



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

(19) Cộng hòa xã hội chủ nghĩa Việt Nam (VN) (11)
CỤC SỞ HỮU TRÍ TUỆ



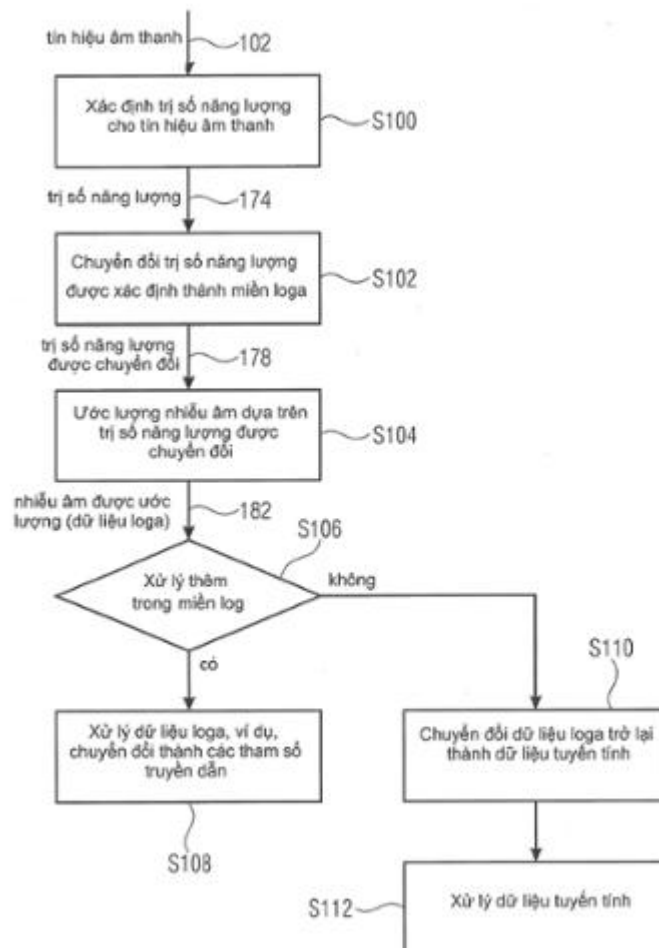
1-0029567

(51)⁷ **G10L 25/03**; G10L 25/21; G10L 19/012; (13) **B**
G10L 21/0216

-
- (21) 1-2017-00625 (22) 21/07/2015
(86) PCT/EP2015/066657 21/07/2015 (87) WO 2016/016051 04/02/2016
(30) 14178779.6 28/07/2014 EP
(45) 25/09/2021 402 (43) 25/05/2017 350A
(73) FRAUNHOFER-GESELLSCHAFT ZUR FOERDERUNG DER ANGEWANDTEN
FORSCHUNG E.V. (DE)
Hansastraße 27c, 80686 Muenchen, Germany
(72) SCHUBERT, Benjamin (DE); JANDER, Manuel (DE); LOMBARD, Anthony (FR);
DIETZ, Martin (DE); MULTRUS, Markus (DE).
(74) CÔNG TY LUẬT TRÁCH NHIỆM HỮU HẠN AMBYS HÀ NỘI (AMBYS
HANOI)
-

(54) PHƯƠNG PHÁP ƯỚC LƯỢNG NHIỀU ÂM TRONG TÍN HIỆU ÂM THANH,
THIẾT BỊ ƯỚC LƯỢNG NHIỀU ÂM, THIẾT BỊ MÃ HÓA ÂM THANH, THIẾT
BỊ GIẢI MÃ ÂM THANH VÀ HỆ THỐNG TRUYỀN CÁC TÍN HIỆU ÂM
THANH

(57) Sáng chế đề cập đến phương pháp ước lượng nhiều âm trong tín hiệu âm thanh, thiết bị ước lượng nhiều âm, thiết bị mã hóa âm thanh, thiết bị giải mã âm thanh, và hệ thống truyền các tín hiệu âm thanh. Phương pháp được mô tả mà ước lượng nhiều âm trong tín hiệu âm thanh (102). Trị số năng lượng (174) cho tín hiệu âm thanh (102) được ước lượng (S100) và được chuyển đổi (S102) thành miền loga. Mức nhiều âm cho tín hiệu âm thanh (102) được ước lượng (S104) dựa trên trị số năng lượng được chuyển đổi.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến lĩnh vực xử lý các tín hiệu âm thanh, cụ thể hơn là đề cập đến phương thức ước lượng nhiễu âm trong tín hiệu âm thanh, ví dụ trong tín hiệu âm thanh sẽ được mã hóa hoặc trong tín hiệu âm thanh mà đã được giải mã. Các phương án mô tả phương pháp ước lượng nhiễu âm trong tín hiệu âm thanh, bộ ước lượng nhiễu âm, bộ mã hóa âm thanh, bộ giải mã âm thanh và hệ thống truyền các tín hiệu âm thanh.

Tình trạng kỹ thuật của sáng chế

Trong lĩnh vực xử lý các tín hiệu âm thanh, ví dụ để mã hóa các tín hiệu âm thanh hoặc để xử lý các tín hiệu âm thanh được giải mã, có các trạng thái trong đó được mong muốn để ước lượng nhiễu âm. Ví dụ, PCT/EP2012/077525 và PCT/EP2012/077527, được kết hợp ở đây bởi sự tham chiếu, mô tả việc sử dụng bộ ước lượng nhiễu âm, ví dụ bộ ước lượng nhiễu âm thống kê cực tiểu, để ước lượng phổ của nhiễu âm nền trong miền tần số. Tín hiệu mã được cấp vào trong thuật toán đã được biến đổi theo khối thành miền tần số, ví dụ bằng biến đổi Fourier nhanh (Fast Fourier transformation - FFT) hoặc bất kỳ giàn lọc phù hợp khác bất kỳ. Việc tạo khung thường giống với việc tạo khung của bộ mã hóa-giải mã, tức là, các biến đổi đã có mặt trong bộ mã hóa-giải mã có thể được sử dụng lại, ví dụ trong bộ mã hóa dịch vụ thoại nâng cao ((Enhanced Voice Services - EVS), FFT được sử dụng cho quá trình xử lý trước. Với mục đích ước lượng nhiễu âm, phổ công suất của FFT được tính toán. Phổ được nhóm thành các băng được thúc đẩy theo cảm quan và các ngăn phổ công suất trong các băng được tích lũy để tạo thành trị số năng lượng trên mỗi băng. Cuối cùng, các trị số năng lượng đạt được bởi phương thức này mà cũng thường được sử dụng để xử lý theo cảm quan tín hiệu âm thanh. Từng băng có thuật toán ước lượng nhiễu âm riêng của nó, tức là, trong từng khung, trị số năng lượng của khung đó được xử lý bằng cách sử dụng thuật toán ước lượng nhiễu âm mà phân tích tín hiệu trên thời gian và cho mức nhiễu âm được ước lượng cho từng băng tại khung đã cho bất kỳ.

Độ phân giải mẫu sử dụng cho tiếng nói chất lượng cao và các tín hiệu âm thanh có thể là 16 bit, tức là tín hiệu có tỷ số tín hiệu trên nhiễu (signal-to-noise-ratio – SNR)

là 96dB. Tính toán phổ công suất có nghĩa là biến đổi tín hiệu thành miền tần số và tính toán bình phương của từng ngăn tần số. Do hàm bình phương, điều này yêu cầu dải động bằng 32 bit. Tổng hợp một số ngăn phổ công suất vào trong các băng yêu cầu khoảng trống phía trên bổ sung cho dải động vì sự phân bố năng lượng trong băng không được biết đến thực sự. Kết quả là, dải động có nhiều hơn 32 bit, thường khoảng 40 bit, cần được hỗ trợ để chạy bộ ước lượng nhiễu âm trên bộ xử lý.

Trong thiết bị xử lý các tín hiệu âm thanh mà hoạt động trên cơ sở năng lượng được nhận từ bộ phận lưu trữ năng lượng, như pin, ví dụ các thiết bị có thể mang theo như điện thoại di động để bảo toàn năng lượng, việc xử lý hiệu quả công suất của các tín hiệu âm thanh là cần thiết cho tuổi thọ của pin. Theo các phương thức đã biết, việc xử lý các tín hiệu âm thanh được thực hiện bởi các bộ xử lý điểm cố định mà, điển hình là, hỗ trợ việc xử lý dữ liệu trong định dạng điểm cố định 16 hoặc 32 bit. Độ phức tạp thấp nhất cho việc xử lý đạt được bằng cách xử lý dữ liệu 16 bit, trong khi xử lý dữ liệu 32 bit đã yêu cầu một số phần bổ sung. Việc xử lý dữ liệu với dải động 40 bit yêu cầu phân tách dữ liệu thành hai, cụ thể là phần định trị và số mũ, cả hai phần phải được giải quyết khi cải biên dữ liệu mà, lần lượt, thu được độ phức tạp tính toán thậm chí cao hơn và yêu cầu lưu trữ thậm chí cao hơn.

