lý do gì.

Fig.5 minh họa cách triển khai ưu tiên cho bộ tổng hợp tín hiệu 300. Bộ tổng hợp tín hiệu 300 nhận, như là đầu vào từ bộ phân tích 200, các giá trị theo thời gian của các dấu hiệu và, thêm vào đó, các thông tin khác về đường bao. Trong mục 310, bộ tổng hợp tín hiệu 300 được minh họa trên Fig.5 suy ra các hệ số định tỷ lệ từ đường bao theo thời gian liên quan đến các giá trị theo thời gian. Do đó, khối 310 nhận thông tin về đường bao như một mặt là các biên độ đường bao, mặt khác là các giá trị theo thời gian. Việc suy ra các hệ số định tỷ lệ được thực hiện, ví dụ, sử dụng hàm nén như hàm căn bậc hai, hàm lũy thừa với lũy thừa nhỏ hơn 1,0 hoặc, ví dụ, hàm logarit.

Bộ tổng hợp tín hiệu 300 bao gồm quy trình đặt 305 các xung vào các giá trị theo thời gian mà, tốt hơn là, chỉ các xung âm hoặc chỉ các xung dương được đặt để có các pha được đồng bộ hóa của các giá trị phổ liên quan được kết hợp với xung. Tuy nhiên, trong các phương án khác, và tùy thuộc vào, ví dụ, các tiêu chí khác được suy ra từ các thông tin phụ điển hình có sẵn về mở rộng băng thông hoặc điền đầy khoảng trống, sự đặt ngẫu nhiên các xung được thực hiện, điển hình là, khi âm sắc của tín hiệu băng gốc không quá cao. Việc đặt các xung âm hoặc dương có thể được kiểm soát bởi cực của dạng sóng ban đầu. Cực của các xung có thể được chọn sao cho bằng với cực của dạng sóng ban đầu có hệ số đỉnh cao nhất. Nói cách khác, điều này có nghĩa là các đỉnh dương được tạo mô hình bởi các xung dương và ngược lại.

Trong bước 315, các xung thu được bởi khối 305 được định tỷ lệ bằng cách sử dụng kết quả của khối 310 và, sự xử lý sau 320 tùy chọn được thực hiện cho các xung. Tín hiệu xung có sẵn và tín hiệu xung được lọc thông cao hoặc lọc thông băng như minh họa ở khối 325 để thu được băng tần số của tín hiệu xung, tức là, để thu được ít

nhất băng của tín hiệu tổng hợp mà được chuyển tiếp đến bộ tổ hợp. Tuy nhiên, sự điều chỉnh đường bao phổ tùy chọn 330 được áp dụng cho đầu ra tín hiệu bởi tầng lọc 325, trong đó sự điều chỉnh đường bao phổ này được thực hiện bằng cách sử dụng hàm đường bao nhất định hoặc sự lựa chọn nhất định các tham số đường bao như được suy ra từ thông tin phụ, hoặc cách khác, được suy ra từ tín hiệu âm thanh nguồn trong ngữ cảnh, ví dụ, các ứng dụng mở rộng băng thông mù.

Fig.6 minh họa phương án ưu tiên của bộ xử lý âm thanh hoặc phương pháp xử lý âm thanh để tạo ra tín hiệu âm thanh được nâng cao tần số. Cách tiếp cận sáng chế được gọi là Kích thích xung đồng bộ hóa đường bao dạng sóng (Waveform Envelope Synchronized Pulse Excitation - WESPE) dựa trên việc tạo ra tín hiệu tương tự chuỗi xung, trong đó vị trí xung thực tế được đồng bộ hóa với đường bao miền thời gian chuyên dụng. Cái gọi là đường bao chung này được suy ra từ tín hiệu LF thu được ở đầu ra của bộ giải mã lõi 20, thông qua tập hợp các tín hiệu thông băng, mà có các đường bao riêng lẻ được tổ hợp thành một đường bao chung.

Fig.6 cho thấy sự hợp nhất điển hình của xử lý WESPE vào bộ giải mã âm thanh có chức năng mở rộng băng thông (bandwidth extension - BWE), đây cũng là phương án ưu tiên của sáng chế. Cách triển khai này vận hành trên các khung thời gian với thời lượng ví dụ, 20ms, tùy chọn với sự chồng lấp theo thời gian giữa các khung hình ở mức, ví dụ, 50%.

Ưu điểm của WESPE BWE mới được đề xuất

· Giảm thiểu độ thô và các thành phần lạ về nhịp điệu

· Sự liên tục sóng hài của các tín hiệu

· Bảo tồn các xung

· Phù hợp để làm giọng nói BWE

· Có thể điều khiển âm nhạc tốt

· Sự giao chéo BWE có thể bắt đầu sẵn ở mức 2kHz hoặc thấp hơn

· BWE tự điều chỉnh về âm sắc, căn chỉnh tần số cơ bản, hòa âm, pha

Sự xử lý WESPE bao gồm các bước sau:

1. Sự ước lượng đường bao theo thời gian 100: Tín hiệu LF thu được từ bộ giải mã lõi 20 được chia tách (105) thành tập hợp các tín hiệu thông băng. Tiếp theo, đường bao theo thời gian được xác định 120 cho mỗi tín hiệu thông băng. Tùy chọn,

có thể áp dụng sự chuẩn hóa hoặc lọc các đường bao riêng lẻ. Sau đó, tất cả các đường bao theo thời gian được tổ hợp 130 thành đường bao chung. Tốt hơn là, thuật toán tổ hợp là quy trình tính trung bình.

2. Sự đặt xung được đồng bộ hóa: Đường bao chung được suy ra từ bước 1 được phân tích 205 bởi việc áp dụng tính toán đường cong, tốt hơn là cho vị trí của các cực đại lân cận của nó. Các ứng cử cực đại thu được có thể được lựa chọn sau hoặc được ổn định tùy chọn xét theo 210 khoảng cách theo thời gian của chúng. Xung Dirac được đặt 305 vào tín hiệu "thô" được ước lượng cho sự tạo ra HF tại mỗi vị trí cực đại. Tùy chọn, quy trình này có thể được hỗ trợ bởi thông tin phụ.

3. Định tỷ lệ độ lớn xung riêng lẻ được suy ra từ đường bao: Chuỗi xung được hợp dịch trong bước 2 trước đó được gán trọng số 315 bởi các trọng số theo thời gian được suy ra từ đường bao chung.

4. Sự xử lý sau, trích xuất HF hoặc lựa chọn điền đầy khoảng trống: Tín hiệu "thô" được tạo ra ở bước 3 được xử lý sau tùy chọn 320, ví dụ, bằng cách thêm nhiễu âm, và được lọc 325 để sử dụng như HF trong BWE hoặc như tín hiệu ô đích điền đầy khoảng trống.

5. Sự điều chỉnh năng lượng: Sự phân bố năng lượng phổ của tín hiệu được lọc từ bước 4 được điều chỉnh 330 để sử dụng như HF trong BWE hoặc làm tín hiệu ô đích điền đầy khoảng trống. Tại đây, thông tin phụ 40 từ dòng bit về sự phân bố năng lượng mong muốn được sử dụng.

6. Sự trộn HF hoặc tín hiệu điền đầy khoảng trống với LF: Cuối cùng, tín hiệu được điều chỉnh từ bước 5 được trộn 400 với đầu ra bộ mã hóa lõi 30 tương ứng với các nguyên lý BWE hoặc điền đầy khoảng trống thông thường, tức là đi qua bộ lọc HP và bổ sung cho LF hoặc điền đầy các lỗ phổ trong các vùng phổ điền đầy khoảng trống.

Sau đây, hàm của từng bước nằm trong sự xử lý WESPE sẽ được giải thích thêm khi đưa ra các tín hiệu ví dụ và ảnh hưởng của chúng đến kết quả xử lý.

Sự ước lượng đường bao chung theo thời gian thích hợp là phần quan trọng của WESPE. Đường bao chung cho phép sự ước lượng các đặc tính cảm quan được tính trung bình và do đó là tiêu biểu của từng khung thời gian riêng lẻ.

Nếu tín hiệu LF rất có âm sắc với tần số cơ bản f0 và phổ vạch âm bội được đặt cách Δf0 mạnh, một số vạch sẽ xuất hiện trong mỗi tín hiệu thông băng riêng lẻ, nếu chiều rộng băng thông của chúng có thể chứa chúng, tạo ra các biến điệu đường bao nhất quán mạnh thông qua các nhịp điệu trong tất cả các băng thông băng. Việc tính trung bình các đường bao theo thời gian sẽ bảo toàn cấu trúc biên điệu đường bao nhất quán này được tìm thấy trên các đường bao thông băng và sẽ dẫn đến các đỉnh mạnh tại các vị trí gần như cách đều nhau $\Delta T0 = 1/(\Delta f0)$. Sau đó, thông qua áp dụng tính toán đường cong, các xung mạnh sẽ được đặt tại các vị trí đỉnh này, tạo thành chuỗi xung có phổ bao gồm các vạch cách đều rời rạc tại các vị trí $n*\Delta f0$, n=1…N.

Trong trường hợp tín hiệu có âm sắc mạnh hoặc không có âm bội nào hoặc băng thông của các bộ lọc thông băng không thể chứa nhiều hơn một trong các âm bội này trong mỗi băng riêng lẻ, cấu trúc biến điệu sẽ không xuất hiện trong tất cả các tín hiệu thông băng và kết quả là không trội hơn đường bao chung được tính trung bình. Sự đặt xung thu được sẽ dựa trên hầu hết các cực đại có khoảng cách không đều và do đó bị nhiễu.

Điều này cũng đúng đối với các tín hiệu LF nhiễu thể hiện sự đặt các cực đại cục bộ ngẫu nhiên trong tín hiệu đường bao chung: các tín hiệu này dẫn đến sự đặt xung giả ngẫu nhiên.

Các sự kiện nhất thời được bảo toàn vì trong trường hợp này, tất cả các tín hiệu thông băng đều chia sẻ cực đại chung được căn chỉnh theo thời gian, do đó cũng sẽ xuất hiện trong đường bao chung.

Các băng thông nên được định kích thước sao cho chúng trải dài trên băng cảm quan và chúng có thể điều tiết ít nhất 2 âm bội cho tần số cao nhất cần phải được phân giải. Để tính trung bình tốt hơn, các băng thông có thể có sự chồng lấp trong các băng chuyển vị của chúng. Theo cách này, âm sắc của tín hiệu được ước lượng về bản chất là thích ứng với tín hiệu LF. Các thông băng có thể loại trừ các tần số rất thấp ở mức dưới, ví dụ, 20Hz.

Sự đặt và định tỷ lệ xung theo thời gian được đồng bộ hóa là đóng góp quan trọng khác của WESPE. Sự đặt xung được đồng bộ hóa kế thừa các đặc tính cảm quan tiêu biểu được cô đọng trong các biên điệu theo thời gian của đường bao chung và in dấu chúng vào tín hiệu băng đầy đủ thô được thích ứng về mặt cảm quan.

Lưu ý rằng cảm quan của con người về nội dung tần số cao được biết là hoạt động thông qua sự đánh giá các biên điệu trong các đường bao băng mang tính quyết định. Như đã được trình bày chi tiết trên đây, sự đặt xung theo thời gian được đồng bộ hóa với đường bao LF chung thực thi sự tương đồng và căn chỉnh của cấu trúc phổ và theo thời gian có liên quan về mặt cảm quan giữa LF và HF.

Trong trường hợp các tín hiệu rất có âm sắc với các âm bội mạnh và sạch, ví dụ như ống chỉnh tần số cơ bản, WESPE có thể đảm bảo thông qua sự ổn định tùy chọn bổ sung rằng sự đặt xung là chính xác cách đều nhau dẫn đến phổ âm bội HF rất có âm sắc của tín hiệu "thô".

Việc gán trọng số các xung với đường bao chung đảm bảo rằng các biên điệu trội hơn được bảo toàn trong các xung mạnh trong khi các biên điệu ít quan trọng hơn dẫn đến các xung yếu, đóng góp thêm vào đặc tính WESPE về sự thích ứng nội tại của tín hiệu "thô" với tín hiệu LF.

Trong trường hợp tín hiệu nhiễu, nếu sự đặt và gán trọng số cho xung ngày càng trở nên ngẫu nhiên, điều này dẫn đến tín hiệu "thô" dần dần nhiễu hơn, đây là đặc tính rất được mong muốn.

Các bước xử lý còn lại, trích xuất HF, điều chỉnh và trộn năng lượng là các bước tiếp theo cần thiết để tích hợp sự xử lý WESPE mới vào bộ mã hóa-giải mã để phù hợp với chức năng đầy đủ của BWE hoặc điền đầy khoảng trống.

Fig.3a minh họa cách triển khai ưu tiên để xác định đường bao theo thời gian. Như minh họa ở 135, tín hiệu phân tích được xác định bằng cách sử dụng biến đổi Hilbert. Đầu ra của khối 135, tức là, tín hiệu biến đổi Hilbert được sử dụng để tính toán đường bao ENV(t) được minh họa ở 140. Để đạt được điều này, đường bao được tính bằng cách tính bình phương các giá trị theo thời gian của tín hiệu âm thanh nguồn ban đầu tại các thời điểm nhất định và bình phương của các giá trị biến đổi Hilbert tương ứng tại các thời điểm nhất định đó và để cộng các giá trị bình phương và tính toán căn bậc hai từ kết quả của phép cộng đó cho mỗi thời điểm riêng lẻ. Bằng quy trình này, đường bao theo thời gian được xác định với cùng độ phân giải mẫu với tín hiệu âm thanh nguồn ban đầu a(t). Tất nhiên, quy trình như vậy được thực hiện khi đầu vào vào khối 135 và 140 là tín hiệu băng con như thu được bởi khối 105, hoặc như được chọn bởi khối 110 hoặc như được chuẩn hóa và lọc bởi khối 115 ở Fig.2.

Quy trình khác để tính toán đường bao theo thời gian được minh họa trong khối 145 và 150 ở Fig.3b. Để đạt được điều này, dạng sóng của tín hiệu âm thanh nguồn hoặc băng con từ âm thanh nguồn được nắn dòng (145) và tín hiệu được nắn dòng được lọc thông thấp (150), và kết quả của việc lọc thông thấp này là đường bao của tín hiệu âm thanh nguồn hoặc là đường bao của tín hiệu băng con riêng lẻ được tổ hợp với các đường bao khác như vậy từ các băng con khác, tốt hơn là bằng cách tính trung bình như được minh họa ở 130 trên Fig.2.

