



(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)



1-0025860

(51)⁷ G10L 19/028

(13) B

(21) 1-2015-03182

(22) 28/01/2014

(86) PCT/EP2014/051649 28/01/2014

(87) WO 2014/118192 A2 07/08/2014

(30) 61/758,189 29/01/2013 US

(45) 26/10/2020 391

(43) 25/11/2015 332A

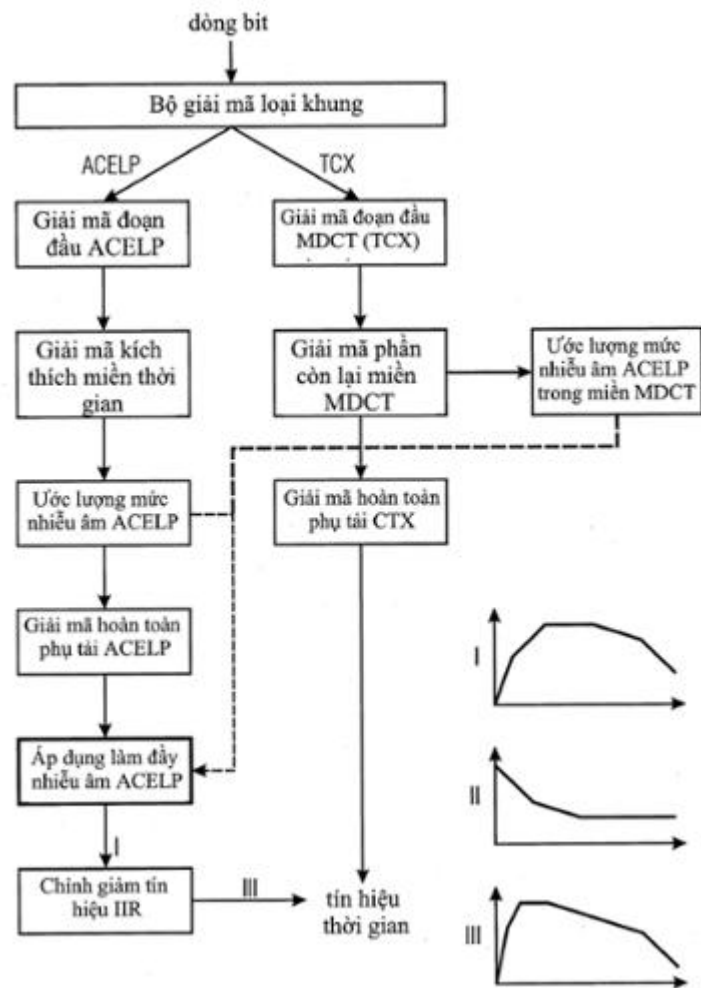
(73) Fraunhofer-Gesellschaft zur Foerderung der angewandten Forschung e.V. (DE)
Hansastraße 27c, 80686 Muenchen, Germany

(72) FUCHS, Guillaume (FR); HELMRICH, Christian (DE); JANDER, Manuel (DE);
SCHUBERT, Benjamin (DE); YOKOTANI, Yoshikazu (JP).

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

(54) BỘ GIẢI MÃ ÂM THANH VÀ PHƯƠNG PHÁP CUNG CẤP THÔNG TIN ÂM THANH ĐƯỢC GIẢI MÃ TRÊN CƠ SỞ THÔNG TIN ÂM THANH ĐƯỢC MÃ HÓA BAO GỒM CÁC HỆ SỐ DỰ BÁO TUYẾN TÍNH

(57) Sáng chế đề cập đến bộ giải mã âm thanh cung cấp thông tin âm thanh được giải mã trên cơ sở thông tin âm thanh được mã hóa bao gồm các hệ số dự báo tuyến tính (LPC), phương pháp tương ứng, chương trình máy tính tương ứng để thực hiện phương pháp này và tín hiệu âm thanh cho vật ghi lưu trữ được lưu tín hiệu âm thanh này, tín hiệu âm thanh được xử lý bằng phương pháp này. Bộ giải mã âm thanh bao gồm bộ điều chỉnh độ nghiêng được tạo cấu hình để điều chỉnh độ nghiêng của nhiều âm sử dụng các hệ số dự báo tuyến tính của khung hiện thời để thu được thông tin độ nghiêng và bộ chèn nhiều âm được tạo cấu hình để bổ sung nhiều âm cho khung hiện thời trong sự phụ thuộc vào thông tin độ nghiêng được thu bởi bộ tính toán độ nghiêng. Bộ giải mã âm thanh khác theo sáng chế bao gồm bộ ước lượng mức nhiều âm được tạo cấu hình để ước lượng mức nhiều âm của khung hiện thời sử dụng hệ số dự báo tuyến tính của ít nhất một khung có trước để thu được thông tin mức nhiều âm; và bộ chèn nhiều âm được tạo cấu hình để bổ sung nhiều âm cho khung hiện thời phụ thuộc vào thông tin mức nhiều âm được cung cấp bởi bộ ước lượng mức nhiều âm. Như vậy, thông tin phụ về nhiều âm nền trong dòng bit có thể được bỏ qua.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến bộ giải mã âm thanh cung cấp thông tin âm thanh được giải mã trên cơ sở thông tin âm thanh được mã hóa bao gồm các hệ số dự báo tuyến tính (linear prediction coefficients - LPC), đề cập đến phương pháp cung cấp thông tin âm thanh được giải mã trên cơ sở thông tin âm thanh được mã hóa bao gồm các hệ số dự báo tuyến tính (LPC), đề cập đến chương trình máy tính chạy trên máy tính, và đề cập đến tín hiệu âm thanh hoặc vật ghi lưu trữ được lưu trữ tín hiệu âm thanh nêu trên, tín hiệu âm thanh được xử lý bằng phương pháp nêu trên.

Tình trạng kỹ thuật của sáng chế

Các bộ mã hóa tiếng nói dạng số tốc độ bit thấp dựa trên nguyên lý mã hóa dự báo tuyến tính kích thích mã (code-excited linear prediction - CELP) thường bị tổn thất vì những thành phần lạ rải rác trong tín hiệu khi tốc độ bit giảm xuống thấp hơn khoảng 0,5 đến 1 bit cho mỗi mẫu, dẫn đến có phần không tự nhiên, âm thanh kim loại. Đặc biệt khi tiếng nói đầu vào có nhiều âm môi trường trong nền, các thành phần lạ tốc độ thấp có thể nghe được rõ ràng: nhiều âm nền sẽ bị tắt dần trong các phần tiếng nói hoạt động. Sáng chế mô tả sơ đồ chèn nhiều âm cho các bộ mã hóa (A) CELP như AMR-WB trong tài liệu tham khảo [1] và G.718 trong tài liệu tham khảo [4, 7] mà, tương tự như các kỹ thuật làm đầy nhiều âm được sử dụng trong sự biến đổi trên cơ sở các bộ mã hóa như xHE-AAC trong tài liệu tham khảo [5, 6], bổ sung đầu ra của bộ tạo nhiều âm ngẫu nhiên cho tín hiệu tiếng nói được giải mã để khôi phục nhiều âm nền.

Công bố quốc tế WO 2012/110476 A1 thể hiện khái niệm mã hóa mà việc dự báo tuyến tính dựa trên và sử dụng tạo hình nhiều âm miền tần số. Việc phân tích phổ của tín hiệu đầu vào âm thanh thành ảnh phổ bao gồm chuỗi phổ được sử dụng cho cả tính hệ số dự báo tuyến tính cũng như đầu vào cho tạo hình miền tần số dựa trên các hệ số dự báo tuyến tính. Theo tài liệu đối chứng bộ mã hóa âm thanh bao gồm bộ phân tích dự báo tuyến tính để phân tích tín hiệu âm thanh đầu vào để từ đó suy ra các hệ số

dự báo tuyến tính. Bộ tạo hình miền tần số của bộ mã hóa âm thanh được tạo cấu hình để tạo hình phổ hiện tại theo phổ của chuỗi phổ của ảnh phổ dựa trên các hệ số dự báo tuyến tính được cung cấp bởi bộ phân tích dự báo tuyến tính. Phổ đã tạo hình phổ và đã lượng tử hóa được chèn vào trong dòng dữ liệu cùng với thông tin trên các hệ số dự báo tuyến tính được sử dụng trong tạo hình phổ sao cho, tại phía giải mã, việc hủy tạo hình và khử lượng tử hóa có thể được thực hiện. Môđun tạo hình nhiễu âm theo thời gian có thể cũng có mặt để thực hiện tạo hình nhiễu âm theo thời gian.

Xét thấy tình trạng kỹ thuật vẫn còn yêu cầu bộ giải mã âm thanh được cải thiện, phương pháp được cải thiện, chương trình máy tính được cải thiện để thực hiện phương pháp này và tín hiệu âm thanh được cải thiện hoặc vật ghi lưu trữ được lưu trữ tín hiệu âm thanh này, tín hiệu âm thanh được xử lý bằng phương pháp này. Cụ thể hơn, tình trạng kỹ thuật mong muốn tìm ra các giải pháp cải thiện chất lượng âm của thông tin âm thanh được truyền trong dòng bit được mã hóa.

Tài liệu tham khảo

- [1] B. Bessette et al., “The Adaptive Multi-rate Wideband Speech Codec (AMR-WB)”, IEEE Trans. On Speech and Audio Processing, Vol. 10, No. 8, Nov. 2002.
- [2] R. C. Hendriks, R. Heusdens and J. Jensen, “MMSE based noise PSD tracking with low complexity”, in IEEE Int. Conf. Acoust., Speech, Signal Processing, pp. 4266 – 4269, March 2010.
- [3] R. Martin, “Noise Power Spectral Density Estimation Based on Optimal Smoothing and Minimum Statistics”, IEEE Trans. On Speech and Audio Processing, Vol. 9, No. 5, Jul. 2001.
- [4] M. Jelinek and R. Salami, “Wideband Speech Coding Advances in VMR-WB Standard”, IEEE Trans. On Audio, Speech, and Language Processing, Vol. 15, No. 4, May 2007.
- [5] J. Mäkinen et al., “AMR-WB+: A New Audio Coding Standard for 3rd Generation Mobile Audio Services”, in Proc. ICASSP 2005, Philadelphia, USA, Mar. 2005.

[6] M. Neuendorf et al., “MPEG Unified Speech and Audio Coding – The ISO/MPEG Standard for High-Efficiency Audio Coding of All Content Types”, in Proc. 132nd AES Convention, Budapest, Hungary, Apr. 2012. Also appears in the Journal of the AES, 2013.

[7] T. Vaillancourt et al., “ITU-T EV-VBR: A Robust 8 – 32 kbit/s Scalable Coder for Error Prone Telecommunications Channels”, in Proc. EUSIPCO 2008, Lausanne, Switzerland, Aug. 2008.

Bản chất kỹ thuật của sáng chế

Các ký hiệu tham chiếu trong các điểm yêu cầu bảo hộ và trong mô tả chi tiết các phương án theo sáng chế được bổ sung để đơn thuần cải thiện khả năng đọc và không có nghĩa để giới hạn.

Mục đích của sáng chế được giải quyết bởi bộ giải mã âm thanh cung cấp thông tin âm thanh được giải mã trên cơ sở thông tin âm thanh được mã hóa bao gồm các hệ số dự báo tuyến tính (linear prediction coefficients - LPC), bộ giải mã âm thanh bao gồm bộ điều chỉnh độ nghiêng được tạo cấu hình để điều chỉnh độ nghiêng của nhiều âm sử dụng các hệ số dự báo tuyến tính của khung hiện thời để thu được thông tin độ nghiêng và bộ chèn nhiều âm được tạo cấu hình để bổ sung nhiều âm cho khung hiện thời phụ thuộc vào thông tin độ nghiêng được thu bởi bộ tính toán độ nghiêng. Ngoài ra, mục đích của sáng chế được giải quyết bởi phương pháp cung cấp thông tin âm thanh được giải mã trên cơ sở thông tin âm thanh được mã hóa bao gồm các hệ số dự báo tuyến tính (LPC), phương pháp bao gồm điều chỉnh độ nghiêng của nhiều âm sử dụng các hệ số dự báo tuyến tính của khung hiện thời để thu được thông tin độ nghiêng và việc bổ sung nhiều âm cho khung hiện thời phụ thuộc vào thông tin độ nghiêng đã thu được.

