



(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)



1-0028670

(51)⁷ G10L 19/08

(13) B

(21) 1-2016-03835

(22) 03/03/2015

(86) PCT/EP2015/054396 03/03/2015

(87) WO 2015/135797 A1 17/09/2015

(30) 14159811.0 14/03/2014 EP; 14182047.2 22/08/2014 EP

(45) 25/06/2021 399

(43) 25/01/2017 346A

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

Hansastraße 27c, 80686 Muenchen, Germany

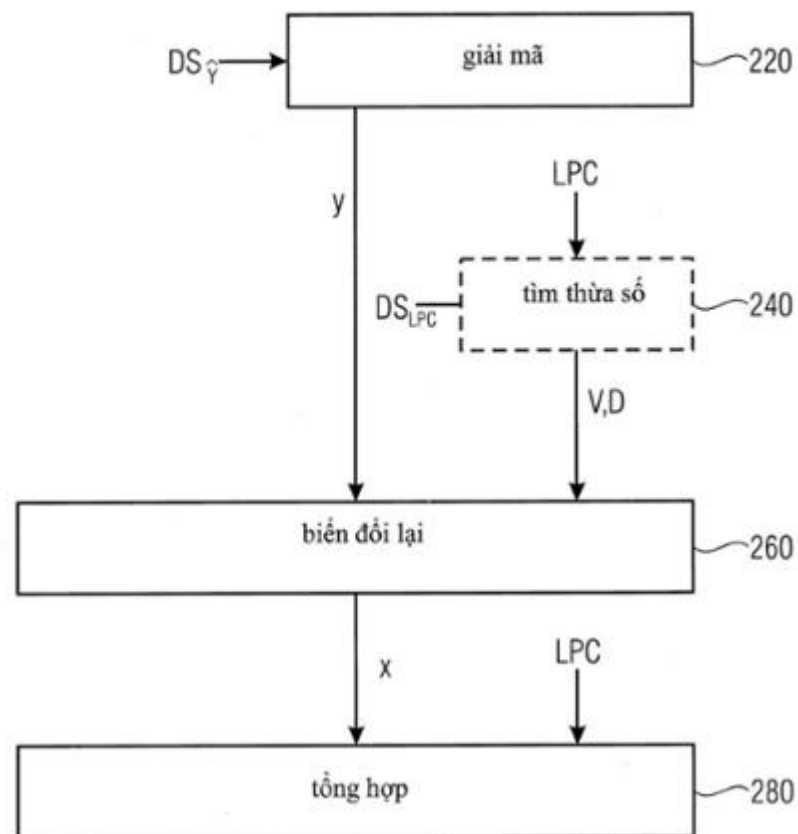
(72) BAECKSTROEM, Tom (FI); FISCHER, Johannes (DE); HELMRICH, Christian
(DE).

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

(54) BỘ MÃ HÓA VÀ PHƯƠNG PHÁP MÃ HÓA TÍN HIỆU ÂM THANH, BỘ GIẢI
MÃ VÀ PHƯƠNG PHÁP GIẢI MÃ DÒNG DỮ LIỆU

(57) Sáng chế đề cập đến bộ mã hóa và phương pháp mã hóa tín hiệu âm thanh, bộ giải mã và phương pháp giải mã dòng dữ liệu. Bộ mã hóa để mã hóa tín hiệu âm thanh thành dòng dữ liệu bao gồm bộ dự báo, bộ tìm thừa số, bộ biến đổi và tầng lượng tử hóa và mã hóa. Bộ dự báo được tạo cấu hình để phân tích tín hiệu âm thanh để thu được các hệ số dự báo mô tả sự tương tự phổ của tín hiệu âm thanh và đưa tín hiệu âm thanh tới hàm bộ lọc phân tích phụ thuộc vào các hệ số dự báo để xuất ra tín hiệu dư của tín hiệu âm thanh. Bộ tìm thừa số được tạo cấu hình để áp dụng việc tìm thừa số ma trận trên ma trận tương quan tự động hoặc ma trận hiệp phương sai của hàm bộ lọc phân tích được định rõ bằng các hệ số dự báo để thu được các ma trận được tìm thừa số. Bộ biến đổi được tạo cấu hình để biến đổi tín hiệu dư dựa trên các ma trận được tìm thừa số để thu được tín hiệu dư được biến đổi. Tầng lượng tử hóa và mã hóa được tạo cấu hình để lượng tử hóa tín hiệu dư được biến đổi để thu được tín hiệu dư được biến đổi được lượng tử hóa hoặc tín hiệu dư được biến đổi được lượng tử hóa được mã hóa.