Công bố "Simple empirical algorithm to obtain signal envelope in the three steps", bởi C. Jarne, ngày 20 tháng 3 năm 2017 minh họa các quy trình khác để tính toán các đường bao theo thời gian, chẳng hạn như tính toán giá trị căn bậc hai trung bình (root mean square - RMS) tức thời của dạng sóng thông qua cửa sổ trượt với sự hỗ trợ hữu hạn. Các quy trình khác bao gồm tính toán làm gần đúng tuyến tính theo mảnh của dạng sóng trong đó đường bao biên độ được tạo ra bằng cách tìm và kết nối các đỉnh của dạng sóng trong cửa sổ di chuyển qua các dữ liệu. Các quy trình khác dựa vào sự xác định các đỉnh không đổi trong tín hiệu âm thanh nguồn hoặc tín hiệu băng con và sự suy ra đường bao bằng phép nội suy.

Các quy trình khác để tính toán đường bao theo thời gian bao gồm diễn giải thông tin phụ thể hiện đường bao hoặc thực hiện dự báo trong miền phổ trên tập hợp gồm các giá trị phổ được suy ra từ khung miền thời gian như được biết từ định hình nhiễu âm theo thời gian (Temporal Noise Shaping – TNS), trong đó các hệ số dự báo tương ứng thể hiện đường bao theo thời gian của khung.

Fig.3c minh họa cách triển khai ưu tiên để xác định tín hiệu phân tích bằng cách sử dụng biến đổi Hilbert được chỉ ra ở 135 trên Fig.3a. Ví dụ, các quy trình để tính toán các biến đổi Hilbert này được minh họa trong "A Praat-Based Algorithm to Extract the Amplitude Envelope and Temporal Fine Structure Using the Hilbert Transform", He, Lei, et al, INTERSPEECH 2016-1447, trang 530-534 . Trong bước 136, phổ phức được tính toán từ tín hiệu a(t), ví dụ, tín hiệu âm thanh nguồn hoặc tín hiệu băng con. Trong bước 137, phần dương của phổ phức được chọn hoặc phần âm được bỏ chọn. Trong bước 138, phần dương của phổ phức được nhân với "- j" và, ở bước 139, kết quả của phép nhân này được chuyển đổi thành miền thời gian và bằng cách lấy đi phần ảo, thu được tín hiệu phân tích $\hat{a}(t)$.

Tất nhiên, nhiều quy trình khác để xác định đường bao theo thời gian cũng có sẵn và cần lưu ý rằng đường bao theo thời gian không nhất thiết phải thực sự "bao" tín hiệu miền thời gian, nhưng tất nhiên, có thể là một số cực đại hoặc cực tiểu của tín hiệu miền thời gian cao hơn hoặc nhỏ hơn giá trị đường bao tương ứng tại thời điểm này.

Fig.4 minh họa phương án ưu tiên của quy trình xác định các giá trị theo thời gian của các đặc điểm nhất định của đường bao theo thời gian. Để đạt được điều này, đường bao theo thời gian trung bình được đưa vào khối 205 để xác định các giá trị theo thời gian ban đầu cho các đặc điểm. Ví dụ, những giá trị theo thời gian ban đầu này có thể là các giá trị theo thời gian của các cực đại thực sự được tìm thấy trong đường bao theo thời gian. Các giá trị theo thời gian cuối cùng của các đặc điểm, mà tại đó các xung thực sự được đặt, được suy ra từ các giá trị theo thời gian thô hoặc từ các giá trị theo thời gian "ban đầu" bằng hàm tối ưu hóa, bằng thông tin phụ hoặc bằng cách chọn hoặc điều khiển các đặc điểm thô như được chỉ ra bởi khối 210. Tốt hơn, khối 210 là sao cho các giá trị ban đầu được điều khiển tương ứng với quy tắc xử lý hoặc tương ứng với hàm tối ưu hóa. Cụ thể, hàm tối ưu hóa hoặc quy tắc xử lý được triển khai để các giá trị theo thời gian được đặt trong đường quét với khoảng cách đường quét T. Cụ thể, khoảng cách đường quét T và/hoặc vị trí của đường quét trong đường bao theo thời gian là sao cho giá trị độ lệch giữa các giá trị theo thời gian và các giá trị theo thời gian ban đầu có đặc trưng được định trước, trong đó, theo các phương án, giá trị độ lệch là tổng các chênh lệch bình phương và/hoặc đặc trưng được định trước là cực tiểu. Do đó, sau khi xác định các giá trị theo thời gian ban đầu, đường quét của các giá trị theo thời gian cách đều được đặt khớp với đường quét không cố định của các giá trị theo thời gian ban đầu càng gần càng tốt, nhưng giờ đây cho thấy dáng điệu âm sắc rõ ràng và lý tưởng. Đường quét có thể được xác định trong miền được lấy mẫu ngược có độ chi tiết theo thời gian tốt hơn so với miền không được lấy mẫu ngược, hoặc người ta có thể sử dụng độ trễ phân đoạn để đặt xung với độ chính xác của mẫu phụ.

Fig.7 minh họa phương án khác của sáng chế trong ngữ cảnh xử lý LPC. Như được minh họa, ví dụ, trên Fig.1 hoặc Fig.6, bộ xử lý âm thanh trên Fig.7 bao gồm bộ xác định đường bao 100, bộ phân tích 200 (cả hai không được hiển thị trên Fig.7) và

bộ tổng hợp tín hiệu 300. Tuy nhiên, trái ngược với Fig.6, dữ liệu đầu ra của bộ giải mã lõi, tức là, đầu ra LF 30 không phải là tín hiệu âm thanh miền thời gian, mà là tín hiệu âm thanh nằm trong miền thời gian LPC. Điển hình, dữ liệu này có thể được tìm thấy trong bộ mã hóa kích thích được mã hóa biến đổi (Transform Coded eXcitation - TCX) như sự biểu diễn tín hiệu nội bộ.

Các dữ liệu TCX được tạo ra bởi bộ giải mã âm thanh 20 trên Fig.7 được chuyển tiếp đến bộ trộn được chỉ ra trên Fig.7 dưới dạng bộ cộng miền LPC 405. Bộ tổng hợp tín hiệu tạo ra các dữ liệu nâng cao tần số TCX. Do đó, tín hiệu tổng hợp được tạo ra bởi bộ tổng hợp tín hiệu được suy ra từ tín hiệu âm thanh nguồn là tín hiệu dữ liệu TCX theo phương án này. Do đó, ở đầu ra của khối 405, tín hiệu âm thanh được nâng cao tần số có sẵn, tuy nhiên, vẫn nằm trong miền thời gian LPC. Bộ lọc tổng hợp LPC 410 được kết nối sau đó thực hiện biến đổi tín hiệu miền thời gian LPC thành miền thời gian.

Bộ lọc tổng hợp LPC được tạo cấu hình để thực hiện bổ sung loại chỉnh giảm nếu cần thiết và ngoài ra, bộ lọc miền thời gian này cũng được tạo cấu hình để thực hiện điều chỉnh đường bao phổ cho băng tín hiệu tổng hợp. Do đó, bộ lọc tổng hợp LPC 410 trên Fig.7 không chỉ thực hiện lọc tổng hợp cho đầu ra dải tần số dữ liệu TCX bởi bộ giải mã âm thanh 20 mà còn thực hiện điều chỉnh đường bao phổ cho các dữ liệu trong băng phổ không có trong đầu ra dữ liệu TCX bởi bộ giải mã âm thanh 20. Điển hình, dữ liệu này cũng thu được từ tín hiệu âm thanh được mã hóa 10 nhờ bộ giải mã âm thanh 20 trích xuất dữ liệu LPC 40a cho dải tần số lõi và trích xuất bổ sung sự điều chỉnh đường bao phổ cho băng cao hoặc, cho IGF (Intelligent Gap Filling – điền đầy khoảng trống thông minh), một hoặc nhiều băng được chỉ ra ở 40b của Fig.7. Do đó, bộ tổ hợp hoặc bộ trộn trên Fig.1 được triển khai bởi bộ cộng miền LPC 405 và bộ lọc tổng hợp LPC 410 được kết nối sau đó ở Fig.7 để đầu ra của bộ lọc tổng hợp LPC 410 được chỉ ra ở 420 là tín hiệu âm thanh miền thời gian được nâng cao tần số. Trái ngược với quy trình trên Fig.6, trong đó sự điều chỉnh đường bao phổ 330 được thực hiện trước khi thực hiện thuật toán trộn bởi bộ tổ hợp 400, Fig.7 thực hiện điều chỉnh đường bao của băng cao hoặc băng điền đầy sau khi trộn hoặc tổ hợp cả hai tín hiệu.

Fig.8 minh họa cách triển khai khác quy trình được minh họa trên Fig.6. Về cơ bản, cách triển khai ở Fig.6 được thực hiện trong miền thời gian để các khối 320, 325,

330, 400 được thực hiện đầy đủ trong miền thời gian. Hoặc, cách triển khai Fig.8 phụ thuộc vào chuyển đổi phổ 105 cho băng thấp, tuy nhiên, đây là ước số tùy chọn, nhưng thuật toán chuyển đổi phổ 105 trên Fig.8 cho băng thấp được sử dụng một cách có lợi thế để triển khai giàn lọc thông băng 105 trên Fig.6. Ngoài ra, cách triển khai ở Fig.8 bao gồm bộ chuyển đổi phổ 345 để chuyển đổi đầu ra của bộ xử lý xung 340 thường bao gồm sự đặt xung 305 và định tỷ lệ xung 315 của Fig.6. Ngoài ra, bộ xử lý xung 340 trên Fig.8 có thể bao gồm khối ổn định 210 như là đặc điểm tùy chọn và khối tìm kiếm cực trị 205 như là đặc điểm tùy chọn.

Tuy nhiên, các quy trình lọc thông cao 325, điều chỉnh đường bao 330 và sự tổ hợp của băng thấp và băng cao được thực hiện bởi giàn lọc tổng hợp, tức là, được thực hiện trong miền phổ, và đầu ra của giàn lọc tổng hợp 400 trên Fig.8 là tín hiệu âm thanh được nâng cao tần số miền thời gian 420. Tuy nhiên, nếu phần tử 400 được triển khai như là bộ tổ hợp đơn giản để tổ hợp các băng khác nhau, thì đầu ra của khối 400 cũng có thể là tín hiệu miền phổ đầy đủ thường bao gồm các khối giá trị phổ sau đó được xử lý thêm theo bất kỳ cách nào cần thiết.

Sau đây, ba ví dụ được đưa ra về các tín hiệu đặc trưng được mở rộng băng thông bằng cách sử dụng WESPE BWE. Tốc độ mẫu là 32kHz, DFT (105 trên Fig.8) với phổ một mặt có 513 vạch được sử dụng để trích xuất 8 tín hiệu thông băng chồng lấp. Để triển khai thông cao 4kHz (325 trên Fig.8), sự điều chỉnh đường bao phổ (330 trên Fig.8) và trộn (400 trên Fig.8) của LF và HF, DFT/IDFT tương tự (345 trên Fig.8) với 50% chồng lấp được sử dụng được tổ chức trong 16 băng hệ số định tỷ lệ đồng nhất. Tín hiệu kết quả hiển thị trong các ảnh phổ là PCM không được mã hóa từ DC đến 4kHz và được tạo ra bởi WESPE từ 4kHz đến 16kHz.

Fig.9a minh họa phần trích ngắn (một khối 1024 mẫu) có dạng sóng, đường bao chung và sự đặt xung được đồng bộ hóa và định tỷ lệ thu được bởi WESPE. Các xung lớn với độ phân tán nhẹ được đặt gần như cách đều nhau trong cấu trúc chu kỳ rộng.

Fig.9b mô tả ảnh phổ của toàn bộ mục thử nghiệm. Cấu trúc xung dọc của tiếng nói hữu thanh tiếp tục được căn chỉnh nhất quán giữa LF và HF, trong khi các âm xát thể hiện nhiễu âm như cấu trúc HF.

Do đó, Fig.9a cho thấy cách WESPE tạo mô hình các xung tiếng nói, hiển thị dạng sóng, đường bao chung và tạo ra xung, trong đó mục là "Tiếng nói nam giới

người Đức". Fig.9b cho thấy cách WESPE tạo mô hình các xung tiếng nói và hiển thị ảnh phổ. Mục là "Tiếng nói nam giới người Đức".

Fig.10a cho thấy phần trích ngắn (một khối 1024 mẫu) có dạng sóng, đường bao chung và sự đặt xung được đồng bộ hóa và định tỷ lệ thu được bởi WESPE. Các xung sắc nét riêng biệt được đặt cách đều nhau trong cấu trúc chu kỳ hẹp. Fig.10b mô tả ảnh phổ của toàn bộ mục thử nghiệm. Cấu trúc vạch ngang của ống chỉnh tần số cơ bản duy trì được căn chỉnh giữa LF và HF, tuy nhiên, HF cũng phần nào nhiễu và sẽ thu lợi từ sự ổn định bổ sung.

Fig.10a cho thấy cách WESPE tạo mô hình các âm bội sóng hài và hiển thị dạng sóng, đường bao chung và sự tạo ra xung. Mục là "Ống chỉnh tần số cơ bản". Fig.10b cho thấy cách WESPE tạo mô hình các âm bội sóng hài và hiển thị ảnh phổ. Mục là "Ống chỉnh tần số cơ bản".

Fig.11a cho thấy phần trích ngắn (một khối 1024 mẫu) có dạng sóng của mục thử nghiệm "Madonna Vogue", đường bao chung và sự đặt xung được đồng bộ hóa và định tỷ lệ kết quả bởi WESPE. Sự đặt và định tỷ lệ các xung có cấu trúc gần như ngẫu nhiên. Fig.11b mô tả ảnh phổ của toàn bộ mục thử nghiệm. Các cấu trúc nhất thời dọc của nhạc pop duy trì được căn chỉnh nhất quán giữa LF và HF, trong khi âm sắc HF chủ yếu là thấp.

Fig.11a cho thấy cách WESPE tạo mô hình hỗn hợp nhiễu và hiển thị dạng sóng, đường bao chung và sự tạo ra xung. Mục là "Vogue". Fig.11b cho thấy cách WESPE tạo mô hình hỗn hợp nhiễu và hiển thị ảnh phổ. Mục là "Vogue".