Như giải pháp thứ hai theo sáng chế, sáng chế đề xuất bộ giải mã âm thanh cung cấp thông tin âm thanh được giải mã trên cơ sở thông tin âm thanh được mã hóa bao gồm các hệ số dự báo tuyến tính (LPC), bộ giải mã âm thanh bao gồm bộ ước lượng mức nhiều âm được tạo cấu hình để ước lượng mức nhiều âm cho khung hiện thời sử dụng hệ số dự báo tuyến tính của ít nhất một khung có trước để thu được thông tin mức nhiều âm, và bộ chèn nhiều âm được tạo cấu hình để bổ sung nhiều âm cho

khung hiện thời phụ thuộc vào thông tin mức nhiễu âm được cung cấp bởi bộ ước lượng mức nhiễu âm. Ngoài ra, mục đích của sáng chế được giải quyết bởi phương pháp cung cấp thông tin âm thanh được giải mã trên cơ sở thông tin âm thanh được mã hóa bao gồm các hệ số dự báo tuyến tính (LPC), phương pháp bao gồm ước lượng mức nhiễu âm cho khung hiện thời sử dụng hệ số dự báo tuyến tính của ít nhất một khung có trước để thu được thông tin mức nhiễu âm, và bổ sung nhiễu âm cho khung hiện thời phụ thuộc vào thông tin mức nhiễu âm được cung cấp bởi việc ước lượng mức nhiễu âm. Ngoài ra, mục đích của sáng chế được giải quyết bởi chương trình máy tính để thực hiện phương pháp này, trong đó chương trình máy tính chạy trên máy tính, và tín hiệu âm thanh hoặc vật ghi lưu trữ được lưu tín hiệu âm thanh này, tín hiệu âm thanh được xử lý bằng phương pháp này.

Các giải pháp đã đề xuất tránh việc cung cấp thông tin phụ trong dòng bit CELP để điều chỉnh nhiễu âm được cung cấp lên phía bộ giải mã trong xử lý làm đầy nhiễu âm. Điều này có nghĩa là lượng dữ liệu để được chuyển với dòng bit có thể bị giảm trong khi chất lượng của nhiễu âm được chèn có thể được tăng lên chỉ trên cơ sở các hệ số dự báo tuyến tính của khung được giải mã hiện tại hoặc có trước. Nói cách khác, thông tin phụ liên quan đến nhiễu âm mà sẽ làm tăng lượng dữ liệu để được truyền với dòng bit có thể được bỏ qua. Sáng chế cho phép cung cấp bộ mã hóa dạng số tốc độ bit thấp và phương pháp có thể sử dụng bằng thông tin ít hơn liên quan đến dòng bit và cung cấp chất lượng nhiễu âm nền được cải thiện khi so sánh với các giải pháp của tình trạng kỹ thuật.

Tốt hơn là bộ giải mã âm thanh bao gồm bộ xác định loại khung để xác định loại khung của khung hiện thời, bộ xác định loại khung được tạo cấu hình để kích hoạt bộ điều chỉnh độ nghiêng để điều chỉnh độ nghiêng của nhiễu âm khi loại khung của khung hiện thời được phát hiện là thuộc loại tiếng nói. Trong một số phương án, bộ phát hiện loại khung được tạo cấu hình để nhận dạng khung như là khung loại tiếng nói khi khung là ACELP hoặc CELP được mã hóa. Việc tạo hình nhiễu âm theo độ nghiêng của khung hiện thời có thể cung cấp nhiễu âm nền tự nhiên hơn và có thể làm giảm những tác động không mong muốn của việc nén âm thanh đối với nhiễu âm nền của tín hiệu mong muốn được mã hóa trong dòng bit. Khi các tác động nén không mong muốn đó và các thành phần lạ thường trở nên có thể được nhận thấy đối với

nhiều âm nền của thông tin tiếng nói, điều đó có thể thuận lợi để nâng cao chất lượng nhiều âm để được bổ sung cho các khung loại tiếng nói này bằng cách điều chỉnh độ nghiêng của nhiều âm trước khi bổ sung nhiều âm cho khung hiện thời. Theo đó, bộ chèn nhiều âm có thể được tạo cấu hình để bổ sung nhiều âm cho khung hiện thời chỉ khi nếu khung hiện thời là khung tiếng nói, vì việc này có thể làm giảm tải làm việc trên phía bộ giải mã nếu chỉ các khung tiếng nói được xử lý bằng cách làm đầy nhiều âm.

Trong phương án được ưu tiên của sáng chế, bộ điều chỉnh độ nghiêng được tạo cấu hình để sử dụng kết quả của phép phân tích bậc một các hệ số dự báo tuyến tính của khung hiện thời để thu được thông tin độ nghiêng. Bằng cách sử dụng phép phân tích bậc một các hệ số dự báo tuyến tính này mà cách sử dụng đó có thể bỏ sót thông tin phụ để biểu thị đặc điểm nhiều âm trong dòng bit. Hơn nữa, việc điều chỉnh của nhiều âm để được bổ sung có thể dựa trên các hệ số dự báo tuyến tính của khung hiện thời mà được truyền với dòng bit như thường để cho phép giải mã thông tin âm thanh của khung hiện thời. Điều này có nghĩa là các hệ số dự báo tuyến tính của khung hiện thời được sử dụng lại theo cách có lợi trong quy trình điều chỉnh độ nghiêng của nhiều âm. Hơn nữa, phép phân tích bậc đơn giản một cách hợp lý sao cho sự phức tạp tính toán của bộ giải mã âm thanh không tăng đáng kể.

Trong một số phương án của sáng chế, bộ điều chỉnh độ nghiêng được tạo cấu hình để thu được thông tin độ nghiêng sử dụng việc tính toán độ khuếch đại g của các hệ số dự báo tuyến tính của khung hiện thời như phép phân tích bậc một. Tốt hơn nữa là, độ khuếch đại g được xác định bằng công thức $g = \Sigma [a_k \cdot a_{k+1}] / \Sigma [a_k \cdot a_k]$, trong đó a_k là các hệ số LPC của khung hiện thời. Trong một số phương án, hai hoặc nhiều hệ số LPC a_k được sử dụng trong phép tính. Tốt hơn là, tổng của 16 hệ số LPC được sử dụng, sao cho $k = 0 \dots 15$. Trong các phương án của sáng chế, dòng bit có thể được mã hóa với nhiều hơn hoặc ít hơn 16 hệ số LPC. Vì các hệ số dự báo tuyến tính của khung hiện thời dễ dàng có mặt trong dòng bit, thông tin độ nghiêng có thể được thu mà không sử dụng thông tin phụ, vì vậy làm giảm lượng dữ liệu để được truyền trong dòng bit. Nhiều âm được bổ sung có thể được điều chỉnh bằng cách sử dụng các hệ số dự báo tuyến tính mà cần thiết để giải mã thông tin âm thanh được mã hóa.

Tốt hơn là, bộ điều chỉnh độ nghiêng được tạo cấu hình để thu được thông tin độ nghiêng sử dụng cách tính hàm truyền của bộ lọc dạng trực tiếp $x(n) - g \cdot x(n-1)$ cho khung hiện thời. Cách tính kiểu này khá đơn giản và không đòi hỏi năng lượng tính toán lớn trên phía bộ giải mã. Độ khuếch đại g có thể được tính toán một cách dễ dàng từ các hệ số LPC của khung hiện thời, như được thể hiện ở trên. Việc này cho phép cải thiện chất lượng nhiều âm cho các bộ mã hoá dạng số tốc độ bit thấp trong khi sử dụng hoàn toàn dữ liệu dòng bit chủ yếu để giải mã thông tin âm thanh đã mã hoá.

Trong phương án được ưu tiên của sáng chế, bộ chèn nhiều âm được tạo cấu hình để áp dụng thông tin độ nghiêng của khung hiện thời cho nhiều âm để điều chỉnh độ nghiêng của nhiều âm trước khi bổ sung nhiều âm cho khung hiện thời. Nếu bộ chèn nhiều âm được tạo cấu hình một cách phù hợp, bộ giải mã âm thanh đã đơn giản hoá có thể được đề xuất. Bằng cách sử dụng đầu tiên thông tin độ nghiêng và sau đó bổ sung nhiều âm đã điều chỉnh cho khung hiện thời, có thể cung cấp phương pháp đơn giản và có hiệu quả cho bộ giải mã âm thanh.

Trong phương án của sáng chế, bộ giải mã âm thanh còn bao gồm bộ ước lượng mức nhiều âm được tạo cấu hình để ước lượng mức nhiều âm cho khung hiện thời sử dụng hệ số dự báo tuyến tính của ít nhất một khung có trước để thu được thông tin mức nhiều âm, và bộ chèn nhiều âm được tạo cấu hình để bổ sung nhiều âm cho khung hiện thời phụ thuộc vào thông tin mức nhiều âm được cung cấp bởi bộ ước lượng mức nhiều âm. Bằng cách này, chất lượng của nhiều âm nền và do đó chất lượng của toàn bộ việc truyền dẫn thông tin có thể được nâng cao vì nhiều âm được bổ sung cho khung hiện thời có thể được điều chỉnh theo mức nhiều âm mà có thể có mặt trong khung hiện thời. Ví dụ, nếu mức nhiều âm cao được mong muốn trong khung hiện thời vì mức nhiều âm cao đã được ước lượng từ các khung có trước, bộ chèn nhiều âm có thể được tạo cấu hình để làm tăng mức của nhiều âm để được bổ sung cho khung hiện thời trước khi bổ sung nó cho khung hiện thời. Do đó, nhiều âm được bổ sung có thể được điều chỉnh để hoặc quá nhỏ hoặc quá lớn khi so sánh với mức nhiều âm mong muốn trong khung hiện thời. Mặt khác, việc điều chỉnh này không dựa trên thông tin phụ dành riêng trong dòng bit mà chỉ sử dụng thông tin dữ liệu cần thiết được truyền trong dòng bit, trong trường hợp này hệ số dự báo tuyến tính của ít nhất một khung có trước mà cũng cung cấp thông tin về mức nhiều âm trong khung có trước. Do đó, tốt

hơn là nhiều âm để được bổ sung cho khung hiện thời được tạo hình sử dụng độ nghiêng được suy ra g và được định tỉ lệ xét đến ước lượng mức nhiều âm. Tốt hơn nhất là, độ nghiêng và mức nhiều âm của nhiều âm để được bổ sung cho khung hiện thời cũng được điều chỉnh khi khung hiện thời thuộc loại tiếng nói. Trong một số phương án, độ nghiêng và/hoặc mức nhiều âm được bổ sung cho khung hiện thời cũng được điều chỉnh khi khung hiện thời là dạng âm thanh thông thường, ví dụ loại TCX hoặc DTX.

Tốt hơn là, bộ giải mã âm thanh bao gồm bộ xác định loại khung để xác định loại khung của khung hiện thời, bộ xác định loại khung được tạo cấu hình để nhận biết loại khung của khung hiện thời hoặc là tiếng nói hoặc là âm thanh thông thường, sao cho việc ước lượng mức nhiều âm có thể được thực hiện phụ thuộc vào loại khung của khung hiện thời. Ví dụ, bộ xác định loại khung có thể được tạo cấu hình để phát hiện khung hiện thời hoặc là khung CELP hoặc khung ACELP, là loại khung tiếng nói, hoặc là khung TCX/MDCT hoặc khung DTX, là các loại khung âm thanh thông thường. Vì các dạng mã hoá đó theo các nguyên lý khác nhau, mà mong muốn xác định loại khung trước khi thực hiện ước lượng mức nhiều âm sao cho các phép tính thích hợp có thể được chọn, phụ thuộc vào loại khung.

Trong một số phương án của sáng chế bộ giải mã âm thanh được làm thích ứng để tính thông tin thứ nhất biểu diễn kích thích không tạo hình phổ của khung hiện thời và để tính thông tin thứ hai liên quan đến định tỉ lệ phổ của khung hiện thời để tính thương số của thông tin thứ nhất với thông tin thứ hai để thu được thông tin mức nhiều âm. Bằng cách này, thông tin mức nhiều âm có thể thu được mà không sử dụng bất kỳ thông tin phụ nào. Do đó, tốc độ bit của bộ mã hoá có thể được duy trì chậm.

Tốt hơn là, bộ giải mã âm thanh được làm thích ứng để giải mã tín hiệu kích thích của khung hiện thời và để tính căn bậc hai trung bình bình phương e_{rms} từ việc biểu diễn miền thời gian của khung hiện thời như thông tin thứ nhất để thu được thông tin mức nhiều âm với điều kiện là khung hiện thời thuộc loại tiếng nói. Tốt hơn là đối với phương án mà bộ giải mã âm thanh được làm thích ứng để thực hiện một cách phù hợp nếu khung hiện thời thuộc loại CELP hoặc ACELP. Tín hiệu kích thích được làm phẳng phổ (trong miền cảm giác) được giải mã từ dòng bit và được sử dụng để cập

nhất ước lượng mức nhiễu âm. Căn bậc hai trung bình bình phương e_{rms} của tín hiệu kích thích cho khung hiện thời được tính sau khi dòng bit được đọc. Loại tính toán này có thể không đòi hỏi năng lượng tính toán cao và do đó có thể thậm chí được thực hiện bởi các bộ giải mã âm thanh có năng lượng tính toán thấp.

Trong phương án được ưu tiên bộ giải mã âm thanh được làm thích ứng để tính mức đỉnh p của hàm truyền của bộ lọc LPC của khung hiện thời như thông tin thứ hai, do đó việc sử dụng hệ số dự báo tuyến tính để thu được thông tin mức nhiễu âm với điều kiện là khung hiện thời thuộc loại tiếng nói. Mặt khác, tốt hơn là khung hiện thời thuộc loại CELP hoặc ACELP. Việc tính mức đỉnh p khá rẻ, và bằng cách sử dụng lại các hệ số dự báo tuyến tính của khung hiện thời, mà cũng được sử dụng để giải mã thông tin âm thanh được chứa trong khung đó, thông tin phụ có thể bị bỏ qua và nhiễu âm nền vẫn có thể được nâng cao mà không làm tăng tốc độ dữ liệu của dòng bit.