200



Lĩnh vực kỹ thuật được đề cập

Các phương án của sáng chế đề cập đến bộ mã hóa để mã hóa tín hiệu âm thanh để thu được dòng dữ liệu và bộ giải mã để giải mã dòng dữ liệu để thu được tín hiệu âm thanh. Các phương án khác đề cập đến phương pháp tương ứng để mã hóa tín hiệu âm thanh và để giải mã dòng dữ liệu. Phương án khác đề cập đến chương trình máy tính thực hiện các bước của các phương pháp để mã hóa và/hoặc giải mã.

Tình trạng kỹ thuật của sáng chế

Tín hiệu âm thanh được mã hóa có thể, ví dụ, là tín hiệu tiếng nói; tức là bộ mã hóa tương ứng với bộ mã hóa tiếng nói và bộ giải mã tương ứng với bộ giải mã tiếng nói. Mô hình được sử dụng thường xuyên nhất trong mã hóa tiếng nói là dự báo tuyến tính xuất phát từ mã đại số (algebraic code excited linear prediction - ACELP) mà được sử dụng trong các tiêu chuẩn họ AMR, G.718 và MPEG USAC. Nó dựa trên việc tạo mô hình tiếng nói sử dụng mẫu nguồn, gồm dự báo tuyến tính (linear predictor - LP) để tạo mô hình đường bao phổ, dự báo dài hạn (long time predictor - LTP) để tạo mô hình tần số cơ bản và bảng mã đại số cho phần dư. Các tham số bảng mã được tối ưu hóa trong miền tổng hợp được gán trọng số theo cảm giác. Mô hình theo cảm giác được dựa trên bộ lọc, trong đó sự ánh xạ từ các phần dư tới đầu ra được gán trọng số được mô tả bằng sự tổ hợp của bộ dự báo tuyến tính và bộ lọc được gán trọng số.

Phần lớn nhất của độ phức tạp tính toán trong các bộ mã hóa-giải mã ACELP được dùng cho việc lựa chọn mục nhập bảng mã đại số, mà dựa trên việc lượng tử hóa phần dư. Sự ánh xạ từ miền phần dư tới miền tổng hợp được gán trọng số về cơ bản là phép nhân bởi ma trận có kích thước $N \times N$, trong đó N là chiều dài vectơ. Do sự ánh xạ này, về SNR (signal to noise ratio - tỉ số tín hiệu trên nhiễu) đầu ra được gán trọng số, các mẫu phần dư được tương quan với nhau và không thể được lượng tử hóa một cách độc lập. Theo đó, mỗi vectơ bảng mã thế vị có thể được đánh giá một cách rõ ràng trong miền tổng hợp được gán trọng số để xác định mục nhập tốt nhất. Phương thức này được biết đến như thuật toán phân tích tổng hợp. Hiệu suất tối ưu chỉ khả thi với việc tìm kiếm vét cạn (brute-force) của bảng mã. Kích thước của bảng mã phụ thuộc vào tốc độ bit không phải tốc độ bit đã cho của B , có 2^B mục nhập để đánh giá

độ phức tạp toàn phần của $O(2^B N^2)$, điều này rõ ràng không thực tế khi B lớn hơn hoặc bằng 11. Vì vậy, trên thực tế, các bộ mã hóa-giải mã sử dụng những sự lượng tử hóa không được tối ưu mà cân bằng giữa độ phức tạp và chất lượng. Một vài thuật toán lặp này để tìm kiếm sự lượng tử hóa tốt nhất mà hạn chế độ phức tạp phải trả bằng độ chính xác đã được trình bày. Để khắc phục hạn chế này, cần có phương thức mới.