Kết thúc chèn C

Hình ảnh đầu tiên trên các hình Fig.9a, Fig.10a, Fig.11a minh họa dạng sóng của khối gồm 1024 mẫu tín hiệu nguồn băng thấp. Ngoài ra, ảnh hưởng của bộ lọc phân tích đối với việc trích xuất khối mẫu được thể hiện ở chỗ dạng sóng bằng 0 ở đầu khối, tức là, ở mẫu 0 và cũng bằng 0 ở cuối khối, tức là, ở mẫu 1023. Ví dụ, dạng sóng này có sẵn ở đầu vào của khối 100 của Fig.1 hoặc ở 30 trên Fig.6. Trục tung trên các hình Fig.9a, Fig.9b, Fig.9c luôn chỉ ra biên độ miền thời gian và trục hoành trong các hình này luôn chỉ ra biến thời gian và cụ thể, số mẫu thường kéo dài từ 0 đến 1023 cho một khối.

Hình ảnh thứ hai trên Fig.9a, Fig.10a, Fig.10b minh họa đường bao băng thấp được tính trung bình và cụ thể, chỉ phần dương của đường bao băng thấp. Tất nhiên, đường bao băng thấp thường là đối xứng và cũng mở rộng sang dải âm. Tuy nhiên, chỉ phần dương của đường bao băng thấp là cần thiết. Có thể thấy từ các hình Fig.9a, Fig.10a và Fig.11a rằng trong phương án này, đường bao chỉ được tính toán trong khi loại trừ cặp mẫu đầu tiên của khối và cặp mẫu cuối cùng của khối, tuy nhiên, điều này hoàn toàn không phải là vấn đề, vì các khối tốt hơn là nên được tính toán theo cách chồng lấp. Do đó, ví dụ, hình ảnh thứ hai của các hình Fig.9a, Fig.10a, Fig.11a, thường minh họa đầu ra của khối 100 của Fig.1 hoặc đầu ra của khối 130 của Fig.2.

Hình ảnh thứ ba trong các hình Fig.9a, Fig.10a, Fig.11a, minh họa tín hiệu tổng hợp sau khi định tỷ lệ xung, tức là, sau xử lý, trong đó các xung được đặt ở các giá trị theo thời gian của các đặc điểm của đường bao và đã được gán trọng số bởi các biên độ tương ứng của đường bao. Các hình Fig.9a, Fig.10a, Fig.11a minh họa rằng các xung được đặt chỉ kéo dài từ mẫu 256 đến mẫu 768. Do đó, tín hiệu bao gồm các xung được gán trọng số chỉ mở rộng trên 512 mẫu và không có bất kỳ phần nào trước các mẫu này và sau các mẫu này, tức là, bao phủ phần giữa của khung. Điều này phản ánh tình trạng khung trước đó có sự chồng lấp và khung sau đó cũng có sự chồng lấp. Với mục đích tạo ra tín hiệu xung với các khối sau đó, tín hiệu xung từ khối tiếp theo cũng sẽ được xử lý trong đó phần tư đầu tiên và phần tư cuối cùng bị thiếu và do đó, tín hiệu xung từ khối tiếp theo sẽ được đặt ngay sau tín hiệu xung được minh họa từ khối hiện tại trên các hình Fig.9a, Fig.10a, Fig.11a. Quy trình này rất hiệu quả, vì bất kỳ thuật toán chồng lấp/cộng nào của tín hiệu xung là không cần thiết. Tuy nhiên, bất kỳ quy trình chồng lấp/cộng nào hoặc bất kỳ quy trình nhiễu xuyên kênh nào từ một khung sang khung tiếp theo liên quan đến tín hiệu xung cũng có thể được thực hiện nếu cần thiết.

Các hình Fig.9b, Fig.10b, Fig.11b minh họa các ảnh phổ. Trục hoành đại diện cho thời gian, nhưng không phải là thời gian đối với các mẫu như trên các hình Fig.9a, 10a, 11a, mà là thời gian đối với các số khối DFT. Trục tung minh họa phổ tần số từ các tần số thấp ở dưới cùng của các hình tương ứng cho đến các tần số cao ở trên cùng của các hình tương ứng. Dải ngang mở rộng từ 0 đến 16kHz để phần tư dưới thể hiện tín hiệu gốc và ba phần tư trên thể hiện tín hiệu tổng hợp. Do đó, các hình Fig.9b,

Fig.10b, Fig.11b minh họa tín hiệu âm thanh được nâng cao tần số trong khi chỉ phần tư dưới của các hình này minh họa tín hiệu âm thanh nguồn.

Các hình chỉ ra rằng cấu trúc băng thấp được phản ánh rất tốt ở băng cao. Điều này đặc biệt dễ thấy, viện dẫn tới Fig.10b, minh họa ống chỉnh tần số cơ bản trong đó ba âm khác nhau của ống chỉnh tần số cơ bản được phát lần lượt từ trái sang phải trên Fig.10b. Cụ thể, phần đầu tiên bên trái của Fig.10b là âm thấp nhất của ống chỉnh tần số cơ bản, phần giữa là âm trung bình của ống chỉnh tần số cơ bản và phần bên phải của Fig.10b là âm cao nhất của ống chỉnh tần số cơ bản. Ống chỉnh tần số cơ bản được đặc trưng bởi phổ rất có âm sắc và dường như sáng chế này đặc biệt hữu ích trong việc tái tạo cấu trúc sóng hài ở mức 12kHz cao hơn rất tốt.

Đối với mục thử nghiệm thứ ba, có thể thấy rằng cấu trúc băng thấp của mục nhạc pop này được chuyển đổi rất tốt thành dải tần số cao bằng quy trình sáng chế.

Fig.12 minh họa phương án khác phần nào tương tự với phương án ở Fig.6. Do đó, các số tham chiếu tương tự trên Fig.6 chỉ ra các mục tương tự như trên Fig.12. Ngoài các đặc điểm trên Fig.6, phương án ở Fig.12 còn bao gồm bộ phân tách LF/HF 160, bộ tạo nhiễu âm ngẫu nhiên hoặc giả ngẫu nhiên 170, chẳng hạn như bảng nhiễu âm hoặc tương tự, và bộ điều chỉnh năng lượng 180.

Bộ phân tách LF/HF 160 thực hiện phân tách đường bao theo thời gian thành đường bao LF và đường bao HF. Tốt hơn là, đường bao LF được xác định bằng lọc thông thấp và đường bao HF được xác định bằng cách trừ đi đường bao LF từ đường bao HF.

Bộ tạo nhiễu âm ngẫu nhiên hoặc giả ngẫu nhiên 170 tạo ra tín hiệu nhiễu âm và bộ điều chỉnh năng lượng 180 điều chỉnh năng lượng nhiễu âm thành năng lượng đường bao HF cũng được ước lượng trong khối 180. Nhiễu âm này, có năng lượng được điều chỉnh thành năng lượng đường bao HF (không có bất kỳ sự đóng góp nào từ đường bao LF), được thêm bởi bộ cộng 335 vào chuỗi xung được gán trọng số như là đầu ra của khối 315. Tuy nhiên, thứ tự của các khối xử lý hoặc các bước 315, 335, ví dụ, cũng có thể được thay đổi.

Mặt khác, các quy trình liên quan đến các mục từ 205 đến 315 chỉ được áp dụng cho đường bao LF như được xác định bởi khối 160.

Phương án ưu tiên dựa trên sự phân tách của đường bao băng đầy đủ thành ít nhất hai phần bao gồm các khối sau đây hoặc các bước trong phần dưới đây hoặc theo bất kỳ thứ tự khả thi về mặt kỹ thuật nào khác:

Sự ước lượng đường bao theo thời gian 100: Sự nắn dòng, sự nén bằng cách sử dụng, ví dụ, hàm x^0,75; sự chia tách sau đó 160 của đường bao trong đường bao LF và đường bao HF. Đường bao LF thu được thông qua lọc thông thấp, trong đó tần số giao chéo là, ví dụ, từ 2 đến 6kHz. Theo phương án, đường bao HF là sự khác biệt giữa đường bao ban đầu và đường bao LF tốt hơn là được điều chỉnh độ trễ.

Sự đặt xung được đồng bộ hóa 300. Đường bao LF được suy ra trong bước được mô tả trước đây được phân tích bằng, ví dụ, thuật toán tính toán đường cong và sự đặt xung được thực hiện trên các vị trí cực đại của đường bao LF.

Sự định tỷ lệ độ lớn xung riêng lẻ 315 được suy ra từ đường bao: Chuỗi xung được kết hợp trong bước được mô tả trước đây được gán trọng số bởi các trọng số theo thời gian suy ra từ đường bao LF.

Năng lượng của đường bao HF được ước lượng và nhiễu âm ngẫu nhiên có cùng năng lượng được thêm 335 vào chuỗi xung được gán trọng số.

Sự xử lý sau, trích xuất HF hoặc lựa chọn điền đầy khoảng trống: Tín hiệu "thô" được tạo ra trong bước được mô tả ở trên ở đầu ra của khối 335 được xử lý sau tùy chọn 320, ví dụ, bằng cách bổ sung nhiễu âm, và được lọc 325 để sử dụng như HF trong BWE hoặc như tín hiệu ô đích điền đầy khoảng trống.

Sự điều chỉnh năng lượng 330: Sự phân bố năng lượng phổ của tín hiệu được lọc từ sự ước lượng năng lượng như đã nêu trong bước trên được điều chỉnh để sử dụng như HF trong BWE hoặc như tín hiệu ô đích điền đầy khoảng trống. Ở đây, thông tin phụ từ dòng bit về sự phân bố năng lượng mong muốn được ưu tiên sử dụng.

Sự trộn 400 HF hoặc tín hiệu điền đầy khoảng trống với LF: Cuối cùng, tín hiệu được điều chỉnh từ bước 5 được trộn với đầu ra mã hóa lõi tương ứng với các nguyên lý BWE hoặc điền đầy khoảng trống thông thường, tức là, đi qua bộ lọc HP và bổ sung cho LF hoặc điền đầy các lỗ phổ trong các vùng phổ điền đầy khoảng trống.

Ở đây cần đề cập rằng tất cả các lựa chọn thay thế hoặc các khía cạnh như đã thảo luận trước đây và tất cả các khía cạnh được định nghĩa bởi các yêu cầu bảo hộ độc lập trong các yêu cầu bảo hộ sau đây đều có thể được sử dụng riêng lẻ, tức là

không có bất kỳ cách thay thế hoặc đối tượng nào khác ngoài cách thay thế, đối tượng hoặc yêu cầu bảo hộ độc lập được dự tính. Tuy nhiên, theo các phương án khác, hai hoặc nhiều cách thay thế hoặc các khía cạnh hoặc các yêu cầu bảo hộ độc lập có thể được kết hợp với nhau và theo các phương án khác, tất cả các khía cạnh hoặc phương án thay thế và tất cả các yêu cầu bảo hộ độc lập có thể được kết hợp với nhau.

Tín hiệu âm thanh được mã hóa theo sáng chế có thể được lưu trữ trên vật ghi lưu trữ kỹ thuật số hoặc vật ghi lưu trữ không tạm thời hoặc có thể được truyền trên vật ghi truyền dẫn như vật ghi truyền dẫn không dây hoặc vật ghi truyền dẫn có dây như Internet.

Mặc dù một vài khía cạnh đã được mô tả trong ngữ cảnh của thiết bị, rõ ràng là các khía cạnh này cũng thể hiện mô tả của phương pháp tương ứng, trong đó khối hoặc thiết bị tương ứng với bước xử lý của phương pháp hoặc đặc điểm của bước xử lý của phương pháp. Tương tự, các khía cạnh được mô tả trong ngữ cảnh của bước xử lý của phương pháp cũng thể hiện sự mô tả của khối hoặc mục hoặc đặc điểm tương ứng của thiết bị tương ứng.

Phụ thuộc vào các yêu cầu triển khai nhất định, các phương án của sáng chế có thể được triển khai trong phần cứng hoặc trong phần mềm. Cách triển khai có thể được thực hiện sử dụng vật ghi lưu trữ số, ví dụ, đĩa mềm, DVD, CD, ROM, PROM, EPROM, EEPROM hoặc bộ nhớ FLASH, có các tín hiệu điều khiển có thể đọc được bằng điện tử được lưu trữ trên đó, mà kết hợp (hoặc có thể kết hợp) với hệ thống máy tính có thể lập trình sao cho phương pháp tương ứng được thực hiện.

Một số phương án theo sáng chế bao gồm vật mang dữ liệu có các tín hiệu điều khiển có thể đọc được bằng điện tử, mà có khả năng kết hợp với hệ thống máy tính có thể lập 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 triển khai như là sản phẩm chương trình máy tính với mã chương trình, mã chương trình có tác dụng để thực hiện một trong số các phương pháp khi sản phẩm chương trình máy tính chạy trên máy tính. Mã chương trình có thể, ví dụ, được lưu trữ trên vật mang có thể đọc được bằng máy.

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

Do đó, nói cách khác, 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 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.

Do đó, phương án khác của các phương pháp theo sáng chế là dòng dữ liệu hoặc chuỗi tín hiệu biểu diễn chương trình máy tính để thực hiện một trong số các phương pháp đã được mô tả ở đây. Ví dụ, dòng dữ liệu hoặc chuỗi tín hiệu có thể 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 công cụ xử lý, ví dụ, máy tính, hoặc thiết bị logic có thể lập trình, được tạo cấu hình để, hoặc thích ứng cho, việc 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 đó để thực hiện một trong số các phương pháp được mô tả ở đây.

Theo một số phương án, thiết bị lôgic có thể lập trình (ví dụ, khả trình 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, 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 hoạ cho các nguyên lý của sáng chế. Cần hiểu rằng các biến thể và biến đổi của các phương án và các chi tiết được mô tả ở đây là rõ ràng đối với người có kỹ năng trong lĩnh vực kỹ thuật này. Do đó, mục đích của sáng chế chỉ bị giới hạn bởi phạm vi của các điểm yêu cầu bảo hộ sắp đưa ra dưới đây và không bị giới hạn bởi các chi tiết cụ thể được được thể hiện bằng cách mô tả và giải thích của các phương án ở đây.