Trong phương án được ưu tiên của sáng chế, bộ giải mã âm thanh được thích ứng để tính mức tối thiểu phổ m_f của khung âm thanh hiện tại bằng cách tính thương số của căn bậc hai trung bình bình phương e_{rms} và mức đỉnh p để thu được thông tin mức nhiễu âm với điều kiện là khung hiện thời thuộc loại tiếng nói. Cách tính này khá đơn giản và có thể cung cấp giá trị bằng số để có thể có ích trong việc ước lượng mức nhiễu âm trên vùng của các khung âm thanh phức tạp. Do đó, mức tối thiểu phổ m_f của các chuỗi trong số các khung âm thanh hiện tại có thể được sử dụng để ước lượng mức nhiễu âm trong khoảng thời gian được chuyển đổi bởi các chuỗi đó trong số các khung âm thanh. Việc này có thể cho phép thu được sự ước lượng tốt cho mức nhiễu âm của khung hiện thời trong khi vẫn giữ được tính phức tạp khá thấp. Mức đỉnh p tốt hơn là được tính toán sử dụng công thức $p = \sum |a_k|$, trong đó a_k là hệ số dự báo tuyến tính tốt hơn là với $k = 0 \dots 15$. Do đó, nếu khung bao gồm 16 hệ số dự báo tuyến tính, p trong một số phương án được tính toán bằng cách cộng toàn bộ các biên độ của tốt hơn là 16 a_k .

Tốt hơn là, bộ giải mã âm thanh được làm thích ứng để giải mã kích thích MDCT không tạo hình của khung hiện thời và tính căn bậc hai trung bình bình phương của nó e_{rms} từ việc biểu diễn miền phổ của khung hiện thời để thu được thông tin mức nhiễu âm như thông tin thứ nhất nếu khung hiện thời thuộc loại âm thanh thông

thường. Đây là phương án được ưu tiên của sáng chế khi khung hiện thời không phải là khung tiếng nói mà là khung âm thanh thông thường. Việc biểu diễn miền phổ trong các khung MDCT hoặc DTX là tương đương ở mức độ lớn với biểu diễn miền thời gian trong các khung tiếng nói, ví dụ các khung CELP hoặc (A)CELP. Các đường khác nhau trong đó MDCT không được xếp vào hàng trong phép toán định lý Parseval. Do đó, tốt hơn là căn bậc hai trung bình bình phương e_{rms} cho khung âm thanh thông thường được tính theo cách thức tương tự như căn bậc hai trung bình bình phương e_{rms} cho các khung tiếng nói. Phép tính này sau đó được ưu tiên để tính các tương đương hệ số LPC của khung âm thanh thông thường như được trình bày trong WO 2012/110476 A1, ví dụ sử dụng phổ lũy thừa MDCT mà đề cập đến bình phương các giá trị MDCT trên thang đo Bark. Trong phương án thay thế, các băng tần số của phổ lũy thừa MDCT có thể có chiều rộng không đổi sao cho thang đo của phổ tương ứng với thang đo tuyến tính. Với thang đo tuyến tính này các tương đương hệ số LPC đã tính toán tương tự như hệ số LPC trong biểu diễn miền thời gian của khung tương tự, bởi vì, ví dụ, được tính toán cho khung ACELP hoặc CELP. Hơn nữa, tốt hơn là, nếu khung hiện thời thuộc loại âm thanh thông thường, mức đỉnh p của hàm truyền của bộ lọc LPC của khung hiện thời được tính toán từ khung MDCT như được trình bày trong WO 2012/110476 A1 được tính như thông tin thứ hai, do đó việc sử dụng hệ số dự báo tuyến tính để thu được thông tin mức nhiễu âm với điều kiện là khung hiện thời thuộc loại âm thanh thông thường. Sau đó, nếu khung hiện thời thuộc loại âm thanh thông thường, tốt hơn là tính mức tối thiểu phổ của khung âm thanh hiện tại bằng cách tính thương số của căn bậc hai trung bình bình phương e_{rms} với mức đỉnh p để thu được thông tin mức nhiễu âm với điều kiện là khung hiện thời thuộc loại âm thanh thông thường. Do đó, thương số mô tả mức tối thiểu phổ m_f của khung âm thanh hiện tại có thể được thu bất kể khung hiện thời thuộc loại tiếng nói hoặc thuộc loại âm thanh thông thường.

Trong phương án được ưu tiên, bộ giải mã âm thanh được làm thích ứng để xếp vào hàng thương số được thu từ khung âm thanh hiện tại trong bộ ước lượng mức nhiễu âm bất kể loại khung nào, bộ ước lượng mức nhiễu âm bao gồm bộ lưu mức nhiễu âm cho hai hoặc nhiều thương số được thu từ các khung âm thanh khác nhau. Việc này có thể thuận lợi nếu bộ giải mã âm thanh được làm thích ứng để chuyển giữa

giải mã các khung tiếng nói với giải mã các khung âm thanh thông thường, ví dụ khi áp dụng tiếng nói đồng nhất độ trễ thấp và giải mã âm thanh (LD-USAC, EVS). Bằng cách này, mức nhiễu âm trung bình trên khắp các khung phức tạp có thể thu được, không liên quan đến loại khung. Tốt hơn là bộ lưu mức nhiễu âm có thể chứa được mười hoặc nhiều thương số được thu từ mười hoặc nhiều khung âm thanh có trước. Ví dụ, bộ lưu mức nhiễu âm có thể chứa chỗ cho các thương số của 30 khung. Do đó, mức nhiễu âm có thể được tính toán cho thời gian mở rộng có trước khung hiện thời. Trong một số phương án, thương số có thể chỉ được xếp vào hàng trong bộ ước lượng mức nhiễu âm khi khung hiện thời được phát hiện thuộc loại tiếng nói. Trong các phương án khác, thương số có thể chỉ được xếp vào hàng trong bộ ước lượng mức nhiễu âm khi khung hiện thời được phát hiện thuộc loại âm thanh thông thường..

Tốt hơn là bộ ước lượng mức nhiễu âm được làm thích ứng để ước lượng mức nhiễu âm trên cơ sở phân tích thống kê hai hoặc nhiều thương số của các khung âm thanh khác nhau. Trong phương án của sáng chế, bộ giải mã âm thanh được làm thích ứng để sử dụng sai số bình phương trung bình cực tiểu dựa trên việc tạo vết mật độ phổ lũy thừa nhiễu âm để phân tích thống kê các thương số. Việc tạo vết được mô tả trong công bố của Hendriks, Heusdens và Jensen trong tài liệu tham khảo [2]. Nếu phương pháp theo tài liệu tham khảo [2] được áp dụng, bộ giải mã âm thanh được làm thích ứng để sử dụng căn bậc hai trung bình bình phương của giá trị vết trong phân tích thống kê, như trong trường hợp này phổ biên độ được tìm kiếm trực tiếp. Trong phương án khác của sáng chế, số liệu thống kê cực tiểu như được biết từ tài liệu tham khảo [3] được sử dụng để phân tích hai hoặc nhiều thương số của các khung âm thanh khác nhau.

Trong phương án được ưu tiên, bộ giải mã âm thanh bao gồm lõi bộ giải mã được tạo cấu hình để giải mã thông tin âm thanh của khung hiện thời sử dụng hệ số dự báo tuyến tính của khung hiện thời để thu được tín hiệu đầu ra bộ mã hóa lõi đã giải mã và bộ chèn nhiễu âm bổ sung nhiễu âm phụ thuộc vào hệ số dự báo tuyến tính được sử dụng trong giải mã thông tin âm thanh của khung hiện thời và/hoặc được sử dụng khi giải mã thông tin âm thanh của một hoặc nhiều khung có trước. Do đó, bộ chèn nhiễu âm sử dụng các hệ số dự báo tuyến tính mà được sử dụng để giải mã thông tin

âm thanh của khung hiện thời. Thông tin phụ để chỉ dẫn bộ chèn nhiều âm có thể được bỏ qua.

Tốt hơn là, bộ giải mã âm thanh bao gồm bộ lọc chỉnh giảm để chỉnh giảm khung hiện thời, bộ giải mã âm thanh được làm thích ứng để áp dụng bộ lọc chỉnh giảm trên khung hiện thời sau khi bộ chèn nhiều âm đã bổ sung nhiều âm cho khung hiện thời. Vì sự chỉnh giảm là IIR bậc một tăng lên các tần số thấp, việc này cho phép cho tính phức tạp thấp, lọc thông cao IIR dốc của nhiều âm được bổ sung tránh các thành phần lạ nhiều âm có thể nghe được ở các tần số thấp.

Tốt hơn là, bộ giải mã âm thanh bao gồm bộ tạo nhiều âm, bộ tạo nhiều âm được làm thích ứng để tạo nhiều âm để được bổ sung cho khung hiện thời bởi bộ chèn nhiều âm. Bộ tạo nhiều âm được chứa để bộ giải mã âm thanh có thể cung cấp bộ giải mã âm thanh thích hợp hơn vì bộ tạo nhiều âm bên ngoài là không cần thiết. Ngoài ra, nhiều âm có thể được cung cấp bởi bộ tạo nhiều âm bên ngoài, mà có thể được kết nối với bộ giải mã âm thanh thông qua giao diện. Ví dụ, các loại phổ của các bộ tạo nhiều âm có thể được áp dụng, phụ thuộc vào nhiều âm nền mà được tăng cường trong khung hiện thời.

Tốt hơn là, bộ tạo nhiều âm được tạo cấu hình để tạo nhiều âm trắng ngẫu nhiên. Nhiều âm này tương tự với các nhiều âm nền thông thường và tương ứng bộ tạo nhiều âm này có thể được cung cấp một cách dễ dàng.

Trong phương án được ưu tiên của sáng chế, bộ chèn nhiều âm được tạo cấu hình để bổ sung nhiều âm cho khung hiện thời với điều kiện là tốc độ của thông tin âm thanh được mã hóa nhỏ hơn 1 bit cho mỗi mẫu. Tốt hơn là tốc độ bit của thông tin âm thanh được mã hóa nhỏ hơn 0,8 bit cho mỗi mẫu. Thậm chí tốt hơn là bộ chèn nhiều âm được tạo cấu hình để bổ sung nhiều âm cho khung hiện thời với điều kiện là tốc độ bit của thông tin âm thanh được mã hóa nhỏ hơn 0,5 bit cho mỗi mẫu.

Trong phương án được ưu tiên, bộ giải mã âm thanh được tạo cấu hình để sử dụng bộ mã hóa dựa trên một hoặc nhiều bộ mã hóa AMR-WB, G.718 hoặc LD-USAC (EVS) để giải mã thông tin âm thanh được mã hóa. Đây là những bộ mã hóa nổi tiếng và trải rộng (A)CELP trong đó việc sử dụng thêm phương pháp làm đầy nhiều âm này có thể hết sức thuận lợi.

Mô tả vắn tắt các hình vẽ

Các phương án của sáng chế được mô tả sau đây cùng với các hình vẽ đi kèm.

Fig.1 thể hiện phương án thứ nhất của bộ giải mã âm thanh theo sáng chế;

Fig.2 thể hiện phương pháp thứ nhất để thực hiện giải mã âm thanh theo sáng chế mà có thể được thực hiện bởi bộ giải mã âm thanh theo Fig.1;

Fig.3 thể hiện phương án thứ hai của bộ giải mã âm thanh theo sáng chế;

Fig.4 thể hiện phương pháp thứ hai để thực hiện giải mã âm thanh theo sáng chế mà có thể được thực hiện bởi bộ giải mã âm thanh theo Fig.3;

Fig.5 thể hiện phương án thứ ba của bộ giải mã âm thanh theo sáng chế;

Fig.6 thể hiện phương pháp thứ ba để thực hiện giải mã âm thanh theo sáng chế mà có thể được thực hiện bởi bộ giải mã âm thanh theo Fig.5;

Fig.7 thể hiện minh họa của phương pháp tính toán mức tối thiểu phổ m_f cho các ước lượng mức nhiễu âm;

Fig.8 thể hiện biểu đồ minh họa độ nghiêng được suy ra từ các hệ số LPC; và

Fig.9 thể hiện biểu đồ minh họa cách các tương đương bộ lọc LPC được xác định từ phổ lũy thừa MDCT.

Mô tả chi tiết sáng chế

Sáng chế được mô tả chi tiết với các hình vẽ từ 1 đến 9. Sáng chế không có nghĩa bị giới hạn bởi việc thể hiện cũng như các phương án được mô tả.