Tài liệu tham khảo

- [1] B. Bessette, R. Salami, R. Lefebvre, M. Jelinek, J. Rotola-Pukkila, J. Vainio, H. Mikkola, and K. Järvinen, "The adaptive multirate wideband speech codec (AMR-WB)," *Speech and Audio Processing, IEEE Transactions on*, vol. 10, no. 8, pp. 620-636, 2002.
- [2] ITU-T G.718, "Frame error robust narrow-band and wideband embedded variable bit-rate coding of speech and audio from 8-32 kbit/s," 2008.
- [3] M. Neuendorf, P. Gournay, M. Multrus, J. Lecomte, B. Bessette, R. Geiger, S. Bayer, G. Fuchs, J. Hilpert, N. Rettelbach, R. Salami, G. Schuller, R. Lefebvre, and B. Grill, "Unified speech and audio coding scheme for high quality at low bitrates," in *Acoustics, Speech and Signal Processing. ICASSP 2009. IEEE Int Conf*, 2009, pp. 1-4.
- [4] J.-P. Adoul, P. Mabillean, M. Delprat, and S. Morissette, "Fast CELP coding based on algebraic codes," in *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP'87*, vol. 12. IEEE, 1987, pp. 1957-1960.
- [5] C. Laamme, J. Adoul, H. Su, and S. Morissette, "On reducing computational complexity of codebook search in CELP coder through the use of algebraic codes," in *Acoustics, Speech, and Signal Processing, 1990. ICASSP-90*, 1990 International Conference on. IEEE, 1990, pp. 177-180.
- [6] F.-K. Chen and J.-F. Yang, "Maximum-take-precedence ACELP: a low complexity search method," in *Acoustics, Speech, and Signal Processing, 2001. Proceedings.(ICASSP'01). 2001 IEEE International Conference on*, vol. 2. IEEE, 2001, pp. 693-696.
- [7] K. J. Byun, H. B. Jung, M. Hahn, and K. S. Kim, "A fast ACELP codebook search method," in *Signal Processing, 2002 6th International Conference on*, vol. 1.

IEEE, 2002, pp. 422-425. [8] N. K. Ha, "A fast search method of algebraic codebook by reordering search sequence," in Acoustics, Speech, and Signal Processing, 1999. 1999 IEEE International Conference on, vol. 1. IEEE, 1999, pp. 21-24.

[9] M. A. Ramirez and M. Gerken, "Efficient algebraic multipulse search," in Telecommunications Symposium, 1998. ITS'98 Proceedings. SBT/IEEE International. IEEE, 1998, pp. 231-236.

[10] T. Bäckström, "Computationally efficient objective function for algebraic codebook optimization in ACELP," in Interspeech 2013, August 2013.

[11] "Vandermonde factorization of Toeplitz matrices and applications in filtering and warping," IEEE Trans. Signal Process., vol. 61, no. 24, pp. 6257-6263, 2013.

[12] G. H. Golub and C. F. van Loan, Matrix Computations, 3rd ed. John Hopkins University Press, 1996.

[13] T. Bäckström, J. Fischer, and D. Boley, "Implementation and evaluation of the Vandermonde transform," in submitted to EUSIPCO 2014 (22nd European Signal Processing Conference 2014) (EUSIPCO 2014), Lisbon, Portugal, Sep. 2014.

[14] T. Bäckström, G. Fuchs, M. Multrus, and M. Dietz, "Linear prediction based audio coding using improved probability distribution estimation," US Provisional Patent US 61/665 485, 6, 2013.