YÊU CẦU BẢO HỘ

1. Bộ xử lý âm thanh để tạo ra tín hiệu âm thanh được nâng cao tần số (420) từ tín hiệu âm thanh nguồn (30), bộ xử lý âm thanh bao gồm:

bộ xác định đường bao (100) để xác định đường bao theo thời gian của ít nhất một phần của tín hiệu âm thanh nguồn;

bộ phân tích (200) để phân tích đường bao theo thời gian để xác định các giá trị có các đặc điểm nhất định của đường bao theo thời gian;

bộ tổng hợp tín hiệu (300) để tạo ra tín hiệu tổng hợp, việc tạo ra bao gồm các xung định vị liên quan đến các giá trị theo thời gian được xác định, trong đó các xung được gán trọng số sử dụng các trọng số được suy ra từ các biên độ của đường bao theo thời gian về các giá trị theo thời gian, trong đó các xung được định vị; và

bộ tổ hợp (400) để tổ hợp ít nhất băng của tín hiệu tổng hợp mà không nằm trong tín hiệu âm thanh nguồn và tín hiệu âm thanh nguồn để thu được tín hiệu âm thanh được nâng cao tần số (420).

2. Bộ xử lý âm thanh theo điểm 1, trong đó bộ phân tích (200) được tạo cấu hình để xác định các giá trị theo thời gian có cực đại cục bộ hoặc cực tiểu cục bộ làm các đặc điểm nhất định.

3. Bộ xử lý âm thanh theo điểm 1 hoặc 2, trong đó bộ xác định đường bao (100) được tạo cấu hình để

phân tách (105) tín hiệu âm thanh nguồn thành nhiều tín hiệu băng con,

tính toán (120) đường bao theo thời gian được lựa chọn của tín hiệu băng con được lựa chọn từ nhiều tín hiệu băng con, đường bao theo thời gian được lựa chọn là đường bao theo thời gian, hoặc

tính toán (120) ít nhất hai đường bao theo thời gian từ ít nhất hai tín hiệu băng con trong số nhiều tín hiệu băng con và tổ hợp (130) ít nhất hai tín hiệu băng con để đạt được đường bao theo thời gian được tổ hợp thành đường bao theo thời gian.

4. Bộ xử lý âm thanh theo điểm 3,

trong đó bộ xác định đường bao (100) được tạo cấu hình để chuẩn hóa (115, 125) hoặc lọc các tín hiệu băng con được lựa chọn hoặc các đường bao theo thời gian trước khi tổ hợp (130), hoặc

trong đó việc tổ hợp bao gồm thuật toán tính trung bình, hoặc

trong đó bộ xác định đường bao (100) được tạo cấu hình để tính toán các đường bao theo thời gian từ tất cả các tín hiệu băng con trong số nhiều tín hiệu băng con, hoặc

trong đó bộ xác định đường bao (100) được tạo cấu hình để xác định đường bao theo thời gian băng rộng đơn lẻ của tín hiệu âm thanh nguồn làm đường bao theo thời gian.

5. Bộ xử lý âm thanh theo một trong số các điểm nêu trên, trong đó bộ xác định đường bao (100) được tạo cấu hình để xác định đường bao theo thời gian

sử dụng bộ nhắc lại đường bao được tạo cấu hình để nắn dòng (145) dạng sóng và lọc thông thấp (150) dạng sóng được nắn dòng, hoặc

sử dụng việc tính toán các giá trị tuyệt đối hoặc các lũy thừa của các giá trị tuyệt đối của dạng sóng số và lọc thông thấp kết quả, hoặc

sử dụng việc tính toán giá trị căn bậc hai trung bình tức thời của dạng sóng thông qua cửa sổ trượt có chiều rộng cửa sổ được xác định, hoặc

sử dụng phép làm gần đúng tuyến tính theo mảnh cho dạng sóng, trong đó đường bao theo thời gian được xác định bằng cách tìm và kết nối các đỉnh của dạng sóng trong cửa sổ trượt di chuyển thông qua kết quả của phép làm gần đúng tuyến tính theo mảnh, hoặc

sử dụng biến đổi Hilbert để tạo ra tín hiệu phân tích cho dạng sóng và tính toán đường bao từ tín hiệu âm thanh nguồn và tín hiệu phân tích sử dụng các thuật toán bình phương, các thuật toán cộng và các thuật toán căn bậc hai.

6. Bộ xử lý âm thanh theo một trong số các điểm nêu trên, trong đó bộ phân tích (200) được tạo cấu hình để

xác định (205) các giá trị theo thời gian ban đầu có các đặc điểm nhất định, và

để suy ra (210), từ các giá trị theo thời gian ban đầu, các giá trị theo thời gian sử dụng hàm tối ưu hóa, hoặc sử dụng thông tin phụ được kết hợp với tín hiệu âm thanh nguồn, hoặc lựa chọn hoặc điều khiển các giá trị bổ sung theo quy tắc xử lý.

7. Bộ xử lý âm thanh theo điểm 6, trong đó quy tắc xử lý hoặc hàm tối ưu hóa được triển khai sao cho các giá trị theo thời gian được đặt trong đường quét với khoảng cách đường quét (T), trong đó khoảng cách đường quét (T) và vị trí của đường quét trong

đường bao theo thời gian là để giá trị độ lệch giữa các giá trị theo thời gian và các giá trị theo thời gian ban đầu có đặc trưng được định trước.

8. Bộ xử lý âm thanh theo điểm 7, trong đó giá trị độ lệch là tổng của các hiệu số bình phương và trong đó đặc trưng được định trước là đặc trưng tối thiểu.

9. Bộ xử lý âm thanh theo một trong số các điểm nêu trên, trong đó bộ tổng hợp tín hiệu (300) được tạo cấu hình để

đặt (35) chỉ các xung dương hoặc chỉ các xung âm để thu được chuỗi xung, và sau đó gán trọng số (315) cho các xung trong chuỗi xung, hoặc

gán trọng số cho chỉ các xung âm hoặc chỉ các xung dương sử dụng các phép gán trọng số tương ứng với các giá trị theo thời gian của các xung trong chuỗi xung, và

đặt các xung được gán trọng số tại các giá trị theo thời gian tương ứng để thu được chuỗi xung.

10. Bộ xử lý âm thanh theo một trong số các điểm nêu trên,

trong đó bộ tổng hợp tín hiệu (300) được tạo cấu hình để suy ra (310) các trọng số từ các biên độ, sử dụng hàm nén, hàm nén là hàm từ nhóm các hàm bao gồm:

hàm lũy thừa với lũy thừa nhỏ hơn 1, hàm logarit, hàm căn bậc hai, và hàm không tuyến tính được tạo cấu hình để làm giảm các giá trị cao hơn và tăng các giá trị thấp hơn.

11. Bộ xử lý âm thanh theo một trong số các điểm nêu trên, trong đó bộ tổng hợp tín hiệu (300) được tạo cấu hình để thực hiện hàm xử lý sau (320), hàm xử lý sau bao gồm ít nhất một hàm trong số nhóm các hàm gồm thêm nhiễu âm, thêm sóng hài thiếu, lọc nghịch đảo, và điều chỉnh đường bao (330).

12. Bộ xử lý âm thanh theo một trong số các điểm nêu trên,

trong đó bộ xác định đường bao (100) được tạo cấu hình để phân tách (160) đường bao theo thời gian thành phần tần số thấp và phần tần số cao,

trong đó bộ phân tích (200) được tạo cấu hình để sử dụng phần tần số thấp của đường bao theo thời gian để phân tích.

13. Bộ xử lý âm thanh theo điểm 12, trong đó bộ tổng hợp tín hiệu (300) được tạo cấu hình để tạo ra (170) nhiễu âm được điều chỉnh năng lượng (180), và thêm (335) nhiễu âm được điều chỉnh năng lượng (180) vào tín hiệu bao gồm các xung được gán trọng số hoặc không được gán trọng số để thu được tín hiệu tổng hợp.

14. Bộ xử lý âm thanh theo một trong số các điểm nêu trên,

trong đó bộ tổng hợp tín hiệu (300) được tạo cấu hình để lọc thông cao (325) hoặc để lọc thông băng (325) tín hiệu bao gồm các xung được đặt và được gán trọng số để thu được ít nhất băng của tín hiệu tổng hợp mà không nằm trong tín hiệu âm thanh nguồn và để thực hiện sự điều chỉnh đường bao phổ (330) với băng của tín hiệu tổng hợp, hoặc trong đó sự điều chỉnh đường bao phổ được thực hiện sử dụng các giá trị điều chỉnh đường bao được suy ra từ thông tin phụ được kết hợp với tín hiệu âm thanh nguồn hoặc sử dụng các giá trị điều chỉnh đường bao được suy ra từ tín hiệu âm thanh nguồn hoặc được kết hợp với hàm điều chỉnh đường bao được định trước.

15. Bộ xử lý âm thanh theo một trong số các điểm nêu trên, trong đó tín hiệu âm thanh nguồn là tín hiệu âm thanh miền thời gian,

trong đó ít nhất băng của tín hiệu tổng hợp là tín hiệu âm thanh miền thời gian, và

trong đó bộ tổ hợp (400) được tạo cấu hình để thực hiện tổ hợp miền thời gian sử dụng phép cộng từng mẫu các mẫu của ít nhất một băng của tín hiệu tổng hợp và các mẫu tương ứng của tín hiệu âm thanh nguồn (30).

16. Bộ xử lý âm thanh theo một trong số các điểm nêu trên, trong đó tín hiệu âm thanh nguồn là tín hiệu kích thích trong miền mã hóa dự báo tuyến tính (linear prediction coding – LPC),

trong đó ít nhất một băng của tín hiệu tổng hợp là tín hiệu kích thích trong miền LPC,

trong đó bộ tổ hợp (400) được tạo cấu hình để tổ hợp (405) tín hiệu âm thanh nguồn và ít nhất một băng bằng phép cộng từng mẫu trong miền LPC,

trong đó bộ tổ hợp (400) được tạo cấu hình để lọc (410) kết quả của phép cộng từng mẫu sử dụng bộ lọc tổng hợp LPC để thu được tín hiệu âm thanh được nâng cao tần số, và

trong đó bộ lọc tổng hợp LPC (410) được điều khiển bởi dữ liệu LPC (40a) được kết hợp với tín hiệu âm thanh nguồn như thông tin phụ, và trong đó bộ lọc tổng hợp LPC (410) còn được điều khiển bởi thông tin đường bao (40b) cho ít nhất một băng của tín hiệu tổng hợp.

17. Bộ xử lý âm thanh theo một trong số các điểm nêu trên,

trong đó bộ phân tích (200), bộ tổng hợp tín hiệu (300) và bộ tổ hợp (400) hoạt động trong miền thời gian hoặc miền thời gian LPC.

18. Bộ xử lý âm thanh theo một trong số các điểm từ 1 đến 16, trong đó bộ xác định đường bao (100) được tạo cấu hình để áp dụng sự chuyển đổi phổ (105) để trích xuất nhiều tín hiệu thông băng cho chuỗi khung,

trong đó bộ tổng hợp tín hiệu (300) được tạo cấu hình để áp dụng chuyển đổi phổ, để trích xuất (325) ít nhất một băng của tín hiệu tổng hợp, và để thực hiện (330) điều chỉnh đường bao cho ít nhất một băng, và

trong đó bộ tổ hợp (400) được tạo cấu hình để tổ hợp trong miền phổ và để áp dụng chuyển đổi thành miền thời gian để thu được tín hiệu âm thanh được nâng cao tần số (420).

19. Phương pháp tạo ra tín hiệu âm thanh được nâng cao tần số (420) từ tín hiệu âm thanh nguồn, phương pháp bao gồm các bước:

xác định (100) đường bao theo thời gian của ít nhất một phần của tín hiệu âm thanh nguồn;

phân tích (200) đường bao theo thời gian để xác định các giá trị theo thời gian của các đặc điểm nhất định của đường bao theo thời gian;

đặt (305) các xung liên quan đến các giá trị theo thời gian được xác định vào tín hiệu tổng hợp, trong đó, trong tín hiệu tổng hợp, các xung được đặt được gán trọng số sử dụng các trọng số được suy ra từ các biên độ của đường bao theo thời gian có liên quan đến các giá trị theo thời gian, tại đó các xung được đặt; và

tổ hợp (400) ít nhất băng của tín hiệu tổng hợp mà không nằm trong tín hiệu âm thanh nguồn để thu được tín hiệu âm thanh được nâng cao tần số (420).

20. Vật ghi có thể đọc được bằng máy tính chứa chương trình máy tính để thực hiện, khi chạy trên máy tính hoặc bộ xử lý, phương pháp tạo ra tín hiệu âm thanh được nâng cao tần số theo điểm 19.