Fig.1 thể hiện phương án thứ nhất của bộ giải mã âm thanh theo sáng chế. Bộ giải mã âm thanh được làm thích ứng để cung cấp thông tin âm thanh được giải mã trên cơ sở thông tin âm thanh được mã hóa. Bộ giải mã âm thanh được tạo cấu hình để sử dụng bộ mã hóa mà có thể dựa trên AMR-WB, G.718 và LD-USAC (EVS) để giải mã thông tin âm thanh được mã hóa. Thông tin âm thanh được mã hóa bao gồm các hệ số dự báo tuyến tính (LPC), mà có thể được ký hiệu riêng là hệ số a_k . Bộ giải mã âm thanh bao gồm bộ điều chỉnh độ nghiêng được tạo cấu hình để điều chỉnh độ nghiêng của nhiễu âm sử dụng các hệ số dự báo tuyến tính của khung hiện thời để thu được thông tin độ nghiêng và bộ chèn nhiễu âm được tạo cấu hình để bổ sung nhiễu âm cho

khung hiện thời phụ thuộc vào thông tin độ nghiêng được thu bởi bộ tính toán độ nghiêng. Bộ chèn nhiều âm được tạo cấu hình để bổ sung nhiều âm cho khung hiện thời với điều kiện là tốc độ bit của thông tin âm thanh được mã hóa nhỏ hơn 1 bit cho mỗi mẫu. Hơn nữa, bộ chèn nhiều âm có thể được tạo cấu hình để bổ sung nhiều âm cho khung hiện thời với điều kiện là khung hiện thời là khung tiếng nói. Do đó, nhiều âm có thể được bổ sung cho khung hiện thời để cải thiện chất lượng âm nói chung của thông tin âm thanh được giải mã mà có thể bị làm suy yếu do mã hóa các thành phần lạ, đặc biệt đối với nhiều âm nền của thông tin tiếng nói. Khi độ nghiêng của nhiều âm được điều chỉnh trong xem xét độ nghiêng của khung âm thanh hiện tại, chất lượng âm nói chung có thể được cải thiện mà không phụ thuộc vào thông phụ trong dòng bit. Do đó, lượng dữ liệu để được truyền với dòng bit có thể bị làm giảm.

Fig.2 thể hiện phương pháp thứ nhất để thực hiện giải mã âm thanh theo sáng chế mà có thể được thực hiện bởi bộ giải mã âm thanh theo Fig.1. Các chi tiết kỹ thuật của bộ giải mã âm thanh đã vẽ trong Fig.1 được mô tả cùng với các đặc tính phương pháp. Bộ giải mã âm thanh được làm thích ứng để đọc dòng bit của thông tin âm thanh được mã hóa. Bộ giải mã âm thanh bao gồm bộ xác định loại khung để xác định loại khung của khung hiện thời, bộ xác định loại khung được tạo cấu hình để kích hoạt bộ điều chỉnh độ nghiêng để điều chỉnh độ nghiêng của nhiều âm khi loại khung của khung hiện thời được phát hiện là thuộc loại tiếng nói. Do đó, bộ giải mã âm thanh xác định loại khung của khung âm thanh hiện tại bằng cách áp dụng bộ xác định loại khung. Nếu khung hiện thời là khung ACELP, bộ xác định loại khung kích hoạt bộ điều chỉnh độ nghiêng. Bộ điều chỉnh độ nghiêng được tạo cấu hình để sử dụng kết quả của phép phân tích bậc một các hệ số dự báo tuyến tính của khung hiện thời để thu được thông tin độ nghiêng. Cụ thể hơn, bộ điều chỉnh độ nghiêng tính toán độ khuếch đại g sử dụng công thức $g = \Sigma [a_k \cdot a_{k+1}] / \Sigma [a_k \cdot a_k]$ như phép phân tích bậc một, trong đó a_k là các hệ số LPC của khung hiện thời. Fig.8 thể hiện biểu đồ minh họa độ nghiêng được suy ra từ các hệ số LPC. Fig.8 thể hiện hai khung thuộc từ "thấy". Đối với chữ "s", mà có lượng lớn các tần số cao, độ nghiêng đi lên. Đối với các chữ "ee", mà có lượng lớn các tần số thấp, độ nghiêng đi xuống. Độ nghiêng phổ được thể hiện trong Fig.8 là hàm truyền của bộ lọc dạng trực tiếp $x(n) - g \cdot x(n-1)$, g được xác định như đã cho ở trên. Do đó, bộ điều chỉnh độ nghiêng sử dụng các hệ số LPC được cung cấp

trong dòng bit và được sử dụng để giải mã thông tin âm thanh đã mã hoá. Thông tin phụ có thể bị bỏ qua một cách hợp lý mà có thể làm giảm lượng dữ liệu được truyền với dòng bit. Hơn nữa, bộ điều chỉnh độ nghiêng được tạo cấu hình để thu được thông tin độ nghiêng sử dụng phép tính hàm truyền của bộ lọc dạng trực tiếp $x(n) - g \cdot x(n-1)$. Theo đó, bộ điều chỉnh độ nghiêng tính toán độ nghiêng của thông tin âm thanh trong khung hiện thời bằng cách tính hàm truyền của bộ lọc dạng trực tiếp $x(n) - g \cdot x(n-1)$ sử dụng độ khuếch đại g đã tính toán trước. Sau khi thu được thông tin độ nghiêng, bộ điều chỉnh độ nghiêng điều chỉnh độ nghiêng của nhiều âm để bổ sung cho khung hiện thời phụ thuộc vào thông tin độ nghiêng của khung hiện thời. Sau đó, nhiều âm được điều chỉnh được bổ sung cho khung hiện thời. Hơn nữa, không được thể hiện trong Fig.2, bộ giải mã âm thanh bao gồm bộ lọc chỉnh giảm để chỉnh giảm khung hiện thời, bộ giải mã âm thanh được làm thích ứng để áp dụng bộ lọc chỉnh giảm trên khung hiện thời sau khi bộ chèn nhiều âm bổ sung nhiều âm cho khung hiện thời. Sau khi chỉnh giảm khung, mà cũng đảm nhiệm như tính phức tạp thấp, việc lọc thông cao IIP dốc của nhiều âm được bổ sung, bộ giải mã âm thanh cung cấp thông tin âm thanh được giải mã. Do đó, phương pháp theo Fig.2 cho phép nâng cao chất lượng âm của thông tin âm thanh bằng cách điều chỉnh độ nghiêng của nhiều âm để được bổ sung cho khung hiện thời để cải thiện chất lượng của nhiều âm nền.

Fig.3 thể hiện phương án thứ hai của bộ giải mã âm thanh theo sáng chế. Bộ giải mã âm thanh lại được làm thích ứng để cung cấp thông tin âm thanh được giải mã trên cơ sở thông tin âm thanh được mã hóa. Bộ giải mã âm thanh lại được tạo cấu hình để sử dụng bộ mã hóa mà có thể dựa trên AMR-WB, G.718 và LD-USAC (EVS) để giải mã thông tin âm thanh được mã hóa. Thông tin âm thanh được mã hoá lại bao gồm các hệ dự báo tuyến tính (LPC), mà có thể được ký hiệu riêng là các hệ số a_k . Bộ giải mã âm thanh theo phương án thứ hai bao gồm bộ ước lượng mức nhiều âm được tạo cấu hình để ước lượng mức nhiều âm cho khung hiện thời sử dụng hệ số dự báo tuyến tính của ít nhất một khung có trước để thu được thông tin mức nhiều âm và bộ chèn nhiều âm được tạo cấu hình để bổ sung nhiều âm cho khung hiện thời trong sự phụ thuộc vào thông tin mức nhiều âm được cung cấp bởi bộ ước lượng mức nhiều âm. Bộ chèn nhiều âm được tạo cấu hình để bổ sung nhiều âm cho khung hiện thời với điều kiện là tốc độ bit của thông tin âm thanh được mã hóa nhỏ hơn 0,5 bit cho mỗi

mẫu. Hơn nữa, bộ chèn nhiễu âm có thể được tạo cấu hình để bổ sung nhiễu âm cho khung hiện thời với điều kiện là khung hiện thời là khung tiếng nói. Do đó, một lần nữa, nhiễu âm có thể được bổ sung cho khung hiện thời để cải thiện chất lượng âm nói chung của thông tin âm thanh được giải mã mà có thể bị làm suy yếu do mã hóa các thành phần lạ, đặc biệt đối với nhiễu âm nền của thông tin tiếng nói. Khi mức nhiễu âm của nhiễu âm được điều chỉnh trong xem xét mức nhiễu âm của ít nhất một khung âm thanh có trước, chất lượng âm nói chung có thể được cải thiện mà không phụ thuộc vào thông phụ trong dòng bit. Do đó, lượng dữ liệu để được truyền với dòng bit có thể bị làm giảm.

Fig.4 thể hiện phương pháp thứ hai để thực hiện giải mã âm thanh theo sáng chế mà có thể được thực hiện bởi bộ giải mã âm thanh theo Fig.3. Các chi tiết kỹ thuật của bộ giải mã âm thanh đã vẽ trong Fig.3 được mô tả cùng với các đặc tính phương pháp. Theo Fig.4, bộ giải mã âm thanh được tạo cấu hình để đọc dòng bit để xác định loại khung của khung hiện thời. Hơn nữa, bộ giải mã âm thanh bao gồm bộ xác định loại khung để xác định loại khung của khung hiện thời, bộ xác định loại khung được tạo cấu hình để nhận biết loại khung của khung hiện thời hoặc là tiếng nói hoặc là âm thanh thông thường, sao cho việc ước lượng mức nhiễu âm có thể được thực hiện phụ thuộc vào loại khung của khung hiện thời. Nói chung, bộ giải mã âm thanh được làm thích ứng để tính thông tin thứ nhất biểu diễn kích thích không tạo hình phổ của khung hiện thời và để tính thông tin thứ hai liên quan đến việc định tỉ lệ phổ của khung hiện thời để tính thương số của thông tin thứ nhất với thông tin thứ hai để thu được thông tin mức nhiễu âm. Ví dụ, nếu loại khung là ACELP, là loại khung tiếng nói, bộ giải mã âm thanh giải mã tín hiệu kích thích của khung hiện thời và tính căn bậc hai trung bình bình phương của nó e_{rms} cho khung hiện thời f từ việc biểu diễn miền thời gian của tín hiệu kích thích. Điều này có nghĩa là, bộ giải mã âm thanh được làm thích ứng để giải mã tín hiệu kích thích của khung hiện thời và để tính căn bậc hai trung bình bình phương e_{rms} của nó từ việc biểu diễn miền thời gian của khung hiện thời như thông tin thứ nhất để thu được thông tin mức nhiễu âm với điều kiện là khung hiện thời thuộc loại tiếng nói. Trong trường hợp khác, nếu loại khung là MDCT hoặc DTX, là loại khung âm thanh thông thường, bộ giải mã âm thanh giải mã tín hiệu kích thích của khung hiện thời và tính căn bậc hai trung bình bình phương của nó e_{rms} cho khung hiện

thời f từ tương đương biểu diễn miền thời gian của tín hiệu kích thích. Điều này có nghĩa là, bộ giải mã âm thanh được làm thích ứng để giải mã kích thích MDCT không tạo hình của khung hiện thời và để tính căn bậc hai trung bình bình phương của nó e_{rms} từ việc biểu diễn miền phổ của khung hiện thời như thông tin thứ nhất để thu được thông tin mức nhiễu âm với điều kiện là khung hiện thời thuộc loại âm thanh thông thường. Cách việc này được thực hiện được mô tả chi tiết trong WO 2012/110476 A1. Hơn nữa, Fig.9 thể hiện biểu đồ minh họa cách tương đương bộ lọc LPC được xác định từ phổ lũy thừa MDCT. Trong khi thang đo được vẽ là thang đo Bark, các tương đương hệ số LPC có thể cũng được thu từ thang đo tuyến tính. Đặc biệt khi chúng được thu từ thang đo tuyến tính, các tương đương hệ số LPC được tính toán rất giống với các tương đương hệ số LPC được tính toán từ sự biểu diễn miền thời gian của cùng một khung, ví dụ khi được mã hóa trong ACELP.