Bản chất kỹ thuật của sáng chế

Bắt đầu từ tình trạng kỹ thuật được thảo luận ở trên, mục đích của sáng chế là cung cấp phương thức để ước lượng nhiễu âm trong tín hiệu âm thanh theo cách hiệu quả sử dụng bộ xử lý điểm cố định tránh sự bổ sung tính toán không cần thiết.

Mục đích của sáng chế đạt được nhờ đối tượng như được xác định trong các điểm yêu cầu bảo hộ độc lập.

Sáng chế cung cấp phương pháp ước lượng nhiễu âm trong tín hiệu âm thanh, phương pháp bao gồm xác định trị số năng lượng cho tín hiệu âm thanh, chuyển đổi trị số năng lượng thành miền loga, và ước lượng mức nhiễu âm cho tín hiệu âm thanh dựa trên trị số năng lượng được chuyển đổi.

Sáng chế cung cấp bộ ước lượng nhiễu âm, bao gồm bộ phát hiện được tạo cấu hình để xác định trị số năng lượng cho tín hiệu âm thanh, bộ chuyển đổi được tạo cấu hình để chuyển đổi trị số năng lượng thành miền loga, và bộ ước lượng được tạo cấu hình để ước lượng mức nhiễu âm cho tín hiệu âm thanh dựa trên trị số năng lượng được chuyển đổi.

Sáng chế cung cấp bộ ước lượng nhiễu âm được tạo cấu hình để hoạt động theo phương pháp của sáng chế.

Theo các phương án, miền loga bao gồm miền $\log_2$.

Theo các phương án, việc ước lượng mức nhiễu âm bao gồm thực hiện thuật toán ước lượng nhiễu âm được xác định trước trên cơ sở trị số năng lượng được chuyển đổi trực tiếp trong miền loga. Sự ước lượng nhiễu âm có thể được thực hiện dựa trên thuật toán thống kê cực tiểu được mô tả bởi R. Martin, “Noise Power Spectral Density Estimation Based on Optimal Smoothing and Minimum Statistics”, 2001. Trong các phương án khác, các thuật toán ước lượng nhiễu âm khác có thể được sử dụng, như bộ ước lượng nhiễu âm dựa trên MMSE bởi T. Gerkmann và R. C. Hendriks, “Unbiased MMSE-based noise power estimation with low complexity and low tracking delay”, 2012, hoặc thuật toán được mô tả bởi L. Lin, W. Holmes, và E. Ambikairajah, “Adaptive noise estimation algorithm for speech enhancement”, 2003.

Theo các phương án, việc xác định trị số năng lượng bao gồm thu được phổ công suất của tín hiệu âm thanh bằng cách biến đổi tín hiệu âm thanh thành miền tần số, nhóm phổ công suất thành các băng được thúc đẩy theo cảm quan, và tích lũy các ngân phổ công suất trong băng để tạo ra trị số năng lượng cho từng băng, trong đó trị số năng lượng cho từng băng được chuyển đổi thành miền loga, và trong đó mức nhiễu âm được ước lượng cho từng băng dựa trên trị số năng lượng được chuyển đổi tương ứng.

Theo các phương án, tín hiệu âm thanh bao gồm nhiều khung, và đối với từng khung, trị số năng lượng được xác định và được chuyển đổi thành miền loga, và mức nhiễu âm được ước lượng cho từng băng dựa trên trị số năng lượng được chuyển đổi.

Theo các phương án, trị số năng lượng được chuyển đổi thành miền loga như sau:

$$E_{n_{log}} = \frac{\lfloor (\log_2(1 + E_{n_{lin}})) \cdot 2^N \rfloor}{2^N}$$

$\lfloor x \rfloor$ là số nguyên lớn nhất không vượt quá (x),

$E_{n_{log}}$ trị số năng lượng của băng n trong miền $\log_2$,

$E_{n_{lin}}$ trị số năng lượng của băng n trong miền tuyến tính,

N độ phân giải/độ chính xác.

Theo các phương án, việc ước lượng mức nhiễu âm dựa trên trị số năng lượng

được chuyển đổi mang lại dữ liệu loga, và phương pháp còn bao gồm sử dụng trực tiếp dữ liệu loga cho việc xử lý thêm nữa, hoặc chuyển đổi dữ liệu loga trở lại miền tuyến tính cho việc xử lý thêm nữa.

Theo các phương án, dữ liệu loga được chuyển đổi trực tiếp thành dữ liệu truyền, trong trường hợp việc truyền được thực hiện trong miền loga, và chuyển đổi dữ liệu loga trực tiếp thành dữ liệu truyền sử dụng hàm dịch chuyển cùng với bảng dò tìm hoặc phép gần đúng, ví dụ, $E_{n_lin} = 2^{(E_{n_log}-1)}$.

Sáng chế cung cấp sản phẩm chương trình máy tính không chuyển tiếp bao gồm vật ghi có thể đọc được bằng máy tính lưu trữ các chỉ dẫn mà, khi được chạy trên máy tính, thực hiện phương pháp theo sáng chế.

Sáng chế cung cấp bộ mã hóa âm thanh, bao gồm bộ ước lượng nhiễu âm theo sáng chế.

Sáng chế cung cấp bộ giải mã âm thanh, bao gồm bộ ước lượng nhiễu âm theo sáng chế.

Sáng chế cung cấp hệ thống truyền các tín hiệu âm thanh, hệ thống bao gồm bộ mã hóa âm thanh được tạo cấu hình để tạo ra tín hiệu âm thanh được mã hóa dựa trên tín hiệu âm thanh được nhận, và bộ giải mã âm thanh được tạo cấu hình để nhận tín hiệu âm thanh được mã hóa, để giải mã tín hiệu âm thanh được mã hóa, và để xuất ra tín hiệu âm thanh được giải mã, trong đó ít nhất một trong số bộ mã hóa âm thanh và bộ giải mã âm thanh bao gồm bộ ước lượng nhiễu âm theo sáng chế.

Sáng chế được dựa trên các tìm kiếm của các tác giả mà, trái lại với các phương thức thông thường mà trong đó thuật toán ước lượng nhiễu âm được chạy trên dữ liệu năng lượng tuyến tính, với mục đích ước lượng các mức nhiễu âm trong vật liệu âm thanh/tiếng nói, nó có khả năng để chạy thuật toán cũng dựa trên dữ liệu đầu vào loga. Đối với sự ước lượng nhiễu âm, yêu cầu về độ chính xác dữ liệu không phải là rất cao, ví dụ, khi sử dụng các trị số được ước lượng cho sự tạo ra nhiễu âm nhân tạo như được mô tả trong PCT/EP2012/077525 hoặc PCT/EP2012/077527, cả hai được kết hợp ở đây bởi sự tham chiếu, các tác giả sáng chế đã tìm ra rằng nó là đủ để ước lượng mức nhiễu âm chính xác một cách đại khái trên mỗi khung, tức là, liệu mức nhiễu âm được ước lượng để, ví dụ, cao hơn 0,1dB hay không sẽ là không đáng để chú ý trong tín hiệu cuối cùng. Do đó, trong khi 40 bit có thể được cần để che phủ dải động của dữ liệu, độ chính xác dữ liệu cho các tín hiệu mức trung bình/cao, trong các phương thức thông

thường, cao hơn nhiều mức cần thiết thực tế. Trên cơ sở các tìm kiếm này, theo các phương án, yếu tố chính của sáng chế là để chuyển đổi trị số năng lượng trên mỗi khung thành miền loga, tốt hơn là miền $\log_2$, và để thực hiện sự ước lượng nhiễu âm, ví dụ trên cơ sở thuật toán thống kê cực tiểu hoặc thuật toán phù hợp khác bất kỳ, một cách trực tiếp trong miền loga mà cho phép biểu thị các trị số năng lượng trong 16 bit mà, lần lượt, cho phép xử lý hiệu quả hơn, ví dụ sử dụng bộ xử lý điểm cố định.

Mô tả vắn tắt các hình vẽ

Trong phần sau đây, các phương án của sáng chế sẽ được mô tả dựa vào các hình vẽ đi kèm, trong đó:

Fig.1 thể hiện sơ đồ khối được đơn giản hóa của hệ thống truyền các tín hiệu âm thanh thực hiện phương thức theo sáng chế để ước lượng nhiễu âm trong tín hiệu âm thanh sẽ được mã hóa hoặc trong tín hiệu âm thanh được giải mã,

Fig.2 thể hiện sơ đồ khối được đơn giản hóa của bộ ước lượng nhiễu âm theo phương án mà có thể được sử dụng trong bộ mã hóa tín hiệu âm thanh và/hoặc bộ giải mã tín hiệu âm thanh, và

Fig.3 thể hiện sơ đồ khối mô tả phương thức theo sáng chế để ước lượng nhiễu âm trong tín hiệu âm thanh theo phương án.