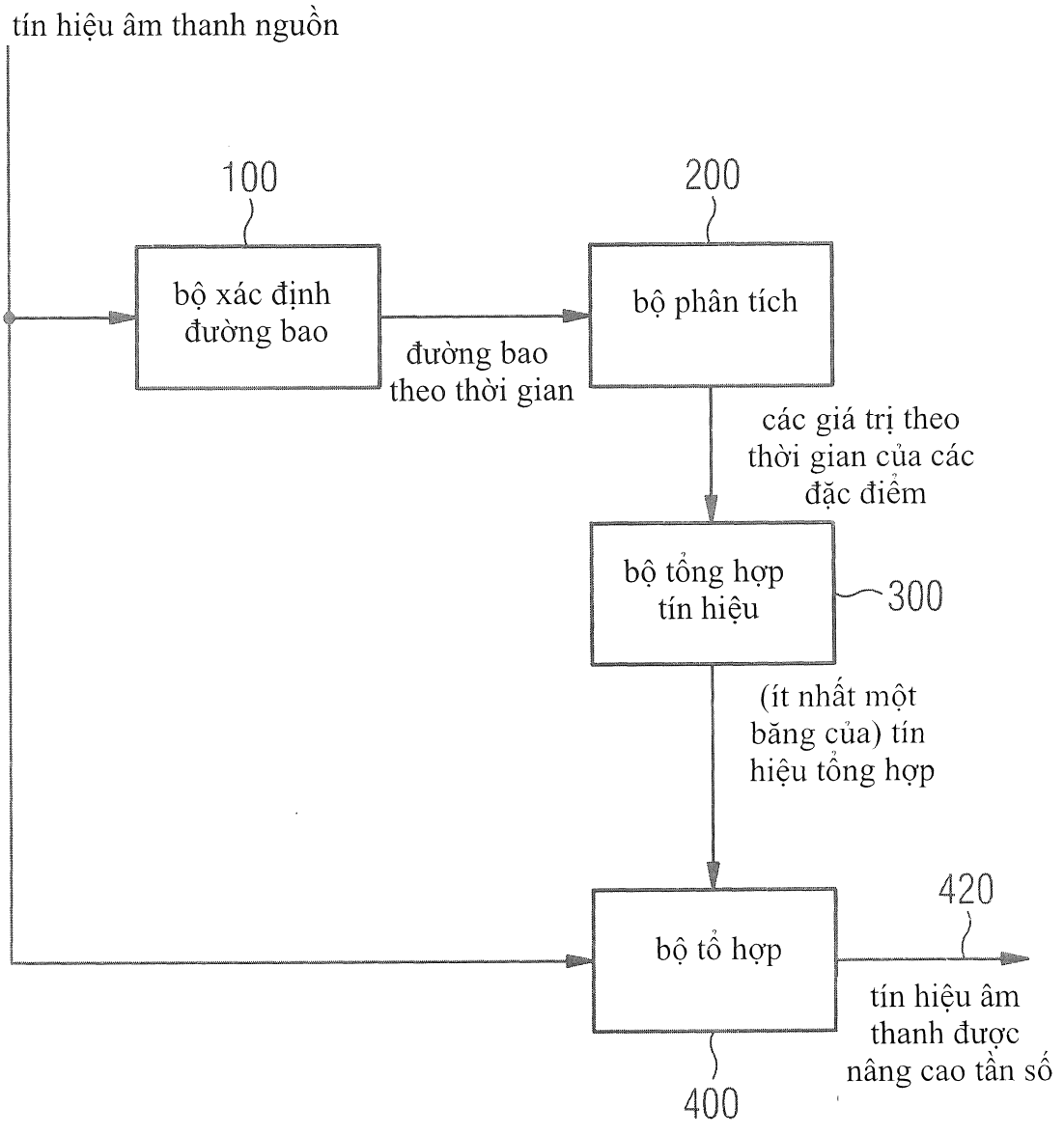
tín hiệu âm thanh nguồn
100
bộ xác định đường bao
đường bao theo thời gian
200
bộ phân tích
các giá trị theo thời gian của các đặc điểm
bộ tổng hợp tín hiệu
300
(ít nhất một băng của) tín hiệu tổng hợp
420
bộ tổ hợp
tín hiệu âm thanh được nâng cao tần số
400

Fig. 1

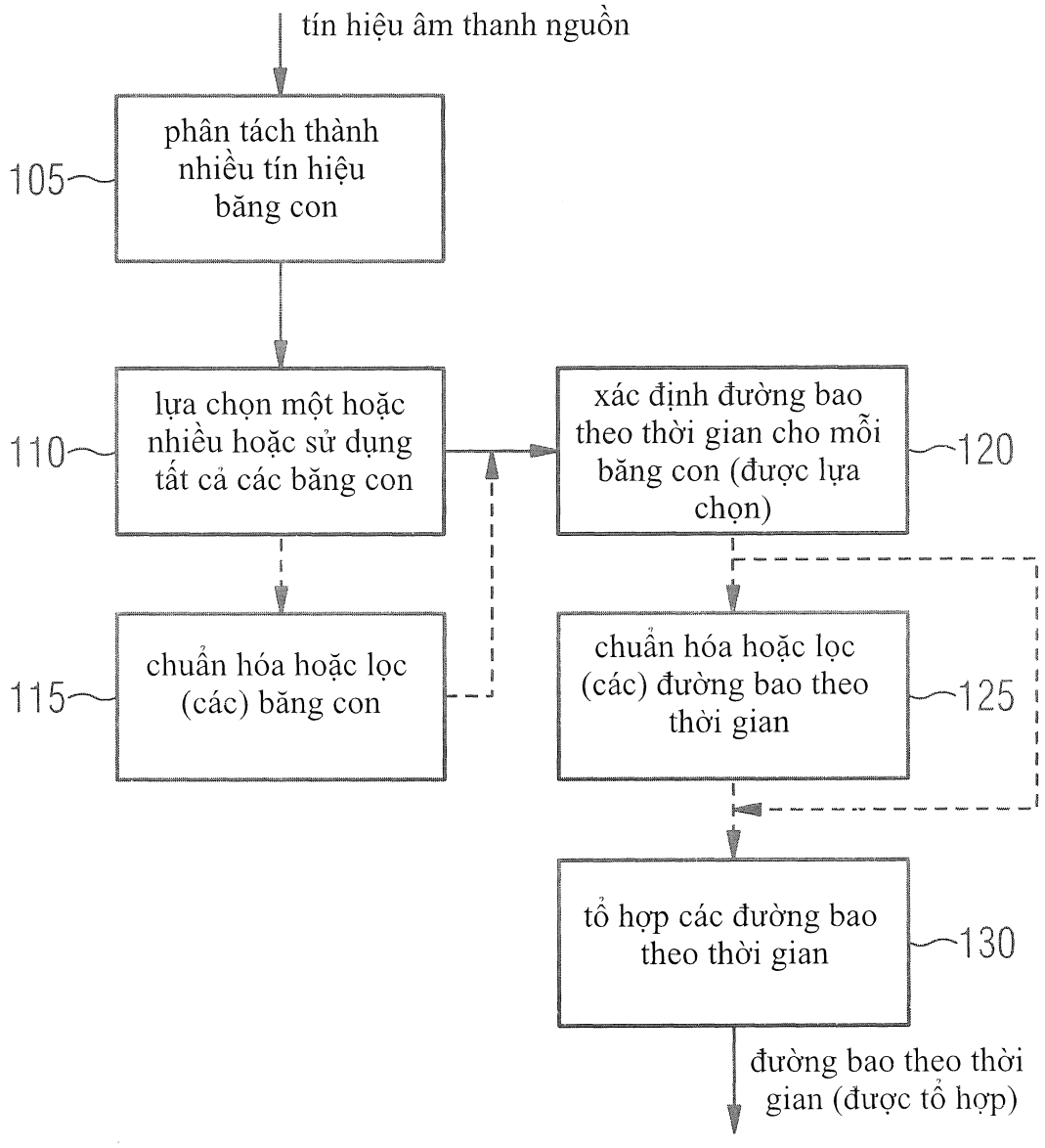
tín hiệu âm thanh nguồn
phân tách thành nhiều tín hiệu băng con
105
lựa chọn một hoặc nhiều hoặc sử dụng tất cả các băng con
110
xác định đường bao theo thời gian cho mỗi băng con (được lựa chọn)
120
chuẩn hóa hoặc lọc (các) băng con
115
chuẩn hóa hoặc lọc (các) đường bao theo thời gian
125
tổ hợp các đường bao theo thời gian
130
đường bao theo thời gian (được tổ hợp)

Fig. 2

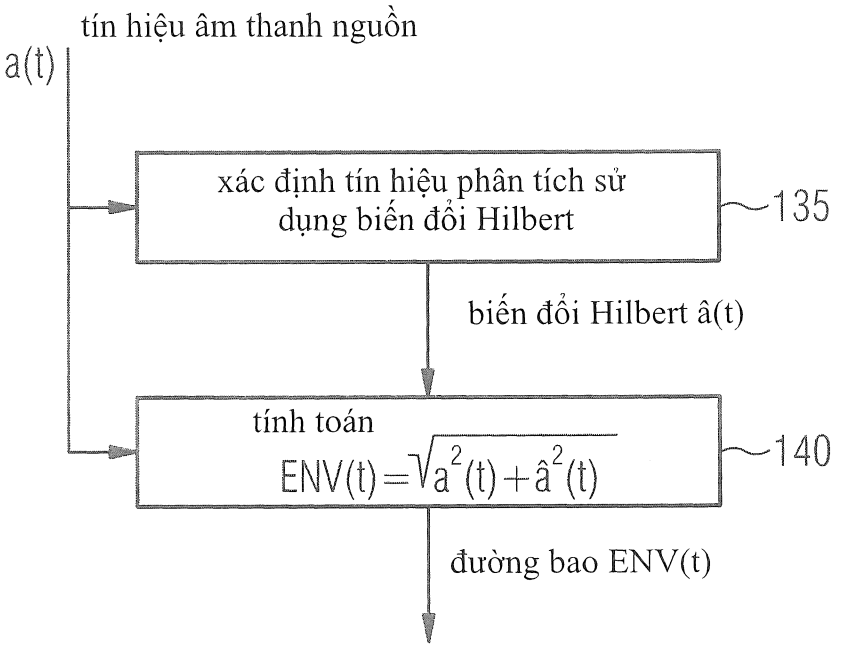
tín hiệu âm thanh nguồn
a(t)
xác định tín hiệu phân tích sử dụng biến đổi Hilbert
135
biến đổi Hilbert â(t)
tính toán
$ENV(t)=\sqrt{a^2(t)+\hat{a}^2(t)}$
140
đường bao ENV(t)

Fig. 3a

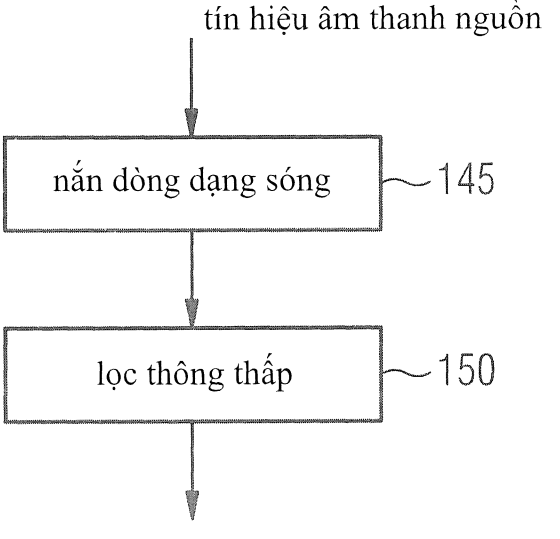
tín hiệu âm thanh nguồn
nắn dòng dạng sóng
145
lọc thông thấp
150

Fig. 3b

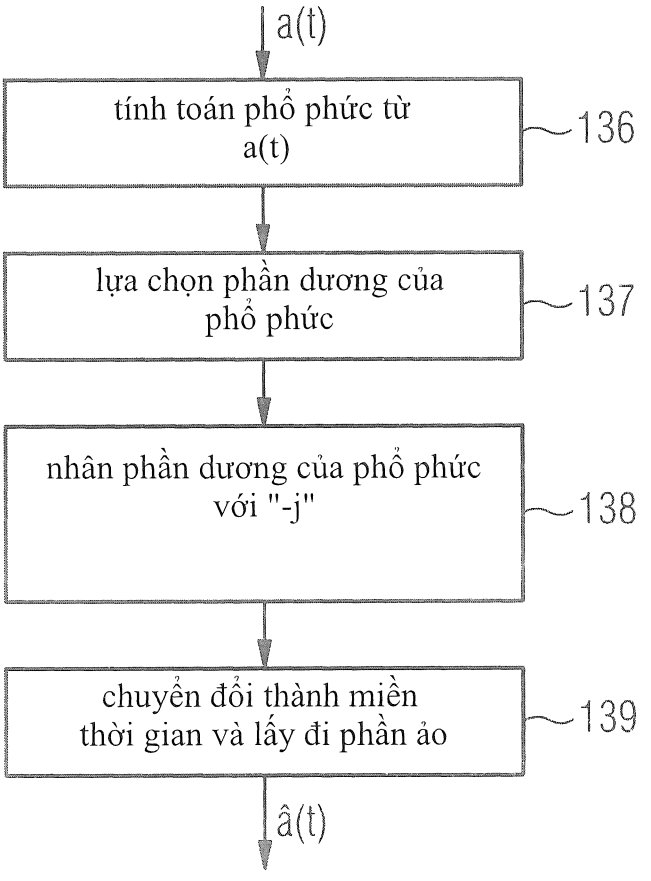
a(t)
tính toán phổ phức từ a(t)
136
lựa chọn phần dương của phổ phức
137
nhân phần dương của phổ phức với "-j"
138
chuyển đổi thành miền thời gian và lấy đi phần ảo
139
â(t)

Fig. 3c

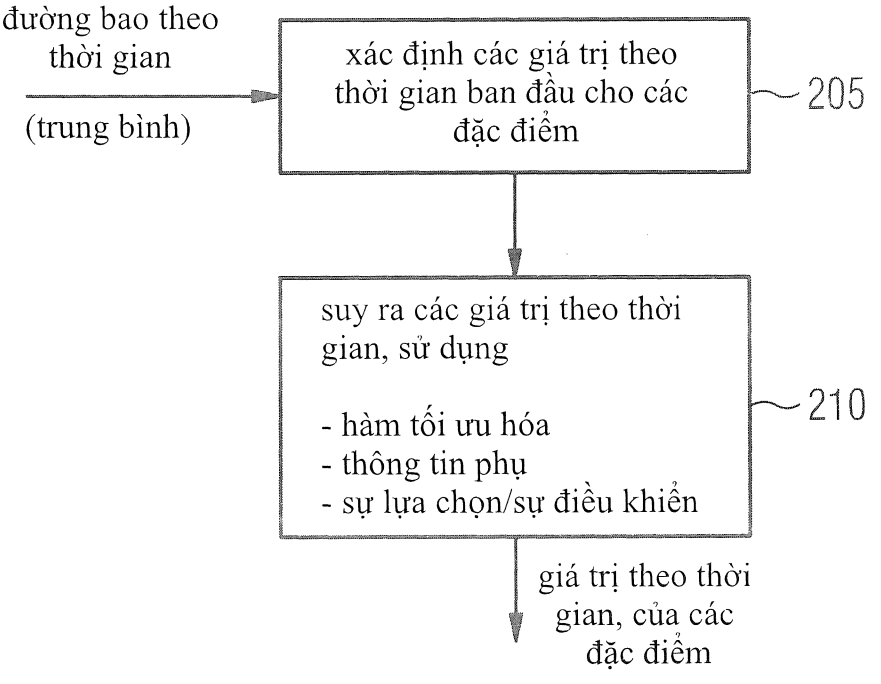
đường bao theo thời gian
(trung bình)
xác định các giá trị theo thời gian ban đầu cho các đặc điểm
205
suy ra các giá trị theo thời gian, sử dụng
- hàm tối ưu hóa
- thông tin phụ
- sự lựa chọn/sự điều khiển
210
giá trị theo thời gian, của các đặc điểm