Ngoài ra, bộ giải mã âm thanh theo Fig.3, như được minh họa bởi biểu đồ phương pháp của Fig.4, được làm thích ứng để tính mức đỉnh p của hàm truyền của bộ lọc LPC của khung hiện thời như thông tin thứ hai, do đó sử dụng hệ số dự báo tuyến tính để thu được thông tin mức nhiễu âm với điều kiện là khung hiện thời thuộc loại tiếng nói. Điều đó có nghĩa là, bộ giải mã âm thanh tính toán mức đỉnh p của hàm truyền của bộ lọc phân tích LPC của khung hiện thời f theo công thức $p = \sum |a_k|$, trong đó a_k là hệ số dự báo tuyến tính với $k = 0 \dots 15$. Nếu khung là khung âm thanh thông thường, các tương đương hệ số LPC được thu từ sự biểu diễn miền phổ của khung hiện thời, như được thể hiện trong Fig.9 và được mô tả trong WO 2012/110476 A1 và được mô tả ở trên. Như quan sát trong Fig.4, sau khi tính mức đỉnh p , mức tối thiểu phổ m_f của khung hiện thời f được tính toán bằng cách chia e_{rms} cho p . Do đó, bộ giải mã âm thanh được làm thích ứng để tính thông tin thứ nhất biểu diễn kích thích không tạo hình phổ của khung hiện thời, trong phương án này e_{rms} , và thông tin thứ hai liên quan đến việc định tỉ lệ phổ của khung hiện thời, trong phương án này mức đỉnh p , để tính thương số của thông tin thứ nhất với thông tin thứ hai để thu được thông tin mức nhiễu âm. Mức tối thiểu phổ của khung hiện thời sau đó được xếp vào hàng bộ ước lượng mức nhiễu âm, bộ giải mã âm thanh được làm thích ứng để xếp vào hàng thương số được thu từ khung âm thanh hiện tại trong bộ ước lượng mức nhiễu âm bất kể loại khung nào và bộ ước lượng mức nhiễu âm bao gồm việc lưu trữ mức nhiễu âm cho hai

hoặc nhiều thương số, trong trường hợp này các mức tối thiểu phổ m_f , được thu từ các khung âm thanh khác nhau. Cụ thể hơn, bộ lưu mức nhiều âm có thể lưu các thương số từ 50 khung để ước lượng mức nhiều âm. Hơn nữa, bộ ước lượng mức nhiều âm được làm thích ứng để ước lượng mức nhiều âm trên cơ sở phân tích thống kê hai hoặc nhiều thương số của các khung âm thanh khác nhau, do tập hợp các mức tối thiểu phổ m_f . Các bước để tính thương số m_f được vẽ chi tiết trong Fig.7, minh họa các bước tính toán cần thiết. Trong phương án thứ hai, bộ ước lượng mức nhiều âm hoạt động dựa trên các số liệu thống kê tối thiểu như đã biết trong tài liệu tham khảo [3]. Nhiều âm được định tỉ lệ theo mức nhiều âm được ước lượng của khung hiện thời dựa trên các số liệu thống kê tối thiểu và sau đó được bổ sung cho khung hiện thời nếu khung hiện thời là khung tiếng nói. Cuối cùng, khung hiện thời được chỉnh giảm (không được thể hiện trong Fig.4). Do đó, phương án thứ hai này cũng cho phép bỏ qua thông tin phụ để làm đầy nhiều âm, cho phép làm giảm lượng dữ liệu được truyền với dòng bit. Theo đó, chất lượng âm của thông tin âm thanh có thể được cải thiện bằng cách nâng cao nhiều âm nền trong giai đoạn giải mã mà không làm tăng tốc độ dữ liệu. Lưu ý rằng vì những biến đổi thời gian/tần số là không cần thiết và vì bộ ước lượng mức nhiều âm chỉ chạy một lần cho mỗi khung (không trên các băng con phức tạp), việc làm đầy nhiều âm đã mô tả thể hiện tính phức tạp thấp trong khi có thể cải thiện việc mã hoá tốc độ bit thấp của tiếng nói có nhiều âm.

Fig.5 thể hiện phương án thứ ba của bộ giải mã âm thanh theo sáng chế.

Bộ giải mã âm thanh được làm thích ứng để cung cấp thông tin âm thanh được giải mã trên cơ sở thông tin âm thanh được mã hóa. Bộ giải mã âm thanh được tạo cấu hình để sử dụng bộ mã hóa thể dựa trên LD-USAC để giải mã thông tin âm thanh được mã hóa. Thông tin âm thanh đã mã hoá bao gồm các hệ dự báo tuyến tính (LPC), mà có thể được ký hiệu riêng như các hệ số a_k . Bộ giải mã âm thanh bao gồm bộ điều chỉnh độ nghiêng được tạo cấu hình để điều chỉnh độ nghiêng của nhiều âm sử dụng các hệ số dự báo tuyến tính của khung hiện thời để thu được thông tin độ nghiêng và bộ ước lượng mức nhiều âm được tạo cấu hình để ước lượng mức nhiều âm cho khung hiện thời sử dụng hệ số dự báo tuyến tính của ít nhất một khung có trước để thu được thông tin mức nhiều âm. Hơn nữa, bộ giải mã âm thanh bao gồm bộ chèn nhiều âm được tạo cấu hình để bổ sung nhiều âm cho khung hiện thời phụ thuộc vào thông tin

độ nghiêng được thu bởi bộ tính toán độ nghiêng và phụ thuộc vào thông tin mức nhiều âm được cung cấp bởi bộ ước lượng mức nhiều âm. Do đó, nhiều âm có thể được bổ sung cho khung hiện thời để cải thiện chất lượng âm nói chung của thông tin âm thanh được giải mã mà có thể bị làm suy yếu do mã hoá các thành phần lạ, đặc biệt đối với nhiều âm nền của thông tin tiếng nói, phụ thuộc vào thông tin độ nghiêng được thu bởi bộ tính toán độ nghiêng và phụ thuộc vào thông tin mức nhiều âm được cung cấp bởi bộ ước lượng mức nhiều âm. Trong phương án này, bộ tạo nhiều âm ngẫu nhiên (không được thể hiện) mà được chứa bởi bộ giải mã âm thanh tạo ra nhiều âm trắng phổ, mà sau đó cả hai được định tỉ lệ theo thông tin mức nhiều âm và được tạo hình sử dụng độ nghiêng được suy ra g, như được mô tả lúc đầu.

Fig.6 thể hiện phương pháp thứ ba để thực hiện giải mã âm thanh theo sáng chế mà có thể được thực hiện bởi bộ giải mã âm thanh theo Fig.5. Dòng bit được đọc và bộ xác định loại khung, được gọi là bộ phát hiện loại khung, xác định khung hiện thời hoặc là khung tiếng nói (ACELP) hoặc là khung âm thanh thông thường (TCX/MDCT). Bất kể thuộc loại khung nào, đoạn đầu khung được giải mã và được làm phẳng phổ, tín hiệu kích thích không tạo hình trong miền cảm giác được giải mã. Trong trường hợp của khung tiếng nói, tín hiệu kích thích này là sự kích thích miền thời gian, như được mô tả lúc đầu. Nếu khung là khung âm thanh thông thường, phần còn lại miền MDCT được giải mã (miền phổ). Việc biểu diễn miền thời gian và biểu diễn miền phổ được sử dụng một cách tương ứng để ước lượng mức nhiều âm như được minh hoạ trong Fig.7 và được mô tả lúc đầu, sử dụng các hệ số LPC cũng như được sử dụng để giải mã dòng bit thay vì sử dụng bất kỳ thông tin phụ nào hoặc các hệ số LPC bổ sung. Thông tin nhiều âm của cả hai loại khung được xếp vào hàng để điều chỉnh độ nghiêng và mức nhiều âm của nhiều âm để được bổ sung cho khung hiện thời với điều kiện là khung hiện thời là khung tiếng nói. Sau khi bổ sung nhiều âm cho khung tiếng nói ACELP (áp dụng làm đầy nhiều âm ACELP) khung tiếng nói ACELP được chỉnh giảm bởi IIR và các khung tiếng nói và các khung âm thanh thông thường được tổ hợp trong tín hiệu thời gian, biểu diễn thông tin âm thanh được giải mã. Hiệu quả thông cao độ dốc của việc chỉnh giảm trên phổ của nhiều âm đã bổ sung được mô tả bởi các hình vẽ đã chèn nhỏ I, II và III trong Fig.6.

Nói cách khác, theo Fig.6, hệ thống làm đầy nhiễu âm ACELP được mô tả ở trên đã được thực hiện trong bộ giải mã LD-USAC (EVS), biến thiên độ trễ thấp của xHE-AAC trong tài liệu tham khảo [6] mà có thể chuyển giữa mã hoá ACELP (tiếng nói) với mã hóa MDCT (âm nhạc/nhiều âm) trên cơ sở mỗi khung. Quy trình chèn theo Fig.6 được tóm tắt như sau:

1. Dòng bit được đọc, và nó được xác định dù khung hiện thời là khung ACELP hoặc khung MDCT hoặc khung DTX. Bất kể loại khung nào, tín hiệu kích thích được làm phẳng phổ (trong miền cảm giác) được giải mã và được sử dụng để cập nhật ước lượng mức nhiễu âm như được mô tả chi tiết ở dưới. Sau đó tín hiệu được khôi phục hoàn toàn lên đến sự chỉnh giảm, mà là bước cuối cùng.
2. Nếu khung là ACELP được mã hóa, độ nghiêng (hình dạng phổ tổng thể) để chèn nhiễu âm được tính bằng phép phân tích LPC bậc một các hệ số bộ lọc LPC. Độ nghiêng được suy ra từ độ khuếch đại g của hệ số LPC thứ 16 a_k mà độ khuếch đại g được xác định bởi

$$g = \Sigma [a_k \cdot a_{k+1}] / \Sigma [a_k \cdot a_k].$$
3. Nếu khung là ACELP được mã hóa, mức tạo hình nhiễu âm và độ nghiêng được sử dụng để thực hiện bổ sung nhiễu âm lên trên khung đã giải mã: bộ tạo nhiễu âm ngẫu nhiên tạo tín hiệu nhiễu âm trắng phẳng, mà sau đó được định tỉ lệ và được tạo hình sử dụng độ nghiêng được suy ra g .
4. Tín hiệu nhiễu âm đã định mức và đã tạo hình cho khung ACELP được bổ sung lên trên tín hiệu đã giải mã chỉ trước bước lọc chỉnh giảm cuối. Vì sự chỉnh giảm là IIR bậc một tăng lên các tần số thấp, việc này cho phép tính phức tạp thấp, việc lọc thông cao IIR dốc của nhiễu âm đã bổ sung, như trong Fig.6, tránh các thành phần lạ nhiễu âm có thể nghe được ở những tần số thấp.

Việc ước lượng mức nhiễu âm trong bước 1 được thực hiện bằng cách tính căn bậc hai e_{rms} của tín hiệu kích thích cho khung hiện thời (hoặc trong trường hợp kích thích miền MDCT tương đương miền thời gian, có nghĩa là e_{rms} sẽ được tính cho khung đó nếu khung đó là khung ACELP) và sau đó chia nó cho mức đỉnh p của hàm truyền của bộ lọc phân tích LPC. Cách tính này mang đến mức m_f của mức tối thiểu

phổ của khung f như trong Fig.7. m_f được xếp vào hàng sau cùng trong bộ ước lượng mức nhiễu âm hoạt động dựa trên ví dụ các số liệu thống kê tối thiểu trong tài liệu tham khảo [3]. Lưu ý rằng vì các biến đổi thời gian/tần số là không cần thiết và vì bộ ước lượng mức chỉ chạy một lần cho mỗi khung (không ở trên các băng con phức tạp), hệ thống làm đầy nhiễu âm CELP đã mô tả thể hiện tính phức tạp rất thấp trong khi có thể cải thiện việc mã hóa tốc độ bit thấp của tiếng nói có nhiễu âm.

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

Tín hiệu âm thanh đã mã hóa theo sáng chế có thể được lưu trên vật ghi lưu trữ dạng số hoặc có thể được truyền dẫn trên môi trường truyền dẫn như môi trường truyền dẫn không dây hoặc môi trường truyền dẫn có dây như mạng liên mạng.

Phụ thuộc vào các yêu cầu thực hiện nhất định, các phương án của sáng chế có thể được thực hiện dưới dạng phần cứng hoặc phần mềm. Việc tiến hành có thể được thực hiện sử dụng vật ghi lưu trữ dạng số, ví dụ đĩa mềm, DVD, Blu-Ray, CD, ROM, PROM, EPROM, EEPROM hoặc bộ nhớ FLASH, có các tín hiệu điều khiển có thể đọc điện tử được lưu trên đó, mà phối hợp (hoặc có khả năng phối hợp) với hệ thống máy tính có thể lập trình được để thực hiện phương pháp tương ứng. Do đó, vật ghi lưu trữ dạng số có thể đọc được bằng máy tính.

Một số phương án theo sáng chế bao gồm vật mang dữ liệu có các tín hiệu điều khiển có thể đọc được theo cách 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.

Thông thường, các phương án của sáng chế có thể được thực hiện dưới dạng sản phẩm chương trình máy tính với mã chương trình, mã chương trình có tác dụng 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 trên vật mang có thể đọc bằng máy.

Các phương án khác bao gồm chương trình máy tính dùng để thực hiện một trong số các phương pháp được mô tả ở đây, được lưu trong vật mang có thể đọc 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 phương pháp theo sáng chế là, vật mang dữ liệu (hoặc vật ghi lưu trữ dạng số, hoặc vật ghi có thể đọc được bằng máy tính) bao gồm, chương trình máy tính được ghi lên đó, chương trình máy tính này dùng để thực hiện một trong số các phương pháp được mô tả ở đây. Vật mang dữ liệu, vật ghi lưu trữ dạng số hoặc vật ghi được ghi thường là hữu hình và/hoặc không bền vững.

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

Phương án khác bao gồm phương tiện xử lý, ví dụ, máy tính, hoặc thiết bị logic có thể 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 đó, để thực hiện một trong số các phương pháp được mô tả ở đây.

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

máy tính, thiết bị di động, thiết bị nhớ hoặc tương tự. Các thiết bị hoặc hệ thống, ví dụ, bao gồm máy chủ tập tin để truyền chương trình máy tính đến bộ nhận.

Trong một số phương án, thiết bị logic có thể lập trình được (ví dụ, mảng cổng có thể lập trình được bằng 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 có thể lập trình được bằng 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, tốt nhất là các phương pháp được thực hiện bởi các thiết bị phần cứng bất kỳ.