Mô tả chi tiết sáng chế

Trong phần sau đây, các phương án của sáng chế sẽ được mô tả chi tiết hơn và được lưu ý là trong các hình vẽ đi kèm, các phần tử có chức năng giống nhau hoặc tương tự sẽ được ký hiệu bởi cùng ký hiệu tham chiếu.

Fig.1 thể hiện sơ đồ khối được đơn giản hóa của hệ thống truyền các tín hiệu âm thanh thực hiện phương thức theo sáng chế tại phía bộ mã hóa và/hoặc tại phía bộ giải mã. Hệ thống của Fig.1 bao gồm bộ mã hóa 100 nhận tín hiệu âm thanh 104 tại đầu vào 102. Bộ mã hóa bao gồm bộ xử lý mã hóa 106 nhận tín hiệu âm thanh 104 và tạo ra tín hiệu âm thanh được mã hóa mà được cung cấp tại đầu vào 108 của bộ mã hóa. Bộ xử lý mã hóa có thể được lập trình hoặc được xây dựng để xử lý các khung âm thanh liên tiếp của tín hiệu âm thanh và để thực hiện phương thức theo sáng chế để ước lượng nhiễu âm trong tín hiệu âm thanh 104 sẽ được mã hóa. Trong các phương án khác, bộ mã hóa không cần thiết là phần của hệ thống truyền, tuy nhiên, nó có thể là thiết bị độc lập tạo ra các tín hiệu âm thanh được mã hóa hoặc nó có thể là phần của bộ truyền tín hiệu âm thanh. Theo phương án, bộ mã hóa 100 có thể bao gồm ăngten 110

để cho phép sự truyền không dây của tín hiệu âm thanh, như được biểu thị tại 112. Trong các phương án khác, bộ mã hóa 100 có thể xuất ra tín hiệu âm thanh được mã hóa được cung cấp tại đầu ra 108 sử dụng đường kết nối có dây, như ví dụ được biểu thị tại ký hiệu tham chiếu 114.

Hệ thống của Fig.1 còn bao gồm bộ giải mã 150 có đầu vào 152 nhận tín hiệu âm thanh được mã hóa sẽ được xử lý bởi bộ giải mã 150, ví dụ thông qua đường dây 114 hoặc thông qua ăngten 154. Bộ giải mã 150 bao gồm bộ xử lý giải mã 156 hoạt động trên tín hiệu được mã hóa và cung cấp tín hiệu âm thanh được giải mã 158 tại đầu ra 160. Bộ xử lý giải mã có thể được lập trình hoặc được xây dựng để xử lý cho việc thực hiện phương thức theo sáng chế để ước lượng nhiễu âm trong tín hiệu âm thanh được giải mã 104. Trong các phương án khác, bộ giải mã không cần thiết là phần của hệ thống truyền, hơn nữa, nó có thể là thiết bị độc lập để giải mã các tín hiệu âm thanh được mã hóa hoặc nó có thể là phần của bộ nhận tín hiệu âm thanh.

Fig.2 thể hiện sơ đồ khối được đơn giản hóa của bộ ước lượng nhiễu âm 170 theo phương án. Bộ ước lượng nhiễu âm 170 có thể được sử dụng trong bộ mã hóa tín hiệu âm thanh và/hoặc bộ giải mã tín hiệu âm thanh được thể hiện trên Fig.1. Bộ ước lượng nhiễu âm 170 bao gồm bộ phát hiện 172 để xác định trị số năng lượng 174 cho tín hiệu âm thanh 102, bộ chuyển đổi 176 để chuyển đổi trị số năng lượng 174 thành miền loga (xem trị số năng lượng được chuyển đổi 178), và bộ ước lượng 180 để ước lượng mức nhiễu âm 182 cho tín hiệu âm thanh 102 dựa trên trị số năng lượng được chuyển đổi 178. Bộ ước lượng 170 có thể được thực hiện bởi bộ xử lý chung hoặc bởi nhiều bộ xử lý được lập trình hoặc được xây dựng để thực hiện chức năng của bộ phát hiện 172, bộ chuyển đổi 176 và bộ ước lượng 180.

Trong phần sau đây, các phương án của phương thức theo sáng chế có thể được thực hiện trong ít nhất một trong số bộ xử lý mã hóa 106 và bộ xử lý giải mã 156 của Fig.1, hoặc bởi bộ ước lượng 170 của Fig.2 sẽ được mô tả chi tiết hơn.

Fig.3 thể hiện lưu đồ của phương thức theo sáng chế để ước lượng nhiễu âm trong tín hiệu âm thanh. Tín hiệu âm thanh được nhận và, trong bước thứ nhất S100, trị số năng lượng 174 cho tín hiệu âm thanh được xác định, mà sau đó, trong bước S102, được chuyển đổi thành miền loga. Trên cơ sở trị số năng lượng được chuyển đổi 178, trong bước S104, nhiễu âm được ước lượng. Theo các phương án, trong bước S106, nó được xác định là liệu việc xử lý thêm dữ liệu nhiễu âm được ước lượng, mà

được biểu diễn bởi dữ liệu loga 182, có nên ở trong miền loga hay không. Trong trường hợp xử lý thêm trong miền loga được mong muốn (có trong bước S106), dữ liệu loga biểu diễn nhiều âm được ước lượng được xử lý trong bước S108, ví dụ dữ liệu loga được chuyển đổi thành các tham số truyền trong trường hợp sự truyền cũng xảy ra trong miền loga. Nếu không thì (không trong bước S106), dữ liệu loga 182, được chuyển đổi trở lại thành dữ liệu tuyến tính trong bước S110, và dữ liệu tuyến tính được xử lý trong bước S112.

Theo các phương án, trong bước S100, việc xác định trị số năng lượng cho tín hiệu âm thanh có thể được thực hiện như trong các phương thức thông thường. Phổ công suất của FFT, mà đã được áp dụng cho tín hiệu âm thanh, được tính toán và nhóm thành các băng được thúc đẩy theo cảm quan. Các ngăn phổ công suất bên trong băng được tích lũy để tạo ra trị số năng lượng trên mỗi băng sao cho tập hợp các trị số năng lượng thu được. Trong các phương án khác, phổ công suất có thể được tính dựa trên sự biến đổi phổ phù hợp bất kỳ, như biến đổi cosin rời rạc cải biên (Modified Discrete Cosine Transform - MDCT), giàn lọc độ trễ thấp phức tạp (Complex Low-Delay Filterbank - CLDFB), hoặc tổ hợp của một số biến đổi bao quanh các phần phổ khác nhau. Trong bước S100 trị số năng lượng 174 cho từng băng được xác định, và trị số năng lượng 174 cho từng băng được chuyển đổi thành miền loga trong bước S102, theo các phương án, thành miền log2. Các năng lượng băng có thể được chuyển đổi thành miền log2 như sau:

$$E_{n_{log}} = \frac{\lfloor (\log_2(1 + E_{n_{lin}})) \cdot 2^N \rfloor}{2^N}$$

$\lfloor x \rfloor$	là số nguyên lớn nhất không vượt quá (x),
$E_{n_{log}}$	trị số năng lượng của băng n trong miền log2,
$E_{n_{lin}}$	trị số năng lượng của băng n trong miền tuyến tính,
N	độ phân giải/độ chính xác.

Theo các phương án, sự chuyển đổi thành miền log2 được thực hiện mà có lợi ở chỗ hàm (int)log2 có thể thường được tính toán rất nhanh, ví dụ trong một chu kỳ, trên các bộ xử lý điểm cố định sử dụng hàm “chuẩn” mà xác định số lượng các số không dẫn đầu trong số lượng điểm cố định. Đôi khi độ chính xác cao hơn (int)log2 được cần đến, mà được biểu thị trong công thức trên bởi hằng số N. Độ chính xác cao hơn một chút này có thể đạt được với bảng dò tìm đơn giản có các bit có nghĩa nhất sau chỉ dẫn

chuẩn và phép gần đúng, mà là các phương thức chung để đạt được sự tính toán thuật toán phức tạp thấp khi độ chính xác thấp hơn có thể chấp nhận được. Trong công thức trên, hằng số "1" bên trong hàm $\log_2$ được bổ sung để đảm bảo rằng các năng lượng được chuyển đổi vẫn dương. Theo các phương án, điều này có thể quan trọng trong trường hợp bộ ước lượng nhiễu âm dựa trên mô hình thống kê của năng lượng nhiễu âm, như thực hiện ước lượng nhiễu âm trên các trị số âm có thể trái với mô hình này và có thể dẫn đến sự chuyển vận không mong muốn của bộ ước lượng.