Fig. 4

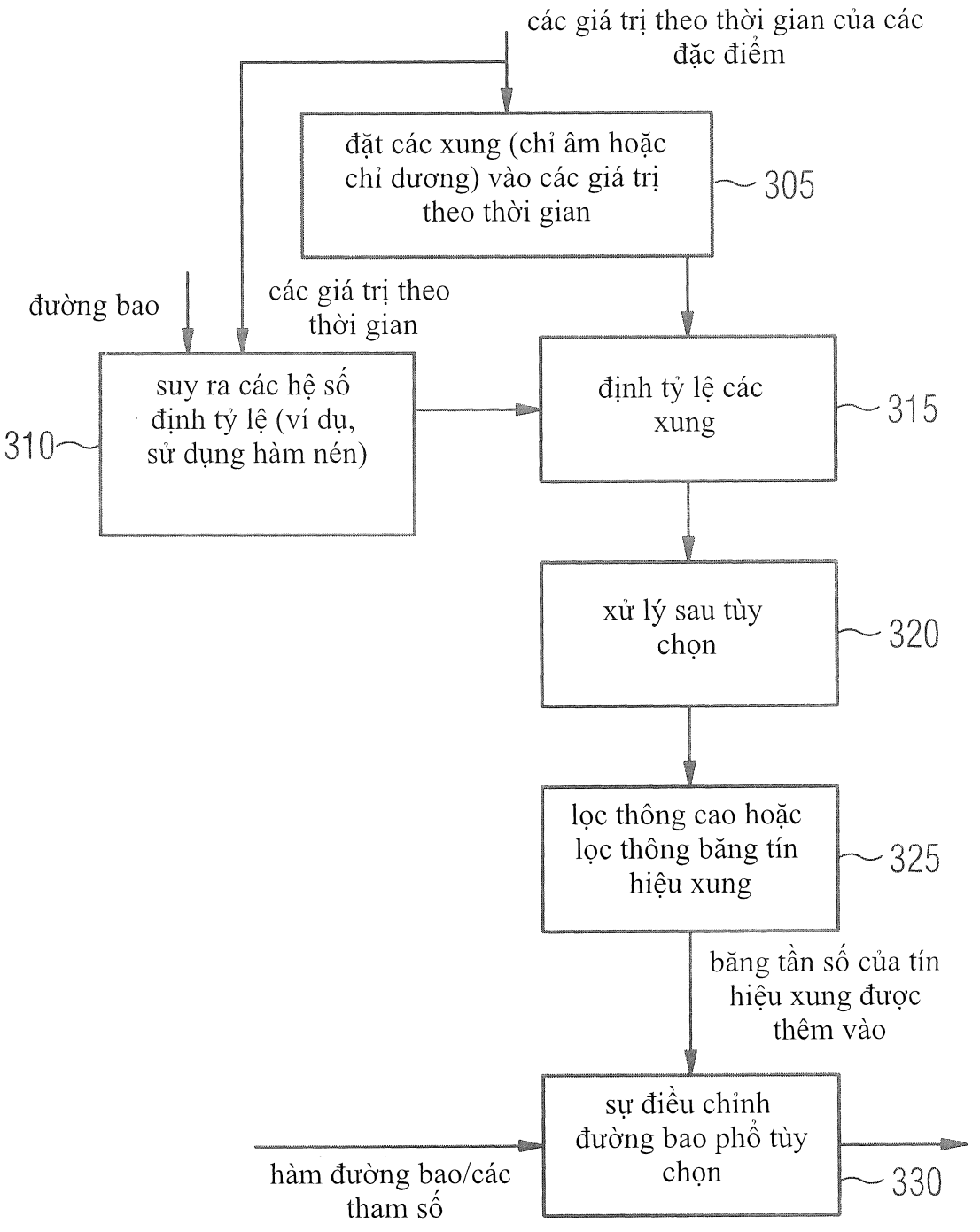
các giá trị theo thời gian của các đặc điểm
đặt các xung (chỉ âm hoặc chỉ dương) vào các giá trị theo thời gian
305
đường bao
các giá trị theo thời gian
suy ra các hệ số định tỷ lệ (ví dụ, sử dụng hàm nén)
310
định tỷ lệ các xung
315
xử lý sau tùy chọn
320
lọc thông cao hoặc lọc thông băng tín hiệu xung
325
băng tần số của tín hiệu xung được thêm vào
sự điều chỉnh đường bao phổ tùy chọn
330
hàm đường bao/các tham số

Fig. 5

6/14

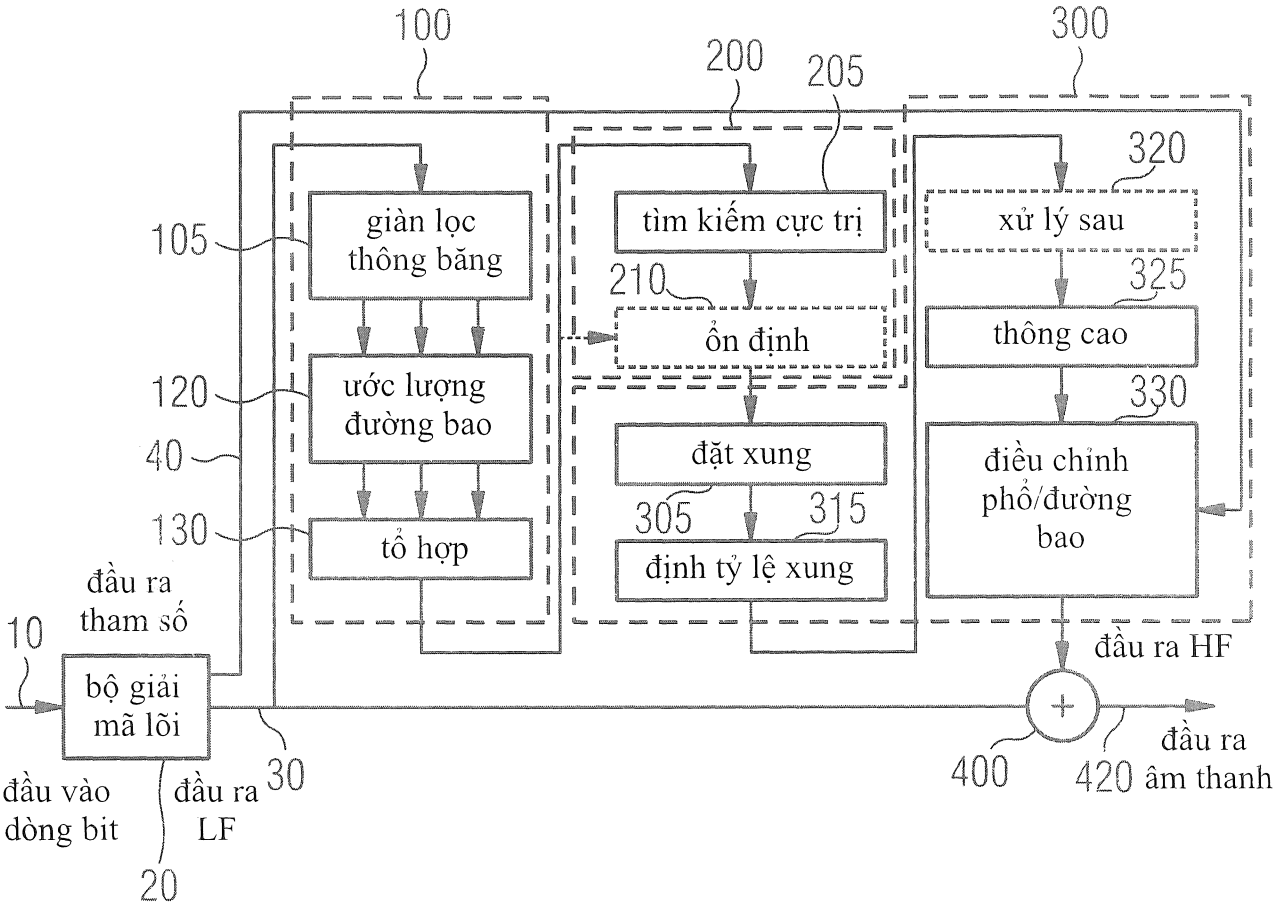

Fig. 6

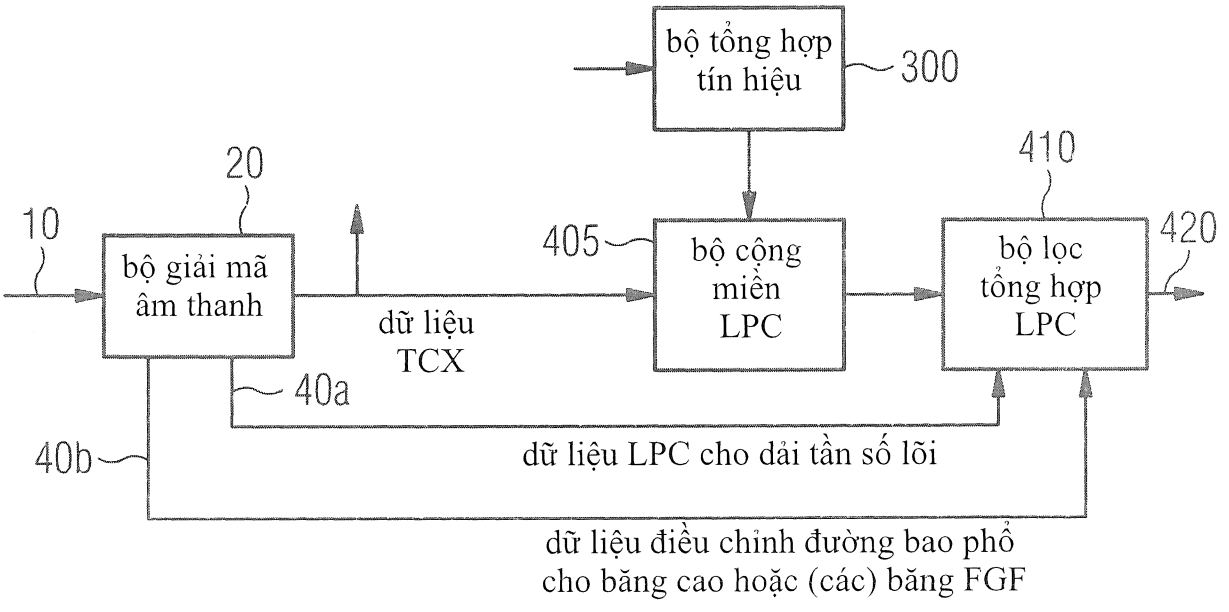

Fig. 7

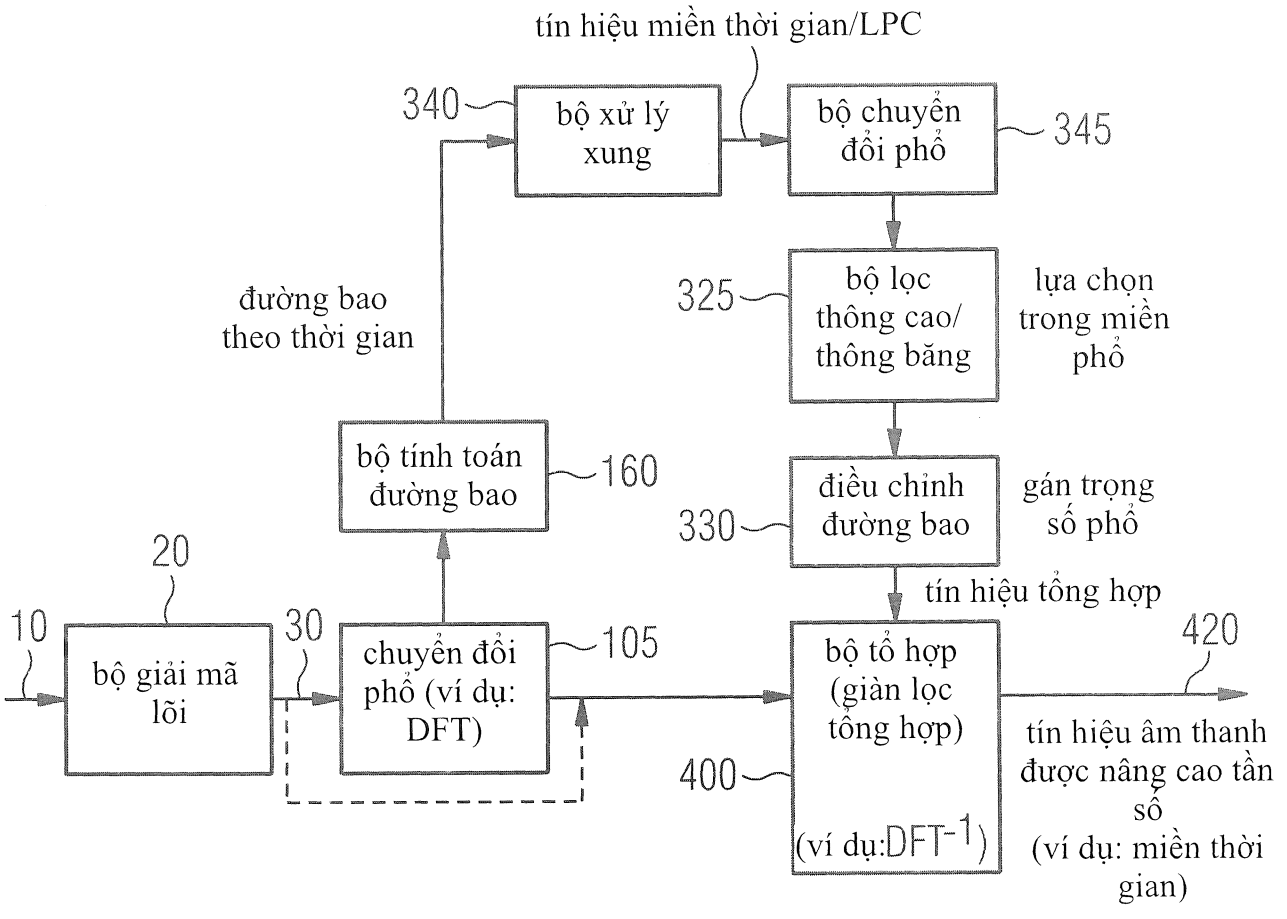
tín hiệu miền thời gian/LPC
340
bộ xử lý xung
bộ chuyển đổi phổ
345
325
bộ lọc thông cao/ thông băng
lựa chọn trong miền phổ
đường bao theo thời gian
bộ tính toán đường bao
160
điều chỉnh đường bao
330
gán trọng số phổ
tín hiệu tổng hợp
20
10
30
bộ giải mã lõi
chuyển đổi phổ (ví dụ: DFT)
105
bộ tổ hợp (giàn lọc tổng hợp)
420
400
(ví dụ:DFT-1)
tín hiệu âm thanh được nâng cao tần số (ví dụ: miền thời gian)