Thiết bị được mô tả ở đây có thể được thực hiện sử dụng thiết bị phần cứng, hoặc sử dụng máy tính, hoặc sử dụng tổ hợp thiết bị phần cứng với máy tính.

Các phương pháp được mô tả ở đây có thể được thực hiện sử dụng thiết bị phần cứng, hoặc sử dụng máy tính, hoặc sử dụng tổ hợp thiết bị phần cứng với máy tính.

Các phương án được mô tả ở trên chỉ mang tính minh họa cho các nguyên lý của sáng chế. Cần hiểu là những biến đổi và những biến thể của các cách lắp đặt và các chi tiết được mô tả ở đây sẽ là rõ ràng với những người có trình độ trong lĩnh vực kỹ thuật. Do đó, mục đích là chỉ bị giới hạn bởi phạm vi của các điểm yêu cầu bảo hộ mà không bị giới hạn bởi các chi tiết cụ thể được thể hiện bằng bản mô tả và việc giải thích các phương án ở đây.

YÊU CẦU BẢO HỘ

1. Bộ giải mã âm thanh cung cấp thông tin âm thanh được giải mã trên cơ sở thông tin âm thanh được mã hóa bao gồm các hệ số dự báo tuyến tính (LPC),

bộ giải mã âm thanh bao gồm:

- bộ điều chỉnh độ nghiêng được tạo cấu hình để điều chỉnh độ nghiêng của nhiều âm nền tùy thuộc vào thông tin độ nghiêng, trong đó bộ điều chỉnh độ nghiêng được tạo cấu hình để sử dụng các hệ số dự báo tuyến tính của khung hiện thời để thu được thông tin độ nghiêng; và

- lõi bộ giải mã được tạo cấu hình để giải mã thông tin âm thanh của khung hiện thời sử dụng các hệ số dự báo tuyến tính của khung hiện thời để thu được tín hiệu đầu ra bộ mã hóa lõi đã giải mã; và

- bộ chèn nhiều âm được tạo cấu hình để thêm nhiều âm nền được điều chỉnh vào khung hiện thời, để thực hiện làm đầy nhiều âm;

đặc trưng ở chỗ:

bộ điều chỉnh độ nghiêng được tạo cấu hình để sử dụng kết quả của phép phân tích bậc một các hệ số dự báo tuyến tính của khung hiện thời để thu được thông tin độ nghiêng, và

trong đó bộ điều chỉnh độ nghiêng được tạo cấu hình để thu thông tin độ nghiêng sử dụng việc tính toán độ khuếch đại g của các hệ số dự báo tuyến tính của khung hiện thời như phép phân tích bậc một,

trong đó:

$$g = \Sigma [a_k \cdot a_{k+1}] / \Sigma [a_k \cdot a_k],$$

trong đó a_k là hệ số dự báo tuyến tính của khung hiện thời, nằm ở chỉ số LPC k .

2. Bộ giải mã âm thanh theo điểm 1, trong đó bộ giải mã âm thanh bao gồm bộ xác định loại khung để xác định loại khung của khung hiện thời, bộ xác định loại khung được tạo cấu hình để kích hoạt bộ điều chỉnh độ nghiêng để điều chỉnh độ nghiêng của nhiều âm nền khi loại khung của khung hiện thời được phát hiện là thuộc loại tiếng nói.

3. Bộ giải mã âm thanh theo điểm bất kỳ trong số các điểm nêu trên, trong đó bộ giải mã âm thanh còn bao gồm:

- bộ ước lượng mức nhiễu âm được tạo cấu hình để ước lượng mức nhiễu âm cho khung hiện thời sử dụng nhiều hệ số dự báo tuyến tính của ít nhất một khung có trước để thu được thông tin mức nhiễu âm;

- trong đó bộ chèn nhiễu âm được tạo cấu hình để thêm nhiễu âm nền vào khung hiện thời tùy thuộc vào thông tin mức nhiễu âm được cung cấp bởi bộ ước lượng mức nhiễu âm;

trong đó bộ giải mã âm thanh được làm thích ứng để giải mã tín hiệu kích thích của khung hiện thời và để tính căn bậc hai trung bình bình phương e_{rms} của nó;

trong đó bộ giải mã âm thanh được làm thích ứng để tính mức đỉnh p của hàm truyền của bộ lọc LPC của khung hiện thời;

trong đó bộ giải mã âm thanh được làm thích ứng để tính mức tối thiểu phổ m_f của khung âm thanh hiện thời bằng cách tính thương số của căn bậc hai trung bình bình phương e_{rms} với mức đỉnh p để thu được thông tin mức nhiễu âm;

trong đó bộ ước lượng mức nhiễu âm được làm thích ứng để ước lượng mức nhiễu âm trên cơ sở hai hoặc nhiều hơn hai thương số của các khung âm thanh khác nhau.

4. Bộ giải mã âm thanh theo điểm bất kỳ trong số các điểm nêu trên, trong đó bộ giải mã âm thanh bao gồm bộ lọc chỉnh giảm để chỉnh giảm khung hiện thời, bộ giải mã âm thanh được làm thích ứng để áp dụng bộ lọc chỉnh giảm trên khung hiện thời sau khi bộ chèn nhiễu âm đã bổ sung nhiễu âm cho khung hiện thời.

5. Bộ giải mã âm thanh theo điểm bất kỳ trong số các điểm nêu trên, trong đó bộ giải mã âm thanh bao gồm bộ tạo nhiễu âm, bộ tạo nhiễu âm được làm thích ứng để tạo nhiễu âm để thêm vào khung hiện thời bởi bộ chèn nhiễu âm.

6. Bộ giải mã âm thanh theo điểm bất kỳ trong số các điểm nêu trên, trong đó bộ giải mã âm thanh bao gồm bộ tạo nhiễu âm được tạo cấu hình để tạo nhiễu âm trắng ngẫu nhiên.

7. Bộ giải mã âm thanh theo điểm bất kỳ trong số các điểm nêu trên, trong đó bộ giải mã âm thanh được tạo cấu hình để sử dụng bộ giải mã dựa trên một hoặc nhiều bộ giải mã trong số các bộ giải mã AMR-WB, G.718 hoặc LD-USAC (EVS) để giải mã thông tin âm thanh được mã hóa.

8. Phương pháp cung cấp thông tin âm thanh được giải mã trên cơ sở thông tin âm thanh được mã hóa bao gồm các hệ số dự báo tuyến tính (LPC),

phương pháp bao gồm các bước:

- điều chỉnh độ nghiêng của nhiễu âm nền tùy thuộc vào thông tin độ nghiêng, trong đó các hệ số dự báo tuyến tính của khung hiện thời được sử dụng để thu được thông tin độ nghiêng; và

- giải mã thông tin âm thanh của khung hiện thời sử dụng các hệ số dự báo tuyến tính của khung hiện thời để thu được tín hiệu đầu ra bộ mã hóa lỗi đã giải mã; và

- thêm nhiễu âm nền được điều chỉnh vào khung hiện thời, để thực hiện làm đầy nhiễu âm;

đặc trưng ở chỗ:

kết quả của phép phân tích bậc một các hệ số dự báo tuyến tính của khung hiện thời được sử dụng để thu được thông tin độ nghiêng, và

trong đó thông tin độ nghiêng thu được bằng cách sử dụng việc tính toán độ khuếch đại g của các hệ số dự báo tuyến tính của khung hiện thời như phép phân tích bậc một,

trong đó:

$$g = \Sigma [a_k \cdot a_{k+1}] / \Sigma [a_k \cdot a_k],$$

trong đó a_k là hệ số dự báo tuyến tính của khung hiện thời, nằm ở chỉ số LPC k .

9. Vật ghi có thể đọc được bằng máy tính chứa chương trình máy tính để thực hiện phương pháp theo điểm 8, trong đó chương trình máy tính chạy trên máy tính.