Theo phương án, trong công thức trên, N được thiết lập đến 6, mà tương đương với $2^6 = 64$ bit của dải động. Dải động này lớn hơn dải động 40 bit được mô tả ở trên và, do đó, là đủ. Để xử lý dữ liệu, mục đích là để sử dụng dữ liệu 16 bit, mà cho phép 9 bit cho phần định trị và một bit cho phần dấu. Định dạng này được ký hiệu chung là định dạng "6Q9". Theo các khác, vì chỉ các trị số dương có thể được xem xét, bit dấu có thể được tránh khỏi và được sử dụng cho phần định trị để lại tổng 10 bit cho phần định trị, mà được đề cập đến như định dạng "6Q10".

Phần mô tả chi tiết về thuật toán thống kê cực tiểu có thể được tìm thấy trong tài liệu của R. Martin, "Noise Power Spectral Density Estimation Based on Optimal Smoothing and Minimum Statistics", 2001. Về cơ bản nó bao gồm việc theo dõi cực tiểu của phổ công suất được làm nhẵn trên cửa sổ theo thời gian trượt của độ dài đã cho đối với từng băng phổ, thường trên vài giây. Thuật toán cũng bao gồm sự bù xiên để cải thiện độ chính xác của sự ước lượng nhiễu âm. Hơn nữa, để cải thiện sự theo dõi nhiễu âm biến thiên theo thời gian, cực tiểu cục bộ được tính toán trên cửa sổ theo thời gian rất ngắn có thể được sử dụng thay vì cực tiểu ban đầu, với điều kiện là nó mang lại sự tăng vừa phải của năng lượng nhiễu âm được ước lượng. Lượng tăng được cho phép được xác định trong tài liệu của R. Martin, "Noise Power Spectral Density Estimation Based on Optimal Smoothing and Minimum Statistics", 2001 bởi tham số nhiễu âm_độ dốc_cực đại (parameter noise_slope_max). Theo phương án, thuật toán ước lượng nhiễu âm thống kê cực tiểu được sử dụng mà, một cách thông thường, chạy trên dữ liệu năng lượng tuyến tính. Tuy nhiên, theo các nghiên cứu của các tác giả, với mục đích ước lượng mức nhiễu âm trong vật liệu âm thanh hoặc vật liệu tiếng nói, thuật toán có thể được nạp với dữ liệu đầu vào loga để thay thế. Trong khi việc xử lý tín hiệu tự nó vẫn không được cải biên, chỉ cực tiểu của sự điều hướng lại được yêu cầu, mà ở bao gồm tham số nhiễu âm_độ dốc_cực đại để đối phó với dải động giảm

của dữ liệu loga được so sánh với dữ liệu tuyến tính. Đến nay, nó được giả định là thuật toán thống kê cực tiểu, hoặc các kỹ thuật ước lượng nhiễu âm phù hợp khác, cần chạy trên dữ liệu tuyến tính, tức là, dữ liệu mà trong thực tế là sự biểu diễn loga đã được giả định không phù hợp. Trái lại với giả định thông thường này, các tác giả sáng chế đã tìm ra rằng sự ước lượng nhiễu âm quả thực là có thể được chạy trên cơ sở dữ liệu loga mà cho phép sử dụng dữ liệu đầu vào mà chỉ được biểu diễn trên 16 bit mà, kết quả là, cung cấp cho độ phức tạp nhỏ hơn nhiều trong sự thực hiện điểm cố định như hầu hết các thao tác có thể được thực hiện trong 16 bit và chỉ một số phần của thuật toán vẫn yêu cầu 32 bit. Trong thuật toán thống kê cực tiểu, ví dụ, sự bù xiên được dựa trên phương sai của công suất đầu vào, do đó các thống kê bậc bốn mà thường vẫn yêu cầu sự biểu diễn 32 bit.

Như đã được mô tả ở trên về Fig.3, kết quả của quy trình ước lượng nhiễu âm có thể còn được xử lý theo các cách khác nhau. Theo các phương án, cách thứ nhất là để sử dụng trực tiếp dữ liệu loga 182, như được thể hiện trong bước S108, ví dụ bằng cách chuyển đổi trực tiếp dữ liệu loga 182 thành các tham số truyền nếu các tham số này cũng được truyền trong miền loga, điều này thường xảy ra. Cách thứ hai là để xử lý dữ liệu loga 182 sao cho nó được chuyển đổi ngược lại thành miền tuyến tính để xử lý thêm nữa, ví dụ sử dụng các hàm dịch chuyển mà thường là rất nhanh và thường chỉ yêu cầu một chu kỳ trên bộ xử lý, cùng với sự dò tìm bảng hoặc bằng cách sử dụng phép gần đúng, ví dụ:

$$E_{n_lin} = 2^{(E_{n_log}-1)}$$

Trong phần sau đây, ví dụ chi tiết để thực hiện phương thức theo sáng chế để ước lượng nhiễu âm trên cơ sở dữ liệu loga sẽ được mô tả với sự tham chiếu đến bộ mã hóa, tuy nhiên, như được tóm tắt ở trên, phương thức theo sáng chế cũng có thể được áp dụng cho các tín hiệu mà đã được giải mã trong bộ giải mã, như nó là ví dụ được mô tả trong PCT/EP2012/077525 hoặc PCT/EP2012/077527, cả hai được kết hợp ở đây bởi sự tham chiếu. Phương án sau đây mô tả việc thực hiện phương thức theo sáng chế để ước lượng nhiễu âm trong tín hiệu âm thanh trong bộ mã hóa âm thanh, như bộ mã hóa 100 trên Fig.1. Cụ thể hơn, sự mô tả thuật toán xử lý tín hiệu của bộ mã hóa dịch vụ thoại nâng cao (bộ mã hóa EVS) để thực hiện phương thức theo sáng chế để ước lượng nhiễu âm trong tín hiệu âm thanh được nhận tại bộ mã hóa EVS sẽ được cho.

Các khối đầu vào của các mẫu âm thanh có độ dài 20 mili giây được giả định trong định dạng biến điệu mã xung (Pulse Code Modulation - PCM) đồng đều 16 bit. Bốn tốc độ lấy mẫu cũng được giả định, ví dụ, 8000, 16000, 32000 và 48000 mẫu/giây và các tốc độ bit cho dòng bit được mã hóa có thể là 5,9, 7,2, 8,0, 9,6, 13,2, 16,4, 24,4, 32,0, 48,0, 64,0 hoặc 128,0 kbit/giây. Chế độ tương thích (bộ mã hóa-giải mã) thích ứng đa tốc độ băng rộng (Adaptive Multi Rate Wideband - AMR-WB) cũng có thể được cung cấp mà hoạt động tại các tốc độ bit cho dòng bit được mã hóa bằng 6,6, 8,85, 12,65, 14,85, 15,85, 18,25, 19,85, 23,05 hoặc 23,85 kbit/giây.

Với mục đích mô tả sau đây, các quy ước sau đây áp dụng cho các biểu thức toán học:

$\lfloor x \rfloor$ biểu thị số nguyên lớn nhất nhỏ hơn hoặc bằng x : $\lfloor 1,1 \rfloor = 1$, $\lfloor 1,0 \rfloor = 1$ và $\lfloor -1,1 \rfloor = -2$;

$\sum$ biểu thị phép tổng;

Trừ khi có quy định khác, $\log(x)$ ký hiệu loga tại cơ số 10 trong suốt phần mô tả sau đây.

Bộ mã hóa chấp nhận các tín hiệu băng đầy đủ (fullband - FB), băng siêu rộng (superwideband - SWB), băng rộng (wideband - WB) hoặc băng hẹp (narrow-band - NB) được lấy mẫu tại 48, 32, 16 hoặc 8 kHz. Một cách đơn giản, đầu ra bộ giải mã có thể là 48, 32, 16 hoặc 8 kHz, FB, SWB, WB hoặc NB. Tham số R (8, 16, 32 or 48) được sử dụng để biểu thị tốc độ lấy mẫu đầu vào tại bộ mã hóa hoặc tốc độ lấy mẫu đầu ra tại bộ giải mã.

Tín hiệu đầu vào được xử lý sử dụng các khung 20 mili giây. Độ trễ bộ mã hóa-giải mã phụ thuộc vào tốc độ lấy mẫu của đầu vào và đầu ra. Đối với đầu vào WB và đầu ra WB, độ trễ thuật toán toàn bộ là 42,875 mili giây. Nó bao gồm một khung 20 mili giây, độ trễ 1,875 mili giây của các bộ lọc lấy mẫu lại đầu vào và đầu ra, 10 mili giây cho sự kiểm tra trước bộ mã hóa, 1 mili giây của độ trễ bộ lọc sau, và 10 mili giây tại bộ giải mã để cho phép thao tác cộng chồng lấp của phép mã hóa biến đổi lớp cao hơn. Đối với đầu vào NB và đầu ra NB, các lớp cao hơn không được sử dụng, nhưng độ trễ bộ giải mã 10 mili giây được sử dụng để cải thiện hiệu suất bộ mã hóa-giải mã trong sự có mặt của sự xóa khung và cho các tín hiệu âm nhạc. Độ trễ thuật toán toàn bộ cho đầu vào NB và đầu ra NB là 43,875 mili giây – một khung

20 mili giây, 2 mili giây cho bộ lọc lấy mẫu lại đầu vào, 10 mili giây cho sự kiểm tra trước bộ mã hóa, 1,875 mili giây cho bộ lọc lấy mẫu lại đầu ra, và độ trễ 10 mili giây trong bộ giải mã. Nếu đầu ra bị giới hạn đến lớp 2, độ trễ bộ mã hóa-giải mã có thể được giảm 10 mili giây.