[15] K. Hermus, P. Wambacq et al., "A review of signal subspace speech enhancement and its application to noise robust speech recognition," EURASIP Journal on Applied Signal Processing, vol. 2007, no. 1, pp. 195-195, 2007.

Bản chất kỹ thuật của sáng chế

Mục đích của sáng chế là đề xuất khái niệm mã hóa và giải mã các tín hiệu âm thanh trong khi tránh được các hạn chế ở trên.

Mục đích này được giải quyết bằng các điểm yêu cầu bảo hộ độc lập.

Phương án thứ nhất đề xuất bộ mã hóa để mã hóa tín hiệu âm thanh thành dòng dữ liệu. Bộ mã hóa bao gồm bộ dự báo (tuyến tính hoặc dài hạn), bộ tìm thừa số, bộ biến đổi và tầng mã hóa được lượng tử hóa. Bộ dự báo được tạo cấu hình để phân tích tín hiệu âm thanh để thu được các hệ số dự báo (tuyến tính hoặc dài hạn) mô tả đường bao phổ của tín hiệu âm thanh hoặc tần số cơ bản của tín hiệu âm thanh và đưa tín hiệu

âm thanh tới hàm bộ lọc phân tích phụ thuộc vào các hệ số dự báo để xuất ra tín hiệu dư của tín hiệu âm thanh. Bộ tìm thừa số được tạo cấu hình để áp dụng việc tìm thừa số ma trận trên ma trận tương quan tự động hoặc ma trận hiệp phương sai của hàm bộ lọc tổng hợp được định rõ bởi các hệ số dự báo để thu được các ma trận được tìm thừa số. Bộ biến đổi được tạo cấu hình để biến đổi tín hiệu dư dựa trên các ma trận được tìm thừa số để thu được tín hiệu dư được biến đổi. Tầng lượng tử hóa và mã hóa được tạo cấu hình để lượng tử hóa tín hiệu dư biến đổi để thu được tín hiệu dư được biến đổi được lượng tử hóa hoặc tín hiệu dư được biến đổi được lượng tử hóa được mã hóa.

Phương án khác đề xuất bộ giải mã để giải mã dòng dữ liệu thành tín hiệu âm thanh. Bộ giải mã bao gồm tầng giải mã, bộ biến đổi lại và tầng tổng hợp. Tầng giải mã được tạo cấu hình để xuất ra tín hiệu dư biến đổi dựa trên tín hiệu dư biến đổi được lượng tử hóa trở lại hoặc tín hiệu dư biến đổi được lượng tử hóa được mã hóa trở lại. Bộ biến đổi lại được tạo cấu hình để biến đổi lại tín hiệu dư từ tín hiệu dư được biến đổi dựa trên các ma trận được tìm thừa số do việc tìm thừa số ma trận của ma trận tương quan tự động hoặc ma trận hiệp phương sai của hàm bộ lọc tổng hợp được định rõ bởi các hệ số dự báo mô tả đường bao phổ của tín hiệu âm thanh hoặc tần số cơ bản của tín hiệu âm thanh để thu được các ma trận được tìm thừa số. Tầng tổng hợp được tạo cấu hình để tổng hợp tín hiệu âm thanh được dựa trên tín hiệu dư bằng việc sử dụng hàm bộ lọc tổng hợp được định rõ bằng hệ số dự báo.