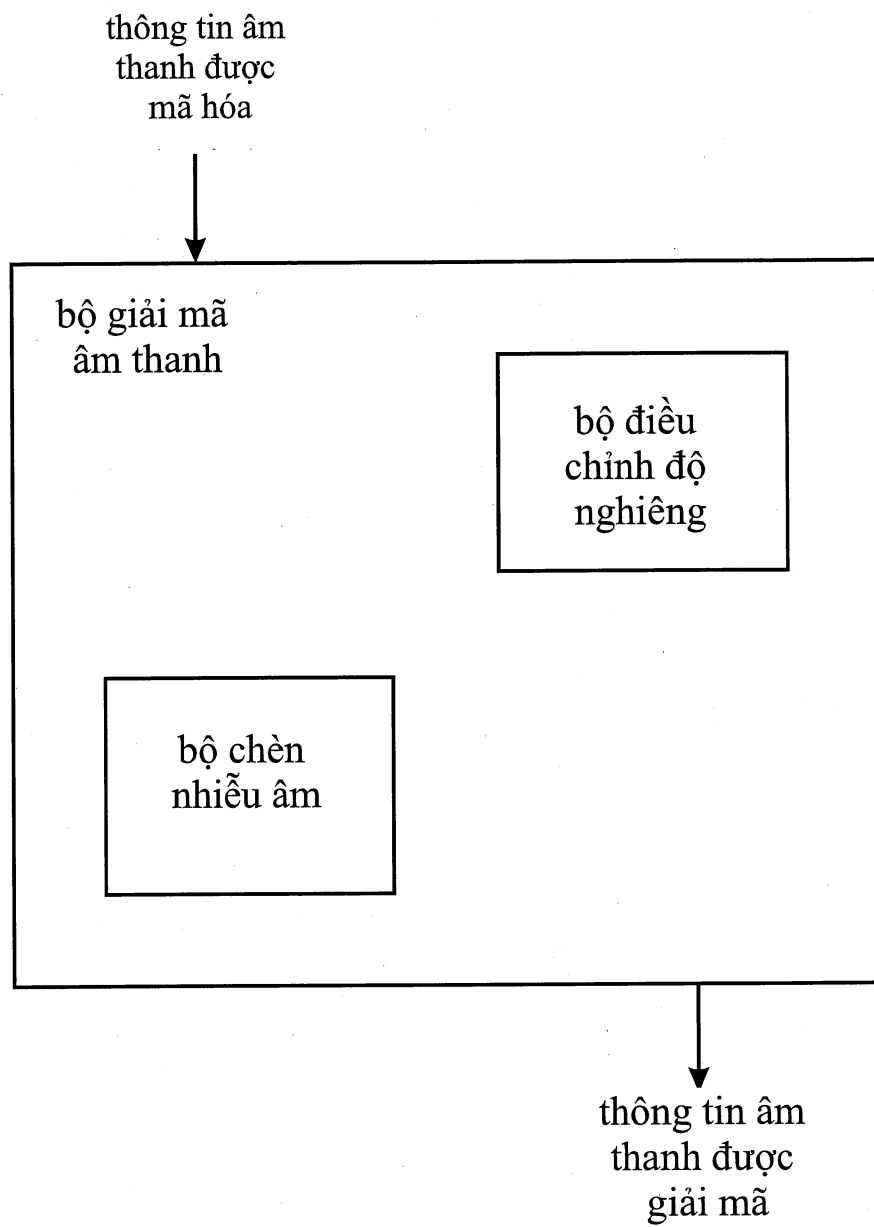


FIG.1

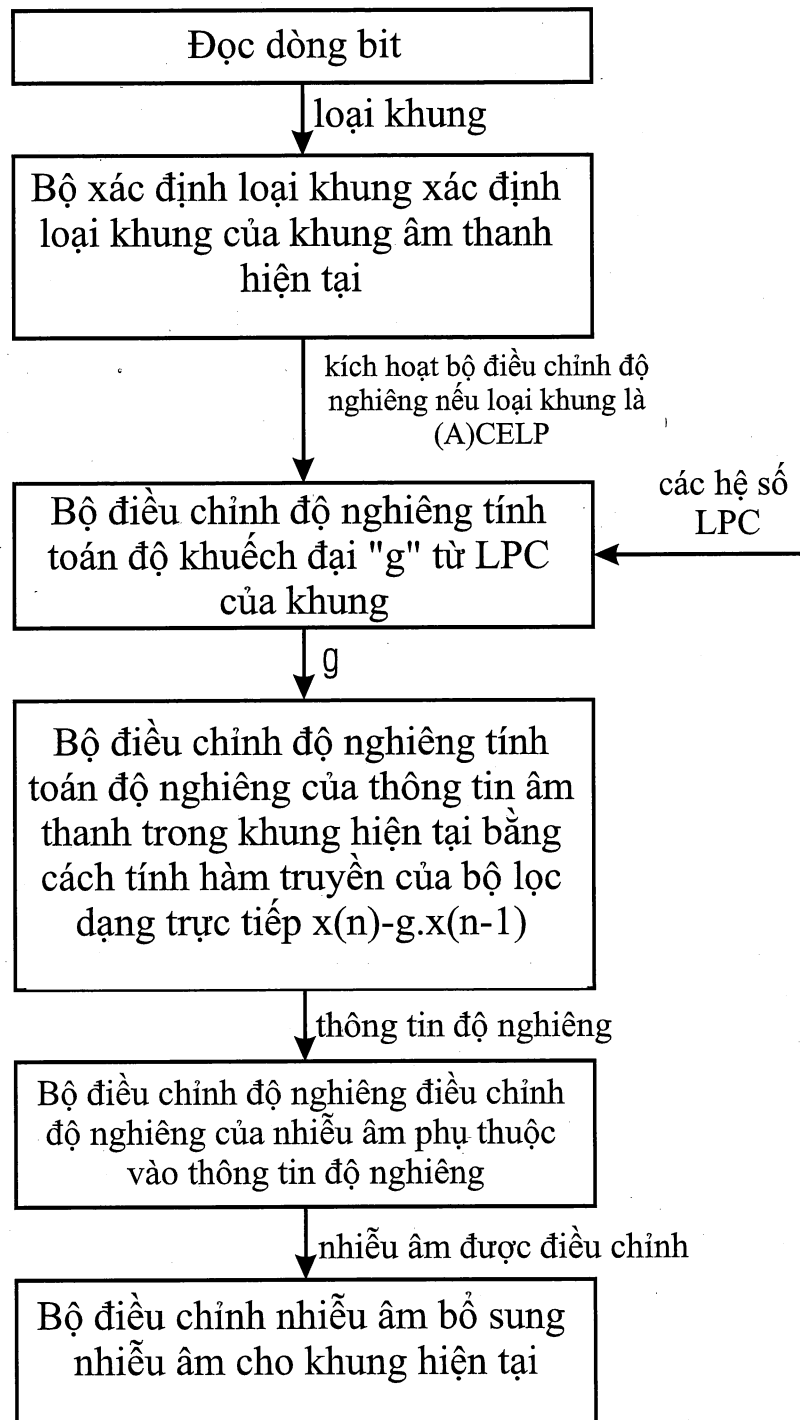


FIG.2

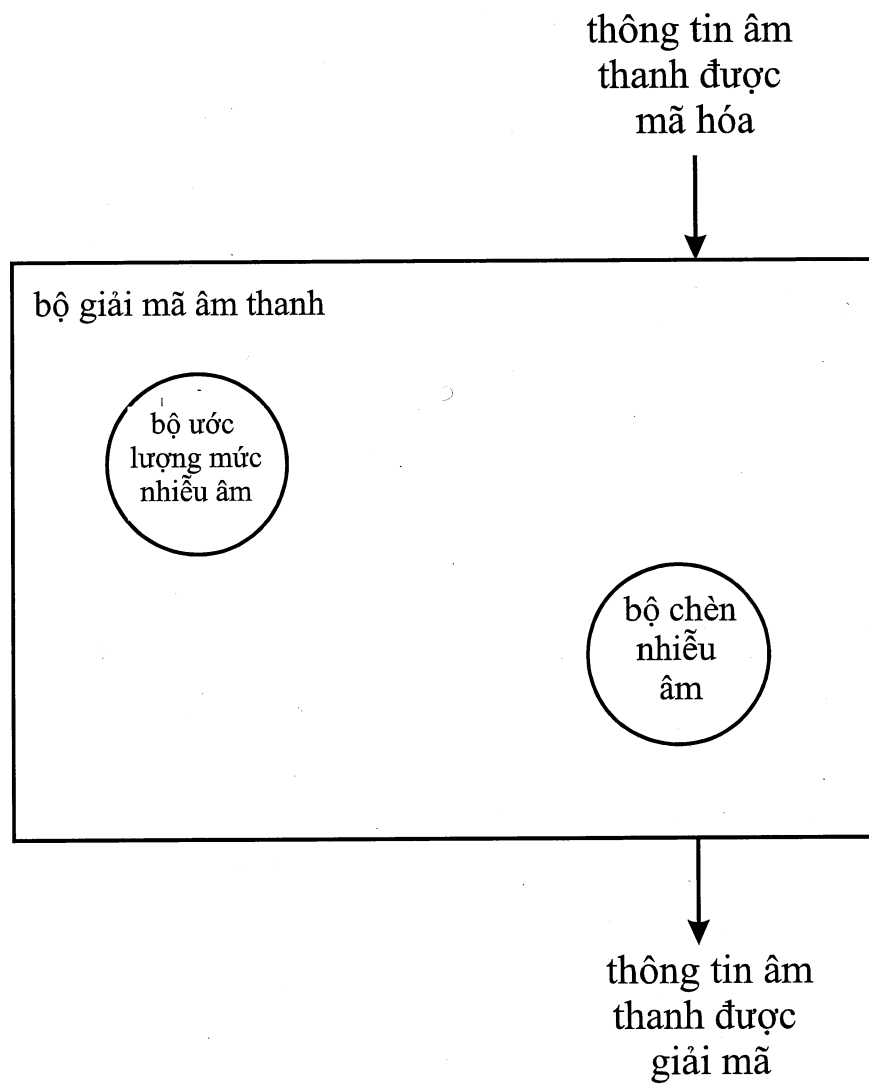


FIG.3

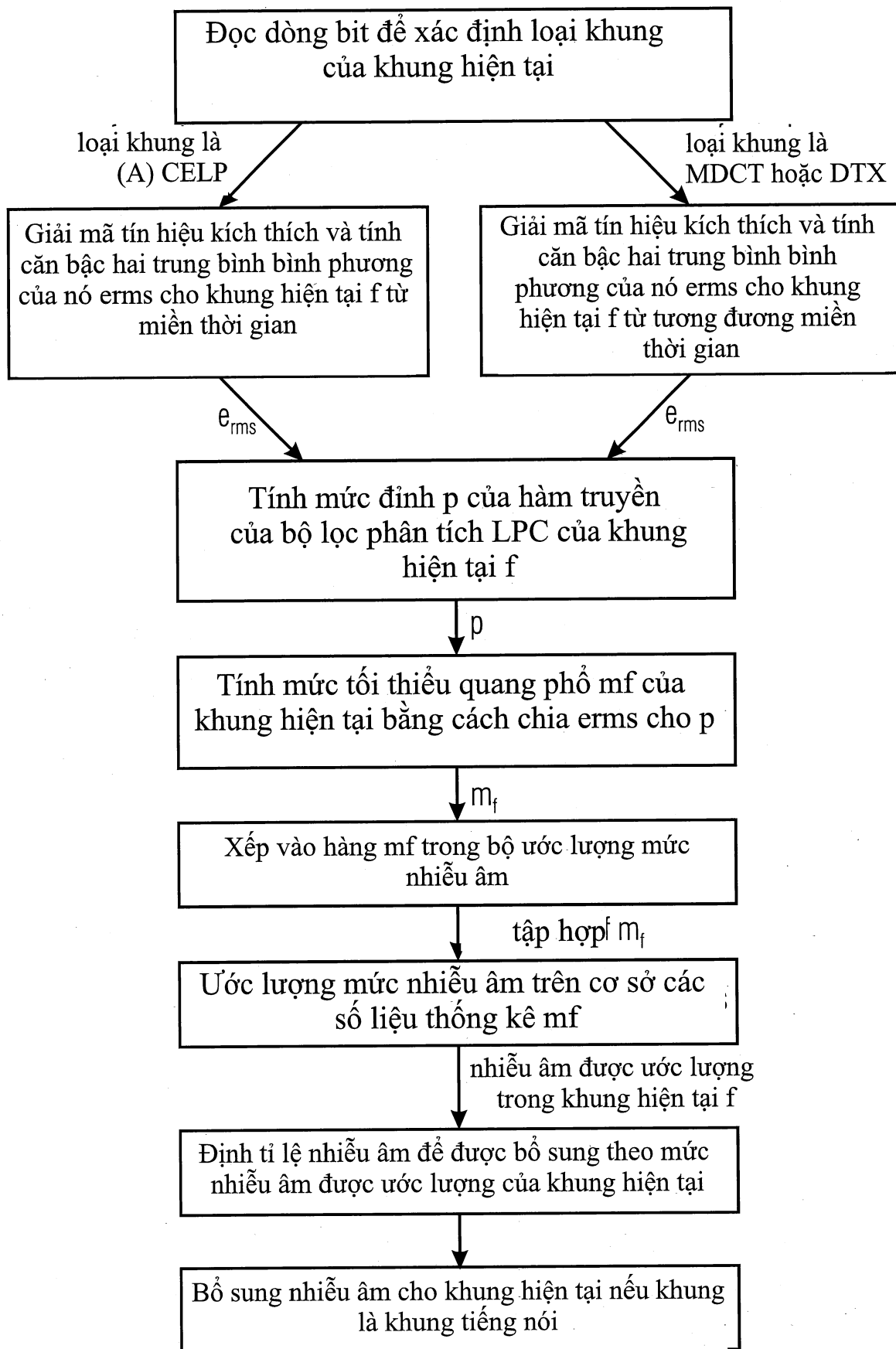


FIG.4

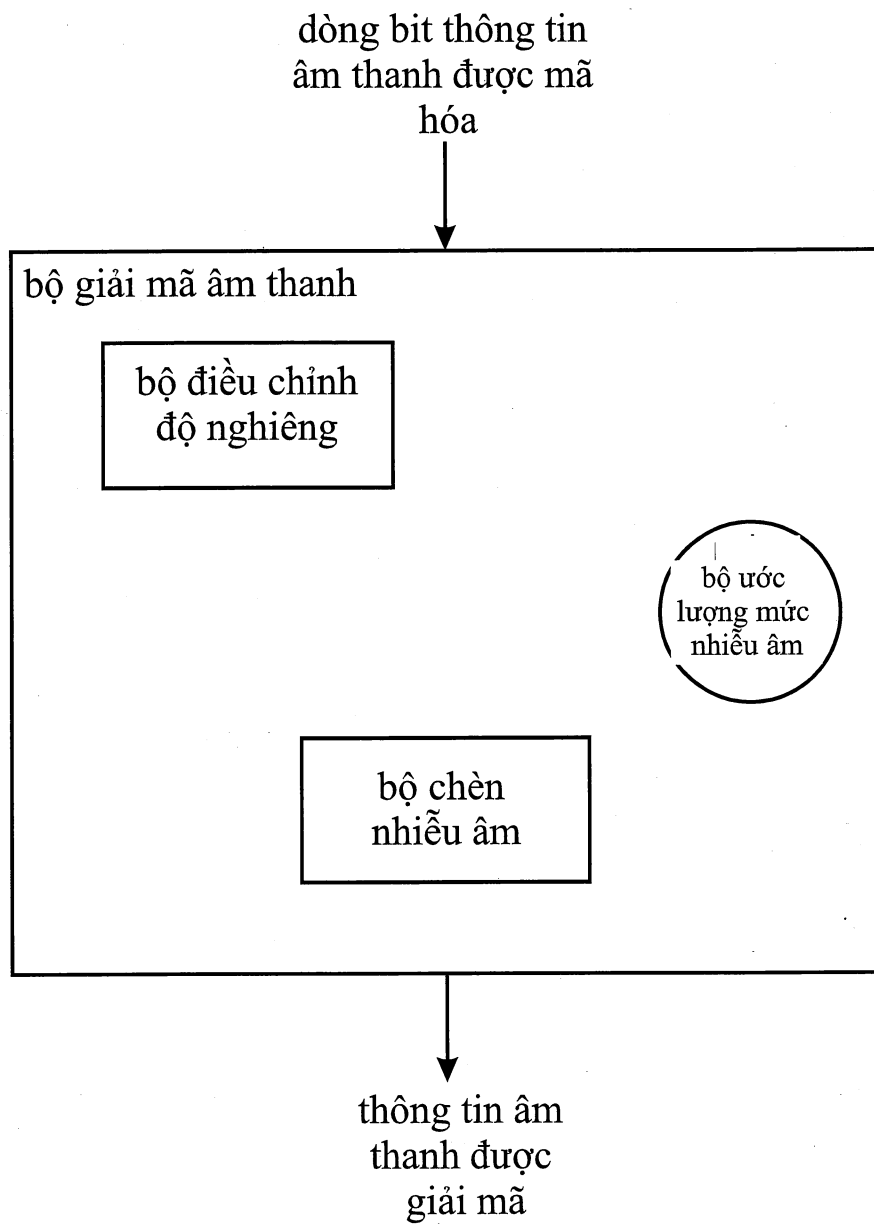


FIG.5

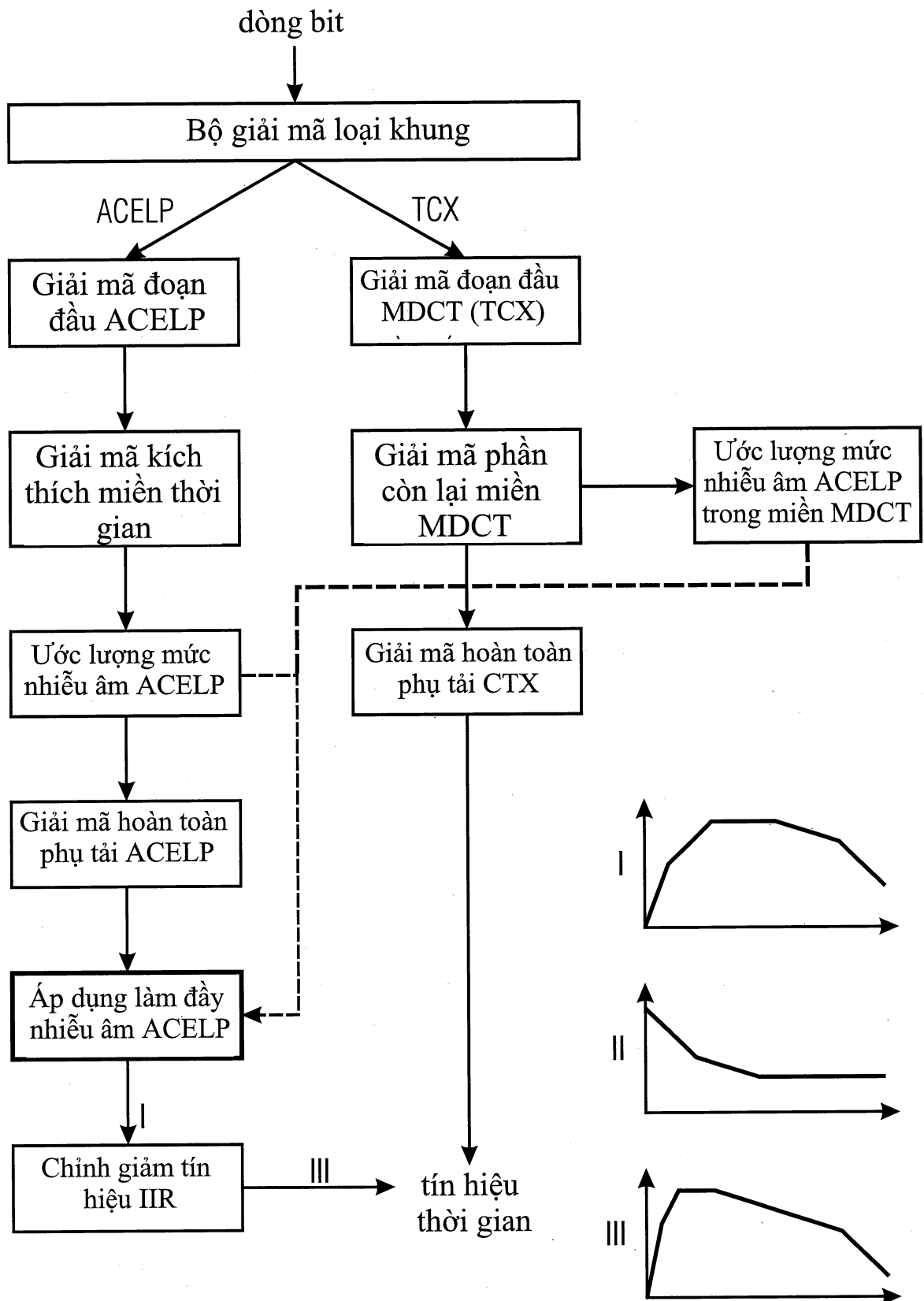


FIG.6

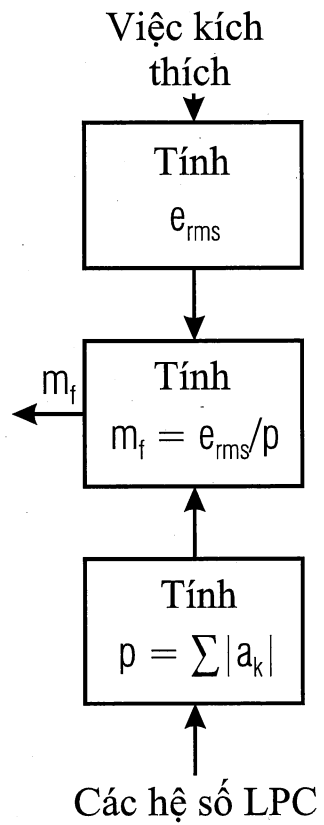


FIG.7A

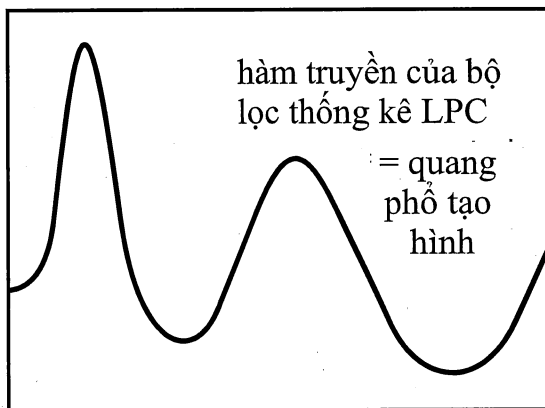