Fig. 8

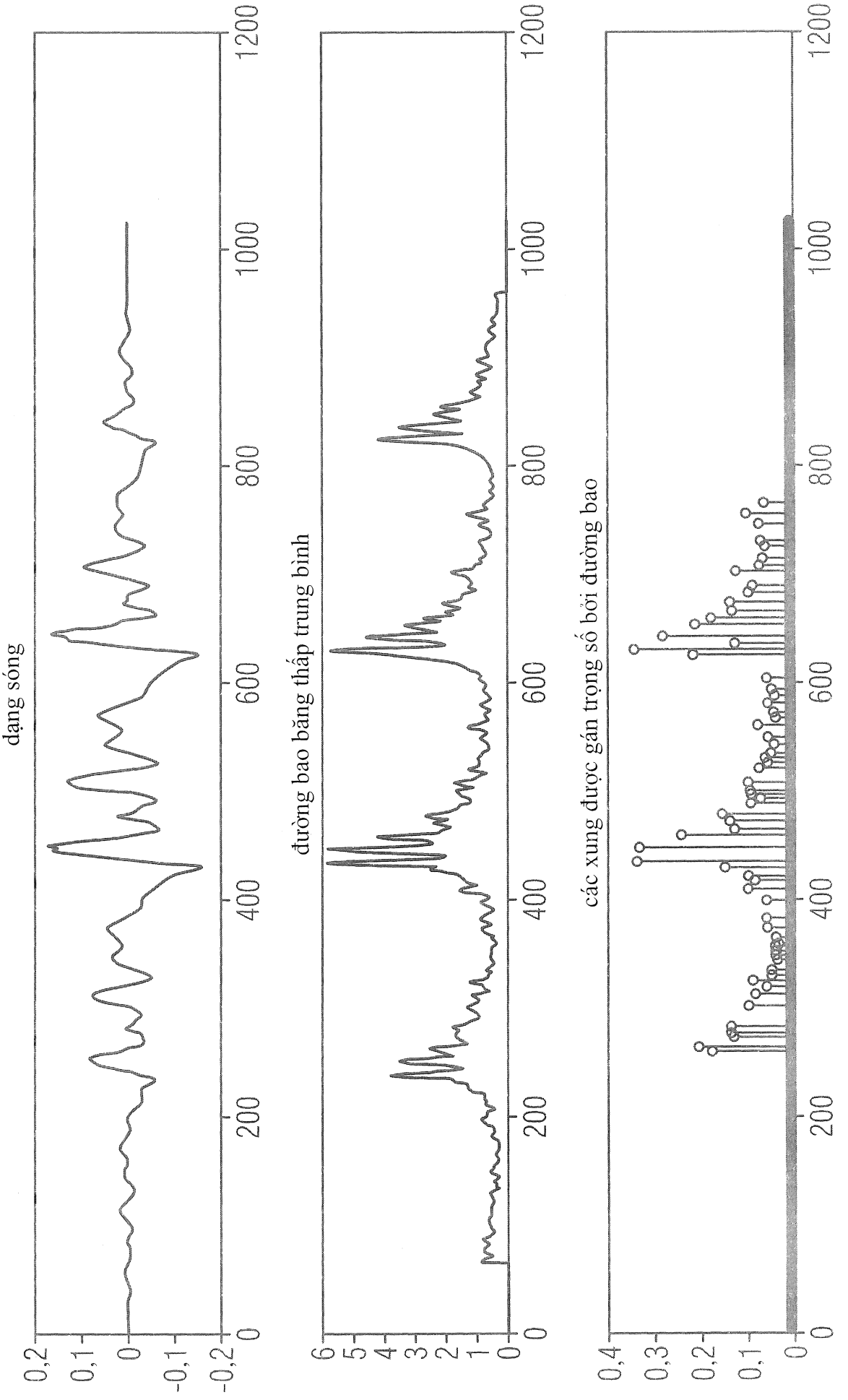
dạng sóng
0,2
0,1
0
-0,1
-0,2
0
200
400
600
800
1000
1200
đường bao băng thấp trung bình
6
5
4
3
2
1
0
0
200
400
600
800
1000
1200
các xung được gán trọng số bởi đường bao
0,4
0,3
0,2
0,1
0
0
200
400
600
800
1000
1200

Fig. 9a

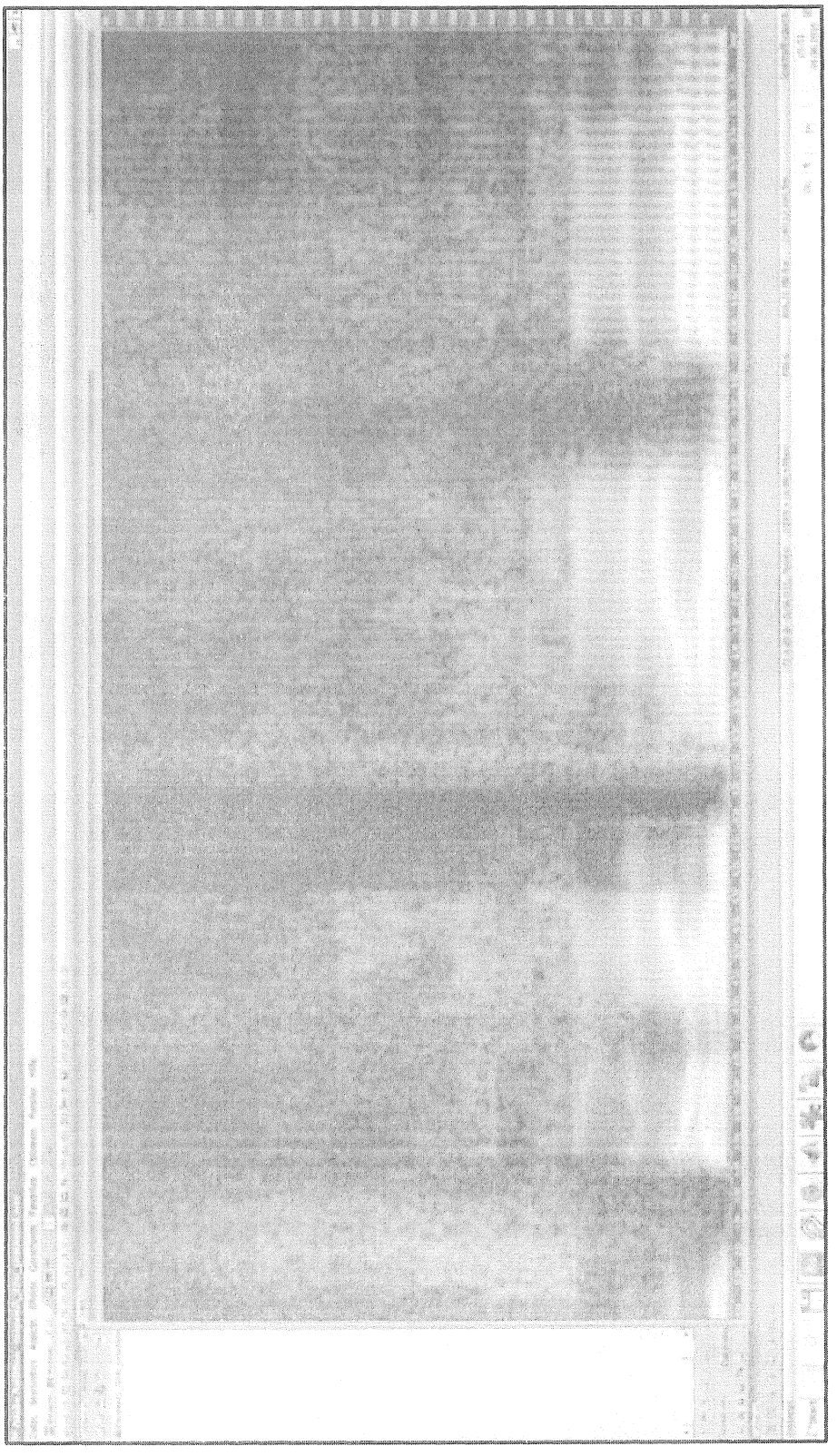

Fig. 9b

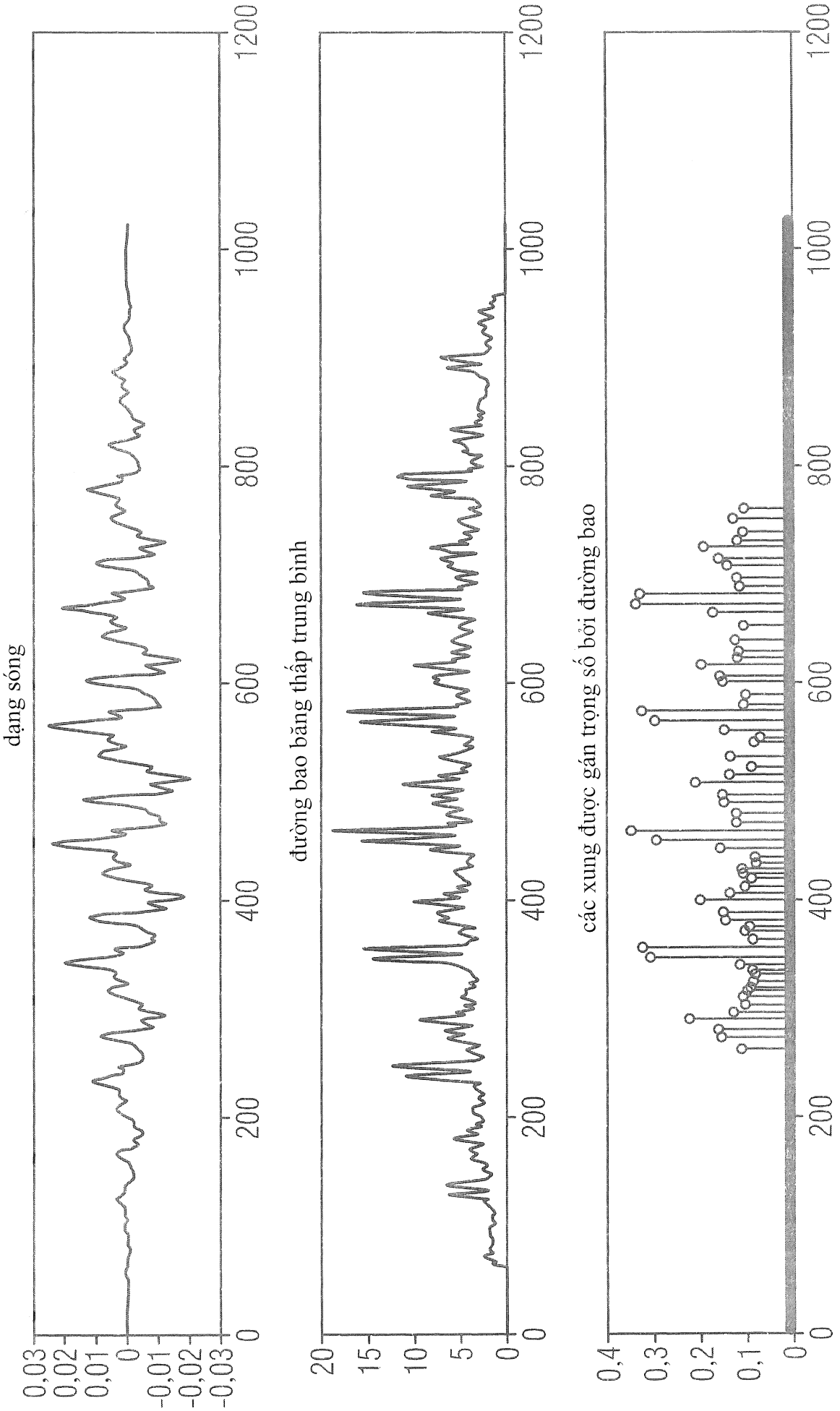
dạng sóng
0,03
0,02
0,01
0
-0,01
-0,02
-0,03
0
200
400
600
800
1000
1200
đường bao băng thấp trung bình
20
15
10
5
0
0
200
400
600
800
1000
1200
các xung được gán trọng số bởi đường bao
0,4
0,3
0,2
0,1
0
0
200
400
600
800
1000
1200