Chức năng chung của bộ mã hóa bao gồm các phân đoạn xử lý sau: xử lý chung, chế độ mã hóa CELP (Code-Excited Linear Prediction), chế độ mã hóa MDCT (Modified Discrete Cosine Transform), chuyển các chế độ mã hóa, thông tin phụ che giấu xóa khung, thao tác DTX/CNG (Discontinuous Transmission/Comfort Noise Generator), lựa chọn tương thích AMR-WB, và mã hóa nhận biết kênh.

Theo phương án của sáng chế, phương thức theo sáng chế được thực hiện trong phân đoạn thao tác DTX/CNG. Bộ mã hóa-giải mã được trạng bị với thuật toán phát hiện hoạt động của tín hiệu (signal activity detection - SAD) để phân loại từng khung đầu vào là hoạt động hay không hoạt động. Nó hỗ trợ hoạt động truyền gián đoạn (DTX) mà trong đó môđun tạo ra nhiễu âm nền nhân tạo miền tần số (frequency-domain comfort noise generation - FD-CNG) được sử dụng để làm gần đúng và cập nhật các thống kê của nhiễu âm nền tại tốc độ bit có thể thay đổi. Do đó, tốc độ truyền trong suốt các chu kỳ tín hiệu không hoạt động có thể thay đổi và phụ thuộc vào mức được ước lượng của nhiễu âm nền. Tuy nhiên, tốc độ cập nhật CNG cũng có thể được cố định bởi tham số dòng lệnh.

Để có thể tạo ra nhiễu âm nhân tạo tương tự với nhiễu âm nền đầu vào thực tế về mặt các đặc tính phổ-theo thời gian, FD-CNG tạo ra việc sử dụng của thuật toán ước lượng nhiễu âm để theo dõi năng lượng của nhiễu âm nền có mặt tại đầu vào bộ mã hóa. Các ước lượng nhiễu âm sau đó được truyền như các tham số trong dạng các khung bộ mô tả chèn im lặng (Silence Insertion Descriptor - SID) để cập nhật biên độ của các chuỗi ngẫu nhiên được tạo ra trong từng băng tại phía bộ giải mã trong suốt các pha không hoạt động.

Bộ ước lượng nhiễu âm FD-CNG dựa trên phương thức phân tích phổ lai. Các tần số thấp tương ứng với băng thông lời được che phủ bởi phép phân tích FFT độ phân giải cao, trong khi các tần số cao hơn còn lại được chụp bởi CLDFB mà thể hiện độ phân giải phổ thấp hơn đáng kể là 400Hz. Lưu ý là CLDFB cũng được sử dụng như công cụ lấy mẫu lại để lấy mẫu giảm tín hiệu đầu vào cho tốc độ lấy mẫu lời.

Tuy nhiên, kích thước của khung SID được giới hạn trong thực tế. Để giảm số lượng các tham số mô tả nhiều âm nền, các năng lượng đầu vào được tính trung bình giữa các nhóm của các băng phổ được gọi là các phân vùng trong đoạn tiếp theo.

1. Các năng lượng phân vùng phổ

Các năng lượng phân vùng được tính toán riêng rẽ cho các băng FFT và CLDFB. Các năng lượng $L_{SID}^{[FFT]}$ tương ứng với các phân vùng FFT và các năng lượng $L_{SID}^{[CLDFB]}$ tương ứng với các phân vùng CLDFB sau đó được nối chuỗi thành mảng đơn E_{FD-CNG} có kích thước $L_{SID} = L_{SID}^{[FFT]} + L_{SID}^{[CLDFB]}$ mà sẽ có vai trò như đầu vào đến bộ ước lượng nhiễu âm được mô tả dưới đây (xem “2. ước lượng nhiễu âm FD-CNG”).

1.1 Tính toán các năng lượng phân vùng FFT

Các năng lượng phân vùng cho các tần số bao quanh băng thông lõi thu được như

$$E_{FD-CNG}(i) = \frac{E_{CB}^{[0]}(i) + E_{CB}^{[1]}(i)}{2} H_{de-emph}(i) \quad i = 0, \dots, L_{SID}^{[FFT]} - 1$$

trong đó $E_{CB}^{[0]}(i)$ và $E_{CB}^{[1]}(i)$ là năng lượng trung bình trong băng tới hạn i lần lượt cho các cửa sổ phân tích thứ nhất và thứ hai. Số lượng của các phân vùng FFT $L_{SID}^{[FFT]}$ chụp được các dải băng thông lõi giữa 17 và 21, theo cấu hình được sử dụng (xem “1.3 các cấu hình bộ mã hóa FD-CNG”). Các trọng số phổ chỉnh giảm $H_{de-emph}(i)$ được sử dụng để bù cho bộ lọc thông cao và được xác định như

$$\{H_{de-emph}(0), \dots, H_{de-emph}(L_{SID}^{[FFT]} - 1)\} = \{9,7461, 9,5182, 9,0262, 8,3493, 7,5764, 6,7838, 5,8377, \\ 4,8502, 4,0346, 3,2788, 2,6283, 2,0920, 1,6304, 1,2850, \\ 1,0108, 0,7916, 0,6268, 0,5011, 0,4119, 0,3637\}.$$

1.2 Tính toán các năng lượng phân vùng CLDFB

Các năng lượng phân vùng cho các tần số trên băng thông lõi được tính như

$$E_{FD-CNG}(i) = \frac{1}{16} \frac{1}{8(A_{CLDFB})^2} \frac{\sum_{j=j_{min}(i)}^{j_{max}(i)} E_{CLDFB}(j)}{j_{max}(i) - j_{min}(i) + 1} \quad i = L_{SID}^{[FFT]}, \dots, L_{SID}^{[FFT]} + L_{SID}^{[CLDFB]} - 1$$

trong đó $j_{min}(i)$ và $j_{max}(i)$ lần lượt là các chỉ số của các băng CLDFB thứ nhất và cuối cùng trong phân vùng thứ i , $E_{CLDFB}(j)$ là năng lượng tổng của băng CLDFB thứ j , và A_{CLDFB} là thừa số tỷ lệ. Hằng số 16 đề cập đến số lượng các khe thời gian trong CLDFB. Số lượng của các phân vùng CLDFB L_{CLDFB} phụ thuộc vào cấu hình được sử dụng, như được mô tả dưới đây.

1.3 Các cấu hình bộ mã hóa FD-CNG

Bảng sau đây liệt kê số lượng các phân vùng và các đường biên phía trên của chúng cho các cấu hình FD-CNG khác nhau tại bộ mã hóa.

	Các tốc độ bit [kbps]	$L_{\text{SID}}^{[\text{FFT}]}$	$L_{\text{SID}}^{[\text{CLDFB}]}$	$f_{\text{max}}(i),$ $i = 0, \dots, L_{\text{SID}}^{[\text{FFT}]} - 1$ [Hz]	$f_{\text{max}}(i),$ $i = L_{\text{SID}}^{[\text{FFT}]} , \dots, L_{\text{SID}} - 1$ [Hz]
NB	•	17	0	100, 200, 300, 400, 500, 600, 750, 900, 1050, 1250, 1450, 1700, 2000, 2300, 2700, 3150, 3975	×
WB	≤ 8	20	0	100, 200, 300, 400, 500, 600, 750, 900, 1050, 1250, 1450, 1700, 2000, 2300, 2700, 3150, 3700, 4400, 5300, 6375	×
	$8 < \bullet \leq 13.2$	20	1	100, 200, 300, 400, 500, 600, 750, 900, 1050, 1250, 1450, 1700, 2000, 2300, 2700, 3150, 3700, 4400, 5300, 6375	8000
	> 13.2	21	0	100, 200, 300, 400, 500, 600, 750, 900, 1050, 1250, 1450, 1700, 2000, 2300, 2700, 3150, 3700, 4400, 5300, 6375, 7975	×
SW B/F B	≤ 13.2	20	4	100, 200, 300, 400, 500, 600, 750, 900, 1050, 1250, 1450, 1700, 2000, 2300, 2700, 3150, 3700, 4400, 5300, 6375	8000, 10000, 12000, 14000

	> 13.2	21	3	100, 200, 300, 400, 500, 600, 750, 900, 1050, 1250, 1450, 1700, 2000, 2300, 2700, 3150, 3700, 4400, 5300, 6375, 7975	10000, 12000, 16000
--	--------	----	---	--	---------------------