FIG.7B

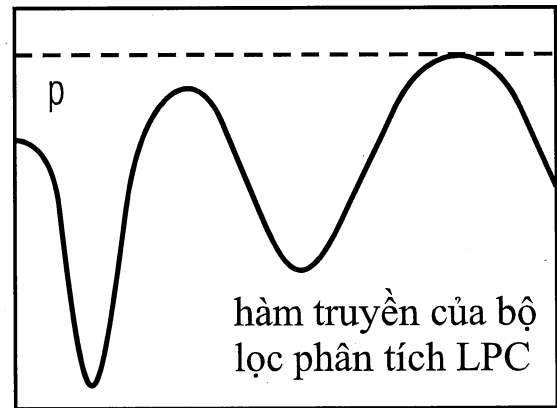
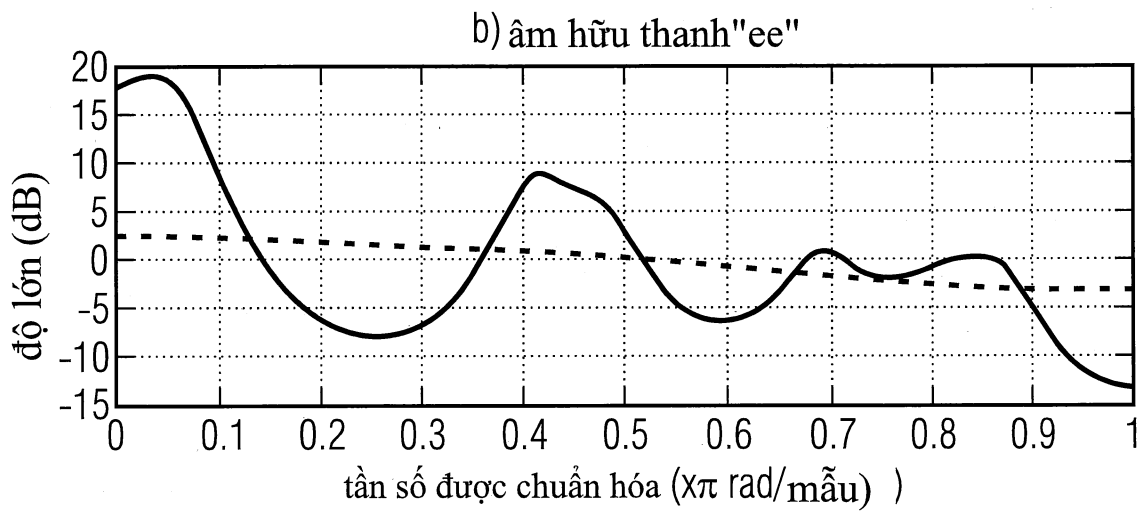
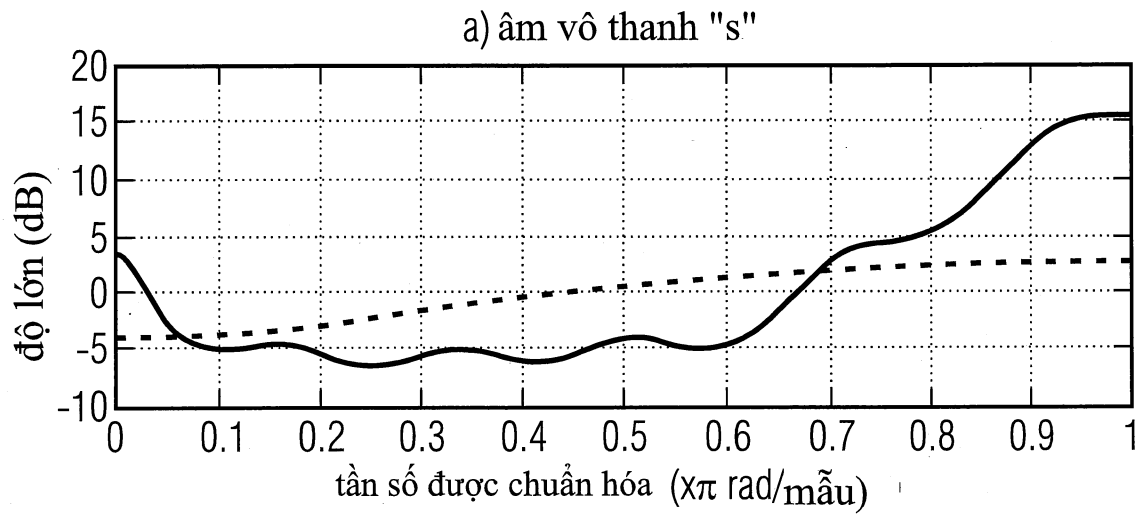
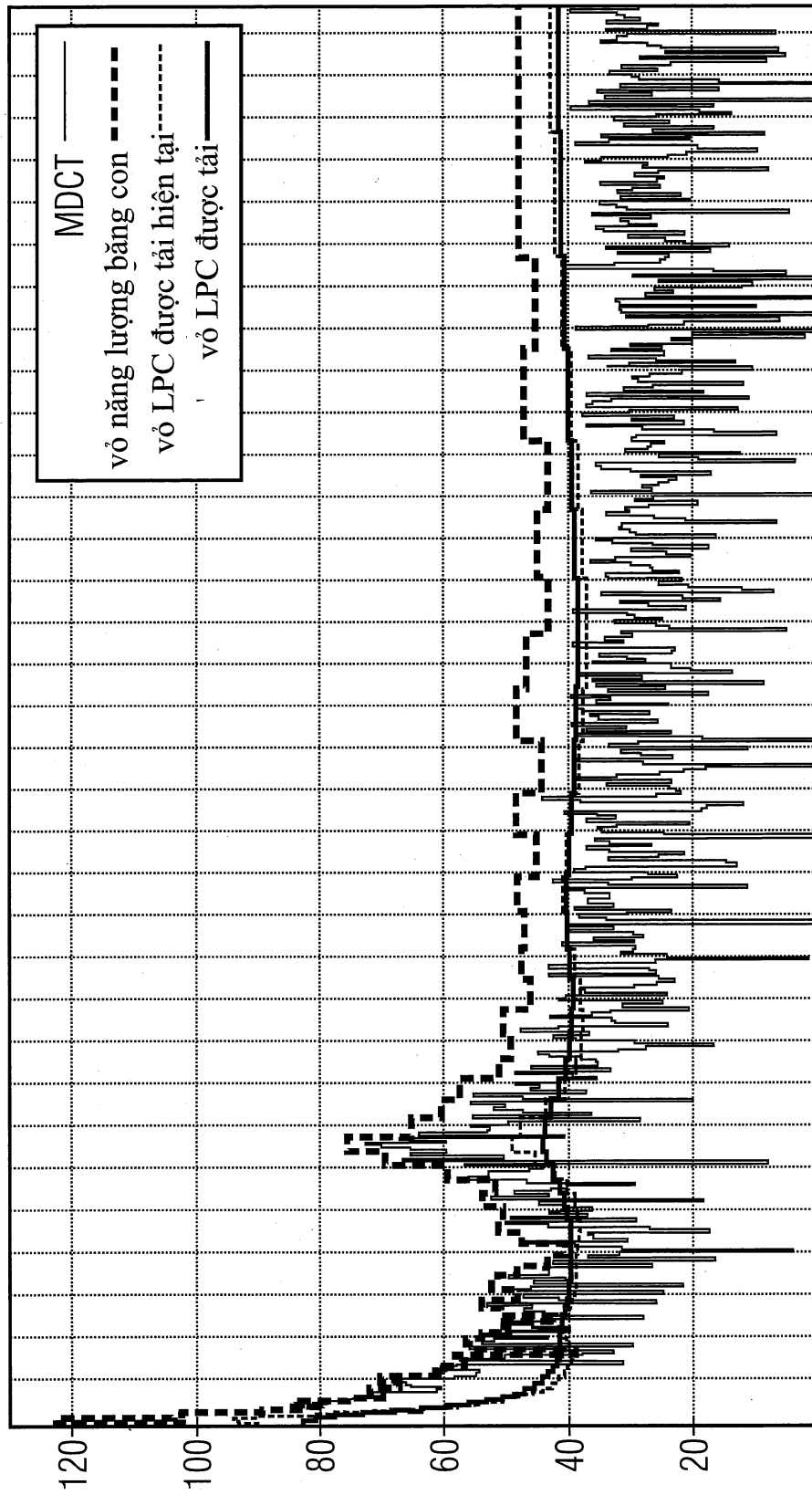


FIG.7C



hàm truyền của(—)bộ lọc LPC
của(- - -) độ nghiêng phổ được suy ra từ các hệ số LPC

FIG.8



việc xác định bộ lọc LPC từ phổ lũy thừa MDCT (bình phương các giá trị MDCT)

FIG.9