Fig. 10a

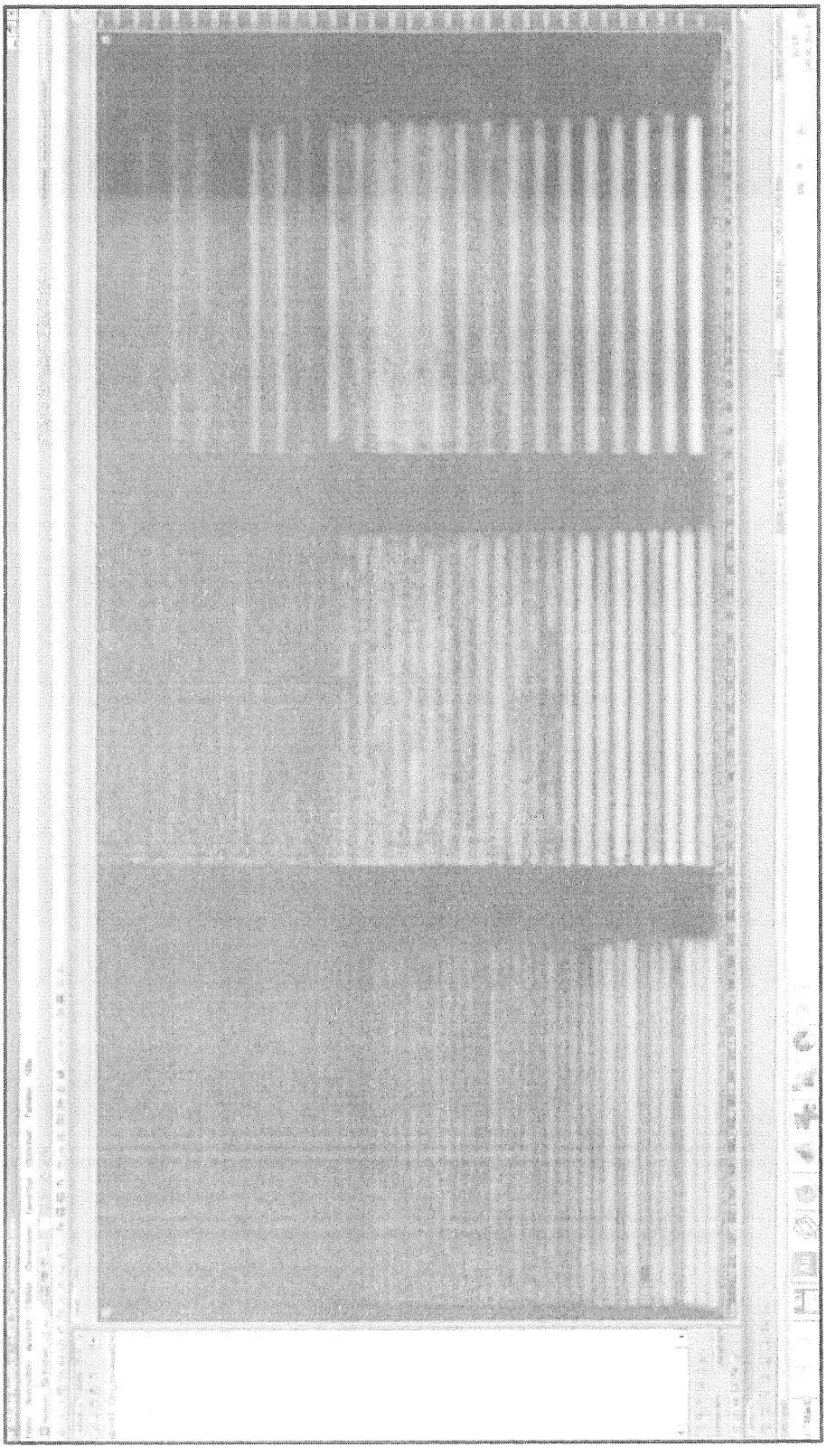

Fig. 10b

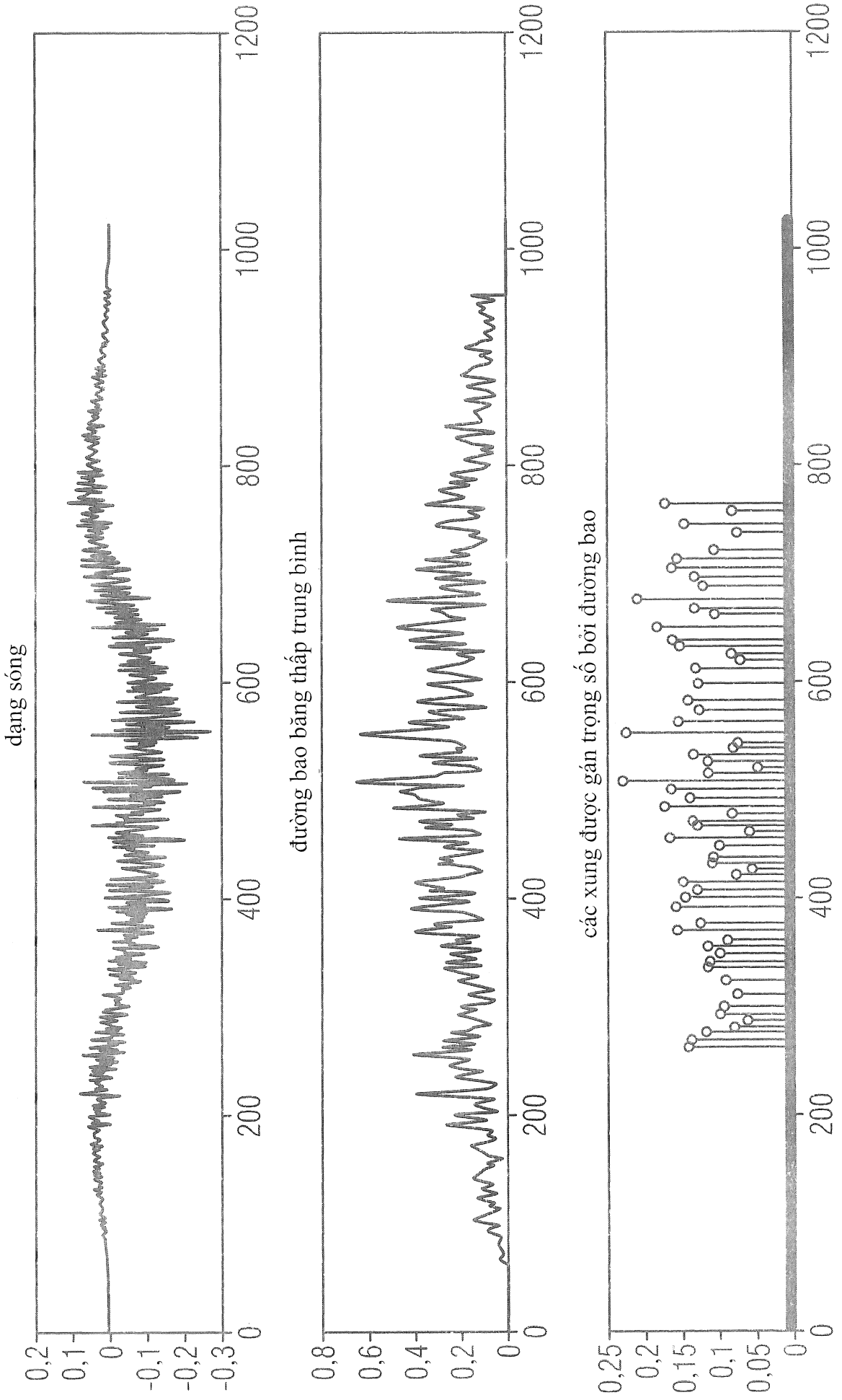
dạng sóng
đường bao băng thấp trung bình
các xung được gán trọng số bởi đường bao
0,2
0,1
0
-0,1
-0,2
-0,3
0,8
0,6
0,4
0,2
0,25
0,15
0,05
0
200
400
600
800
1000
1200

Fig. 11a

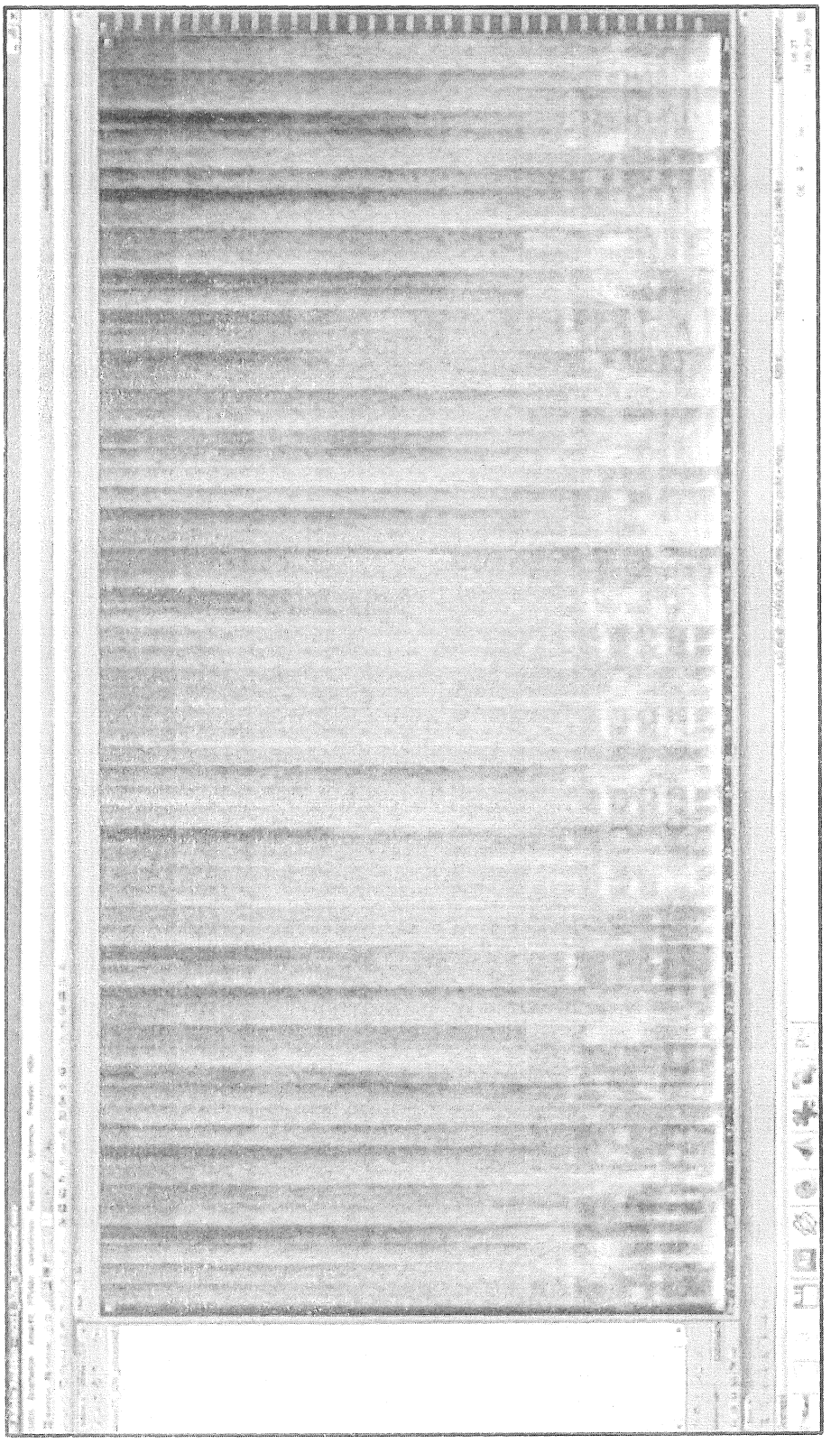

Fig. 11b

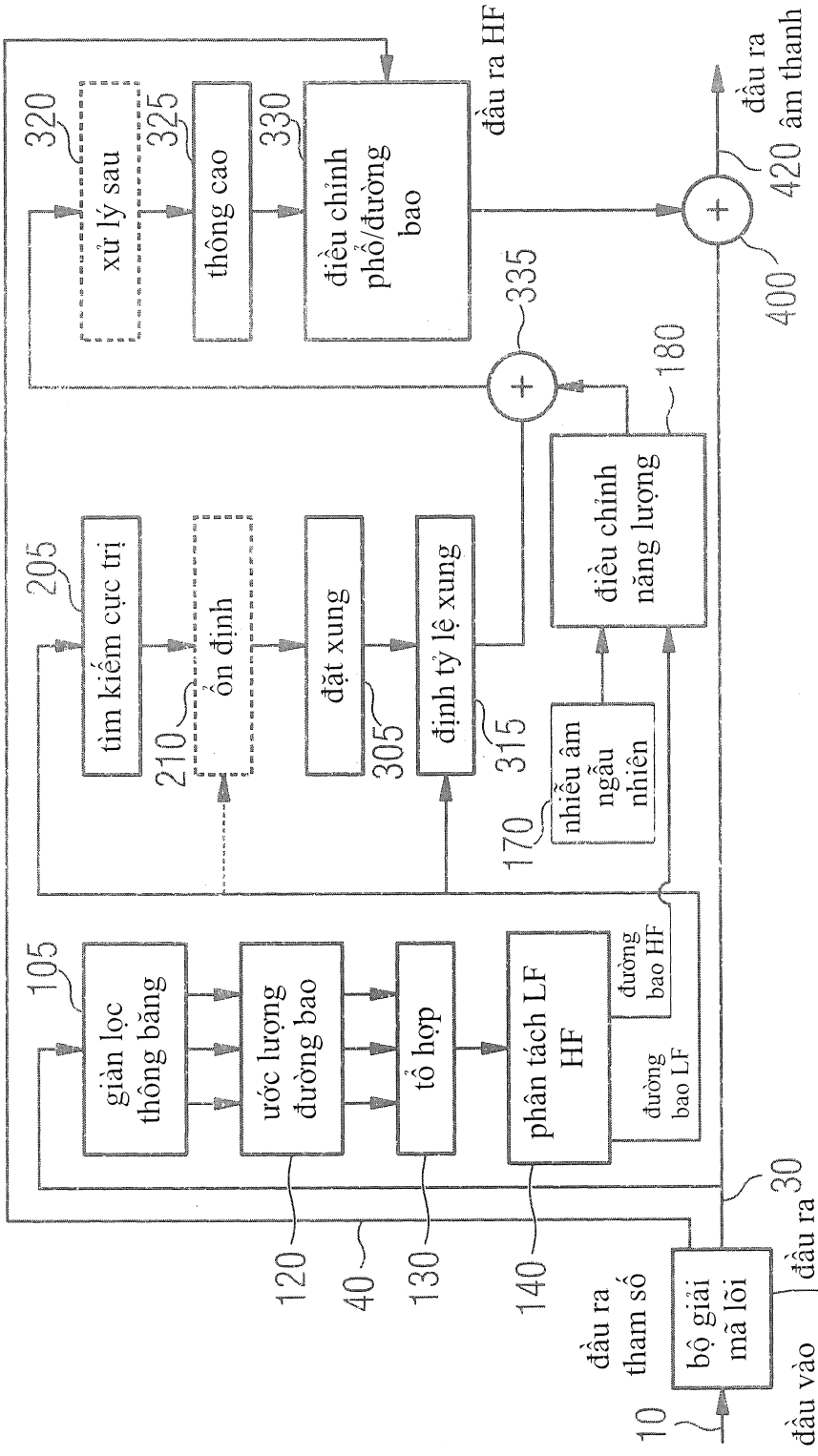
đầu vào dòng bit
đầu ra tham số
10
bộ giải mã lỗi
đầu ra LF
20
30
120
40
130
140
105
giàn lọc thông băng
ước lượng đường bao
tổ hợp
phân tách LF HF
đường bao LF
đường bao HF
205
tìm kiếm cực trị
210
ổn định
305
đặt xung
315
định tỷ lệ xung
170
nhiễu âm ngẫu nhiên
180
điều chỉnh năng lượng
320
xử lý sau
325
thông cao
330
điều chỉnh phổ/đường bao
đầu ra HF
335
400
420
đầu ra âm thanh

Fig. 12