Bảng 1: Các cấu hình của ước lượng nhiễu âm FD-CNG tại bộ mã hóa

Đối với từng phân vùng $i = 0, \dots, L_{\text{SID}} - 1$, $f_{\text{max}}(i)$ tương ứng với tần số của băng cuối cùng trong phân vùng thứ i . Các chỉ số $j_{\text{min}}(i)$ và $j_{\text{max}}(i)$ của các băng thứ nhất và cuối cùng trong từng phân vùng phổ có thể được suy ra như hàm của cấu hình của lõi như sau:

$$j_{\text{max}}(i) = \begin{cases} f_{\text{max}}(i) \frac{\text{core_FFT_length}}{\text{core_sampling_rate}} & i = 0, \dots, L_{\text{SID}}^{\text{[FFT]}} - 1 \\ j_{\text{max}}(L_{\text{SID}}^{\text{[FFT]}} - 1) + \frac{2f_{\text{max}}(i) - \text{core_sampling_rate}}{800} & i = L_{\text{SID}}^{\text{[FFT]}}, \dots, L_{\text{SID}} - 1 \end{cases},$$

$$j_{\text{min}}(i) = \begin{cases} f_{\text{min}}(0) \frac{\text{core_sampling_rate}}{\text{core_FFT_length}} & i = 0 \\ j_{\text{max}}(i - 1) + 1 & i > 0 \end{cases},$$

trong đó $f_{\text{min}}(0) = 50\text{Hz}$ là tần số của băng thứ nhất trong phân vùng phổ. Do đó FD-CNG tạo ra nhiễu âm nhân tạo chỉ trên 50Hz.

2. Ước lượng nhiễu âm FD-CNG

FD-CNG dựa trên bộ ước lượng nhiễu âm để theo dõi năng lượng của nhiễu âm nền có mặt trong phổ đầu vào. Điều này chủ yếu dựa trên thuật toán thống kê cực tiểu được mô tả bởi R. Martin, “Noise Power Spectral Density Estimation Based on Optimal Smoothing and Minimum Statistics”, 2001. Tuy nhiên, để giảm dải động của các năng lượng đầu vào $\{E_{\text{FD-CNG}}(0), \dots, E_{\text{FD-CNG}}(L_{\text{SID}} - 1)\}$ và do đó tạo điều kiện cho sự thực hiện điểm cố định của thuật toán ước lượng nhiễu âm, biến đổi không tuyến tính được áp dụng trước khi ước lượng nhiễu âm (xem “2.1 Nén dải động cho các năng lượng đầu vào”). Phép biến đổi nghịch đảo sau đó được sử dụng trên các ước lượng nhiễu âm thu được để thu hồi dải động ban đầu (xem “2.3 Mở rộng dải động cho các năng lượng nhiễu âm được ước lượng”).

2.1 Nén dải động cho các năng lượng đầu vào

Các năng lượng đầu vào được xử lý bởi hàm không tuyến tính và được lượng tử hóa với độ phân giải 9 bit như sau:

$$E_{MS}(i) = \frac{\lfloor \log_2((1 + E_{FD-CNG}(i))2^9) \rfloor}{2^9} \quad i = 0, \dots, L_{SID}-1$$

2.2 Theo dõi nhiễu âm

Sự mô tả chi tiết của thuật toán thống kê cực tiểu có thể được tìm thấy trong tài liệu của R. Martin, “Noise Power Spectral Density Estimation Based on Optimal Smoothing and Minimum Statistics”, 2001. Về cơ bản nó bao gồm việc theo dõi cực tiểu của phổ công suất được làm nhẵn trên cửa sổ theo thời gian trượt của độ dài đã cho đối với từng băng phổ, thường trên vài giây. Thuật toán cũng bao gồm sự bù xiên để cải thiện độ chính xác của sự ước lượng nhiễu âm. Hơn nữa, để cải thiện sự theo dõi nhiễu âm biến thiên theo thời gian, cực tiểu cục bộ được tính toán trên cửa sổ theo thời gian rất ngắn có thể được sử dụng thay vì cực tiểu ban đầu, với điều kiện là nó mang lại sự tăng vừa phải của năng lượng nhiễu âm được ước lượng. Lượng tăng được cho phép được xác định trong tài liệu của R. Martin, “Noise Power Spectral Density Estimation Based on Optimal Smoothing and Minimum Statistics”, 2001 bởi tham số nhiễu âm_độ dốc_cực đại.

Các đầu vào chính của bộ theo dõi nhiễu âm là các ước lượng nhiễu âm $N_{MS}(i)$, $i = 0, \dots, L_{SID}-1$. Để thu được các chuyển tiếp nhẵn hơn trong nhiễu âm nhân tạo, bộ lọc đệ quy bậc một có thể được áp dụng, tức là $\bar{N}_{MS}(i) = 0,95 \bar{N}_{MS}(i) + 0,05 N_{MS}(i)$.

Hơn nữa, năng lượng đầu vào $E_{MS}(i)$ được tính trung bình trên 5 khung cuối cùng. Điều này được sử dụng để áp dụng giới hạn trên lên $\bar{N}_{MS}(i)$ trong từng phân vùng phổ.

2.3 Mở rộng dải động cho các năng lượng nhiễu âm được ước lượng

Các năng lượng nhiễu âm được ước lượng được xử lý bởi hàm không tuyến tính để bù cho sự nén dải động được mô tả ở trên:

$$N_{FD-CNG}(i) = 2^{\bar{N}_{MS}(i)-1} \quad i = 0, \dots, L_{SID}-1.$$

Theo sáng chế, phương thức được cải thiện để ước lượng nhiễu âm trong tín hiệu âm thanh được mô tả mà cho phép giảm độ phức tạp của bộ ước lượng nhiễu âm, đặc biệt là cho các tín hiệu âm thanh/tiếng nói mà được xử lý trên các bộ xử lý sử dụng số học điểm cố định. Phương thức theo sáng chế cho phép giảm dải động được sử dụng cho bộ ước lượng nhiễu âm cho việc xử lý tín hiệu âm thanh/tiếng nói, ví dụ, trong

môi trường được mô tả trong PCT/EP2012/077527, mà đề cập đến sự tạo ra nhiều âm nhân tạo với độ phân giải phổ-thời gian cao, hoặc trong PCT/EP2012/077527, mà đề cập đến sự bổ sung nhiều âm nhân tạo để mô hình hóa nhiều âm nền tại tốc độ bit thấp. Trong các ngữ cảnh được mô tả, bộ ước lượng nhiều âm được sử dụng hoạt động trên cơ sở thuật toán thống kê cực tiểu để tăng cường chất lượng nhiều âm nền hoặc cho sự tạo ra nhiều âm nhân tạo cho các tín hiệu tiếng nói ồn ào, ví dụ tiếng nói trong sự có mặt của nhiều âm nền mà là trạng thái rất phổ biến trong các cuộc gọi điện thoại và một trong số các hạng mục được kiểm tra của bộ mã hóa-giải mã EVS. Bộ mã hóa-giải mã EVS, theo sự chuẩn hóa, sẽ sử dụng bộ xử lý với số học cố định, và phương thức theo sáng chế cho phép giảm độ phức tạp xử lý bằng cách giảm dải động của tín hiệu mà được sử dụng cho bộ ước lượng nhiều âm thống kê cực tiểu bằng cách xử lý trị số năng lượng cho tín hiệu âm thanh trong miền loga và không còn trong miền tuyến tính.

Mặc dù một số khía cạnh đã được mô tả trong ngữ cảnh của thiết bị, rõ ràng rằng các khía cạnh này cũng biểu diễn sự mô tả của phương pháp tương ứng, mà 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.

Phụ thuộc vào các yêu cầu thực hiện nhất định, các phương án của sáng chế có thể được thực hiện trong phần cứng hoặc trong phần mềm. Các phương án có thể được thực hiện sử dụng môi trường lưu trữ số, ví dụ đĩa mềm, DVD, Blue-Ray, CD, ROM, PROM, EPROM, EEPROM hoặc bộ nhớ FLASH, có các tín hiệu điều khiển đọc được nhờ điện tử được lưu trữ trên đó, mà kết hợp (hoặc có khả năng kết hợp) với hệ thống máy tính lập trình được do đó phương pháp được thực hiện. Do đó, môi trường lưu trữ số có thể đọc được bằng máy tính.

Một số phương án theo sáng chế bao gồm vật mang dữ liệu có các tín hiệu điều khiển có thể đọc được bằng điện tử, mà có khả năng kết hợp với hệ thống máy tính có thể lập trình được, 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 hiện như sản phẩm chương trình máy tính với mã chương trình, mã chương trình có hiệu lực 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.

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 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, có thể tạo cấu hình dòng dữ liệu hoặc chuỗi tín hiệu để được truyền thông qua sự kết nối truyền thông dữ liệu, ví dụ thông qua internet.

Phương án khác gồm có phương tiện xử lý, ví dụ máy tính, hoặc thiết bị logic lập trình được, 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.

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

Các phương án được mô tả ở trên chỉ mang tính minh họa cho các nguyên lý của sáng chế. Điều này được hiểu rằng việc điều chỉnh và biến đổi của sự sắp xếp và các chi tiết được mô tả ở đây sẽ rõ ràng với những người có hiểu biết trung bình trong lĩnh vực kỹ thuật. Do đó, mục đích của sáng chế là chỉ được 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 bản mô tả và giải thích của các phương án ở đây.

YÊU CẦU BẢO HỘ

1. Phương pháp ước lượng nhiễu âm trong tín hiệu âm thanh (102), phương pháp bao gồm:

xác định (S100) trị số năng lượng (174) cho tín hiệu âm thanh (102);

chuyển đổi (S102) trị số năng lượng (174) thành miền log2; và

ước lượng (S104) mức nhiễu âm (182) cho tín hiệu âm thanh (102) dựa trên trị số năng lượng được chuyển đổi (178) trực tiếp trong miền log2,

trong đó trị số năng lượng (174) được chuyển đổi (S102) thành miền log2 như sau:

$$E_{n_{log}} = \frac{\lfloor (\log_2(1 + E_{n_{lin}})) \cdot 2^N \rfloor}{2^N}$$

$\lfloor x \rfloor$ là số nguyên lớn nhất không vượt quá (x),

$E_{n_{log}}$ trị số năng lượng của băng n trong miền log2,

$E_{n_{lin}}$ trị số năng lượng của băng n trong miền tuyến tính,

N độ phân giải lượng tử hóa.

2. Phương pháp theo điểm 1, trong đó bước ước lượng (S104) mức nhiễu âm bao gồm thực hiện thuật toán ước lượng nhiễu âm được xác định trước.

3. Phương pháp theo điểm 2, trong đó thuật toán ước lượng nhiễu âm được xác định trước bao gồm thuật toán thống kê cực tiểu.

4. Phương pháp theo một trong số các điểm từ 1 đến 3, trong đó bước xác định (S100) trị số năng lượng (174) bao gồm thu được phổ công suất của tín hiệu âm thanh (102) bằng cách biến đổi tín hiệu âm thanh (102) thành miền tần số, nhóm phổ công suất thành các băng được thúc đẩy theo cảm quan, và tích lũy các ngăn phổ công suất trong băng để tạo ra trị số năng lượng (174) cho từng băng, trong đó trị số năng lượng (174) cho từng băng được chuyển đổi thành miền log2, và trong đó mức nhiễu âm được ước lượng cho từng băng dựa trên trị số năng lượng được chuyển đổi tương ứng (174).

5. Phương pháp theo điểm 4, trong đó tín hiệu âm thanh (102) bao gồm nhiều khung, và trong đó đối với từng khung, trị số năng lượng (174) được xác định và được chuyển đổi thành miền log2, và mức nhiễu âm được ước lượng cho từng băng của khung dựa trên trị số năng lượng được chuyển đổi (174).

6. Phương pháp theo một trong số các điểm từ 1 đến 5, trong đó bước ước lượng

(S104) mức nhiễu âm dựa trên trị số năng lượng được chuyển đổi (178) mang lại dữ liệu loga, và trong đó phương pháp còn bao gồm các bước:

sử dụng (S108) dữ liệu loga một cách trực tiếp để xử lý thêm, hoặc chuyển đổi (S110, S112) dữ liệu loga trở lại thành miền tuyến tính để xử lý thêm.

7. Phương pháp theo điểm 6, trong đó:

dữ liệu loga được chuyển đổi (S108) trực tiếp thành dữ liệu truyền, trong trường hợp việc truyền được thực hiện trong miền loga, và

chuyển đổi (S110) dữ liệu loga trực tiếp thành dữ liệu truyền sử dụng hàm dịch chuyển cùng với bảng dò tìm hoặc phép gần đúng,

8. Phương pháp theo điểm 7, trong đó phép gần đúng bao gồm:

$$E_{n_lin} = 2^{(E_{n_log}-1)}.$$

9. Vật ghi có thể đọc được bằng máy tính lưu trữ các lệnh mà, khi được chạy trên máy tính, thực hiện phương pháp theo một trong số các điểm từ 1 đến 6.

10. Thiết bị ước lượng nhiễu âm (170) bao gồm:

bộ phát hiện (172) được tạo cấu hình để xác định trị số năng lượng (174) cho tín hiệu âm thanh (102);

bộ chuyển đổi (176) được tạo cấu hình để chuyển đổi trị số năng lượng (174) thành miền log2; và

thiết bị ước lượng (180) được tạo cấu hình để ước lượng mức nhiễu âm (182) cho tín hiệu âm thanh (102) dựa trên trị số năng lượng được chuyển đổi (178) trực tiếp trong miền log2,

trong đó trị số năng lượng (174) được chuyển đổi (S102) thành miền log2 như sau:

$$E_{n_log} = \frac{\lfloor (\log_2(1 + E_{n_lin})) \cdot 2^N \rfloor}{2^N}$$

$\lfloor x \rfloor$ là số nguyên lớn nhất không vượt quá (x),

E_{n_log} trị số năng lượng của băng n trong miền log2,

E_{n_lin} trị số năng lượng của băng n trong miền tuyến tính,

N độ phân giải lượng tử hóa.

11. Thiết bị mã hóa âm thanh (100) bao gồm thiết bị ước lượng nhiễu âm theo điểm 10.

12. Thiết bị giải mã âm thanh (150) bao gồm thiết bị ước lượng nhiễu âm (170) theo

điểm 10.

13. Hệ thống truyền các tín hiệu âm thanh (102), hệ thống bao gồm:

thiết bị mã hóa âm thanh (100) được tạo cấu hình để tạo ra tín hiệu âm thanh được mã hóa (102) dựa trên tín hiệu âm thanh được nhận (102); và

thiết bị giải mã âm thanh (150) được tạo cấu hình để nhận tín hiệu âm thanh được mã hóa (102), để giải mã tín hiệu âm thanh được mã hóa (102), và để xuất ra tín hiệu âm thanh được giải mã (102),

trong đó ít nhất một trong số thiết bị mã hóa âm thanh và thiết bị giải mã âm thanh bao gồm thiết bị ước lượng nhiễu âm (170) theo điểm 10.

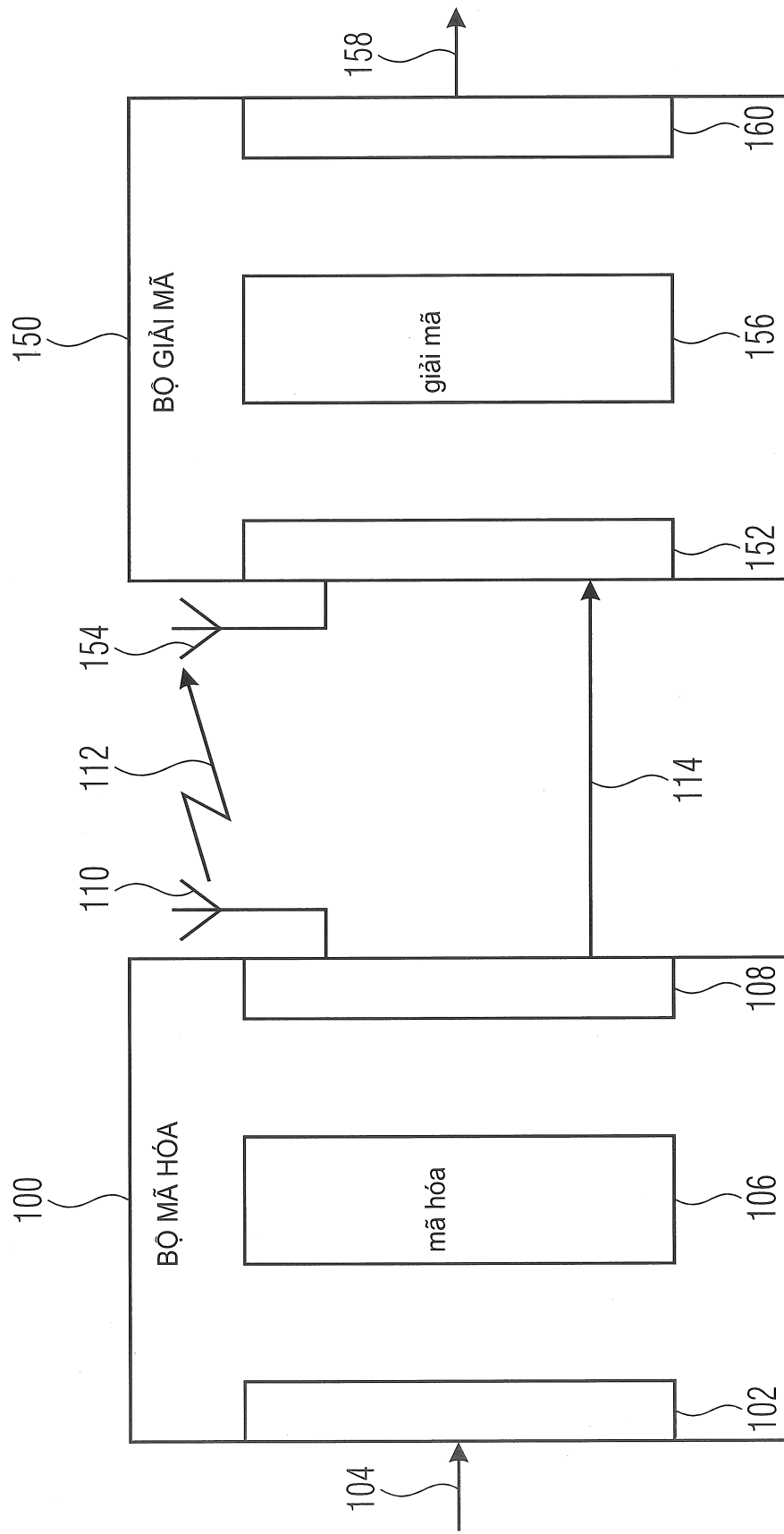


FIG 1

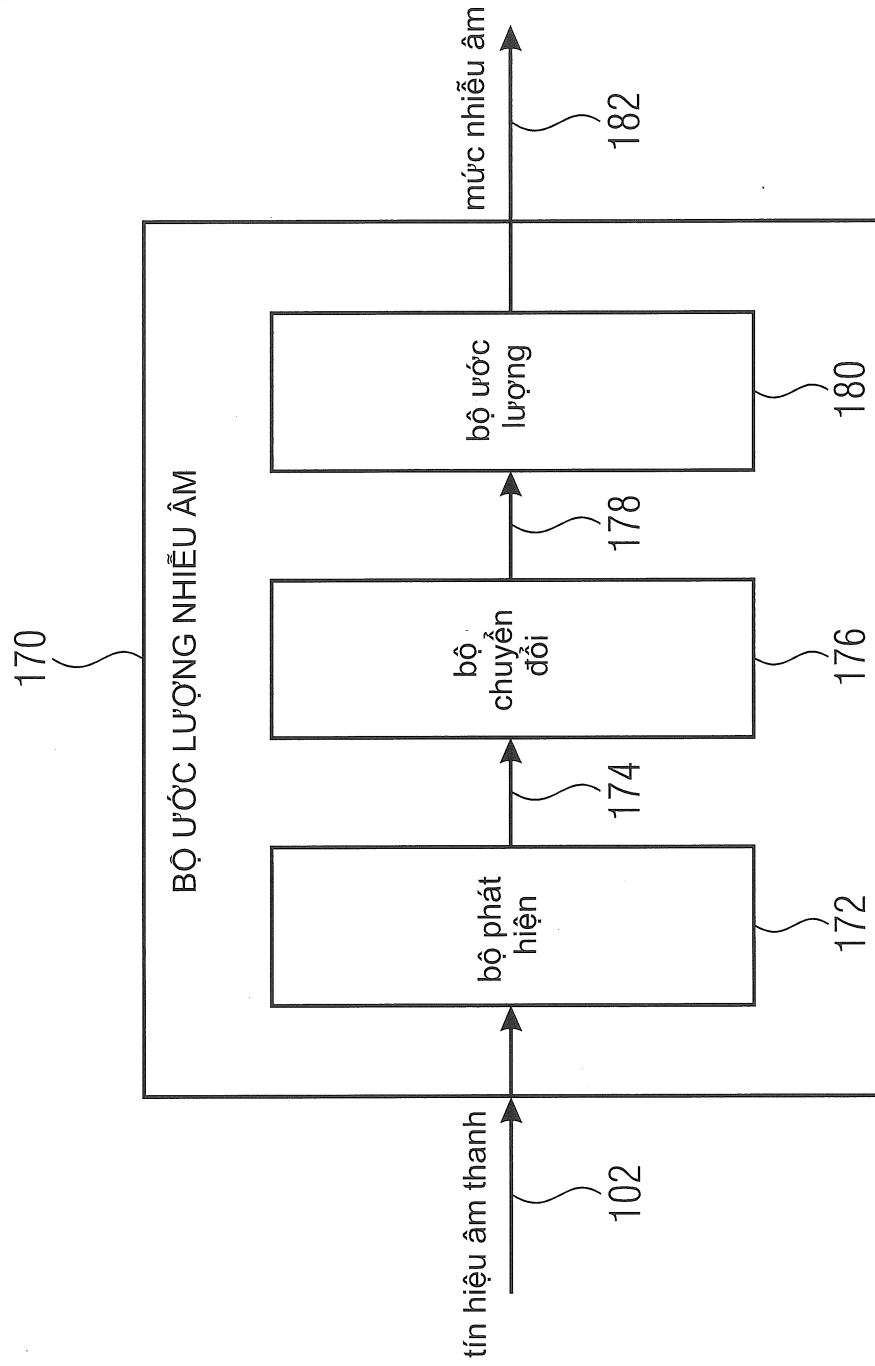


FIG 2

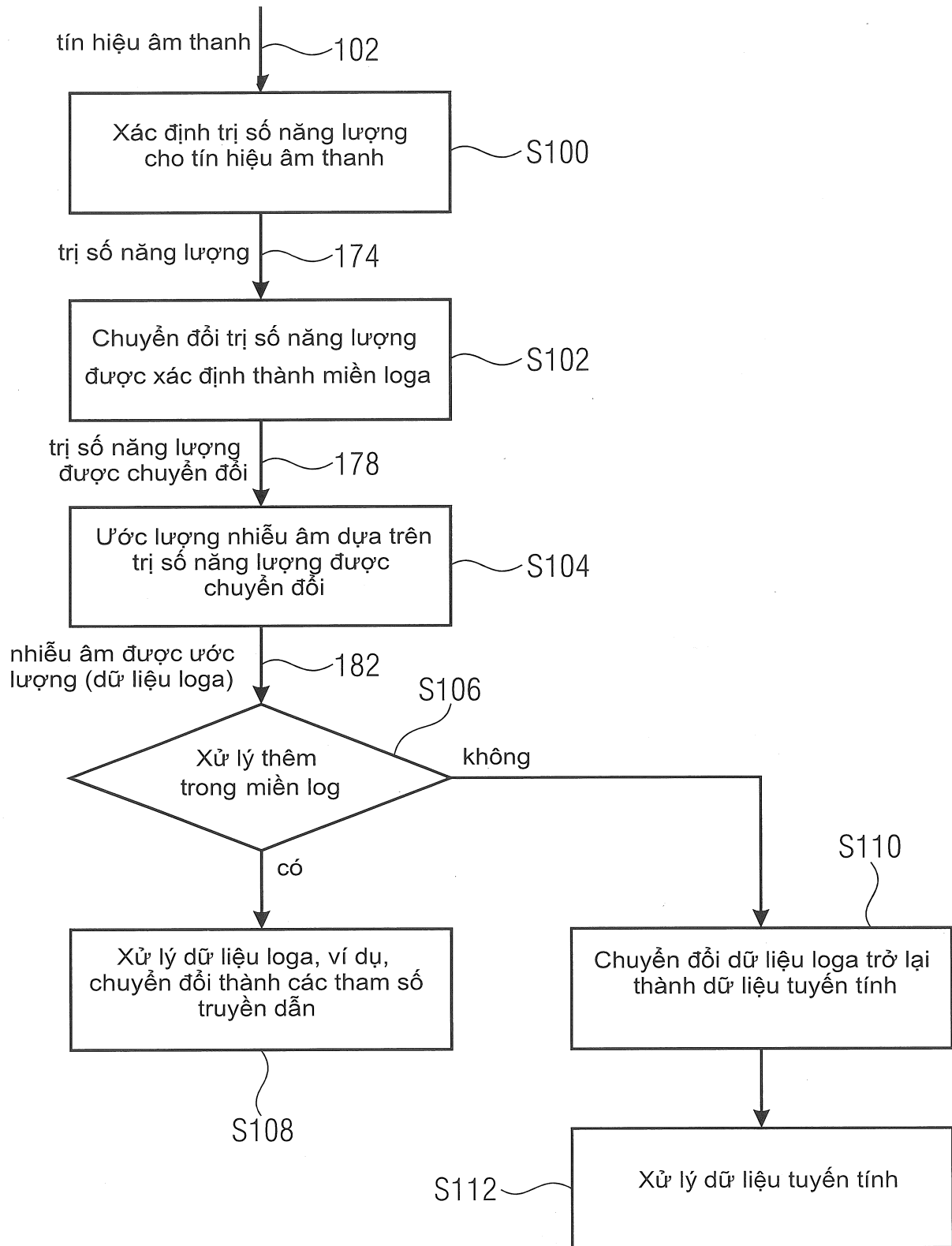


FIG 3