Như có thể thấy trên cơ sở hai phương án này, việc mã hoá và giải mã là quá trình hai tầng, điều này khiến khái niệm này có thể so sánh được với ACELP. Bước thứ nhất cho phép sự lượng tử hóa của phép tổng hợp đối với đường bao phổ hoặc tần số cơ bản, trong đó tầng thứ hai cho phép sự lượng tử hóa hoặc tổng hợp (trực tiếp) của tín hiệu dư, còn được gọi là tín hiệu kích thích và biểu diễn tín hiệu này sau khi lọc tín hiệu với đường bao phổ hoặc tần số cơ bản của tín hiệu âm thanh. Ngoài ra, tương tự với ACELP, sự lượng tử hóa tín hiệu dư hoặc tín hiệu kích thích phù hợp với vấn đề tối ưu hóa, trong đó hàm mục tiêu của vấn đề tối ưu hóa theo các dẫn dắt được bộc lộ trong tài liệu này trong tài liệu này khác biệt đáng kể khi được so sánh với ACELP. Cụ thể, các dẫn dắt của sáng chế được dựa trên nguyên tắc tìm thừa số ma trận được sử dụng để giải tương quan hàm mục tiêu của vấn đề tối ưu hóa, nhờ đó có thể tránh được phép lặp tiêu tốn tính toán và hiệu suất tối ưu được đảm bảo. Việc tìm thừa số ma trận,

mà là một bước trung tâm của các phương án kèm theo, được gồm trong phương án bộ mã hóa và tốt hơn là, nhưng không nhất thiết, được bao gồm trong phương án bộ giải mã.

Việc tìm thừa số ma trận có thể dựa trên các kỹ thuật khác nhau, ví dụ như phân tích trị riêng, việc tìm thừa số Vandermonde hoặc việc tìm thừa số khác, trong đó đối với mỗi kỹ thuật được chọn việc tìm thừa số tìm thừa số là ma trận, ví dụ ma trận tương quan tự động hoặc ma trận hiệp phương sai của hàm bộ lọc tổng hợp, được định rõ bởi các hệ số dự báo (tuyến tính hoặc dài hạn) mà được phát hiện bởi âm thanh thứ nhất trong tầng thứ nhất (bộ dự báo tuyến tính hoặc bộ dự báo dài hạn) của việc mã hóa hoặc giải mã.

Theo các phương án khác, bộ tìm thừa số tìm thừa số hàm bộ lọc tổng hợp, bao gồm các hệ số dự báo mà được lưu trữ sử dụng ma trận, hoặc tìm thừa số phiên bản được gán trọng số của ma trận hàm bộ lọc tổng hợp. Ví dụ, việc tìm thừa số có thể được thực hiện bằng cách sử dụng ma trận Vandermonde V , ma trận đường chéo D và phiên bản biến đổi được liên kết của ma trận Vandermonde V^* . Ma trận Vandermonde có thể được tìm thừa số sử dụng công thức $R = V^*DV$; $C = V^*DV$, trong đó ma trận tương quan tự động R hoặc ma trận hiệp phương sai C được định rõ bởi phiên bản được biến đổi được liên kết của ma trận hàm bộ lọc tổng hợp H^* và phiên bản thông thường của ma trận hàm tổng hợp H , tức là $R = H^*H$ hoặc $C = H^*H$.

Theo phương án khác, bộ biến đổi, bắt đầu từ ma trận đường chéo được xác định trước đó D và ma trận Vandermonde được xác định trước đó V , biến đổi tín hiệu dư x thành tín hiệu dư được biến đổi y sử dụng công thức $y = D^{1/2}Vx$ hoặc công thức $y = DVx$.

Theo phương án khác, tầng tìm thừa số và tầng mã hóa hiện tại có thể lượng tử hóa tín hiệu dư được biến đổi y để thu được tín hiệu dư được biến đổi được lượng tử hóa $\hat{y}$. Sự biến đổi này là vấn đề tối ưu hóa, như được thảo luận ở trên, trong đó hàm mục tiêu $\eta(y) = \frac{(y - \hat{y})^2}{\|\hat{y}\|^2}$ được sử dụng. Trong tài liệu này, điều thuận lợi là hàm mục tiêu giảm độ phức tạp khi được so sánh với các hàm mục tiêu được sử dụng cho các phương pháp mã hóa hoặc giải mã khác nhau, ví dụ hàm mục tiêu được sử dụng trong bộ mã hóa ACELP.

Theo phương án, bộ giải mã nhận các ma trận được tìm thừa số từ bộ mã hóa, ví dụ cùng với dòng dữ liệu, hoặc theo phương án khác bộ giải mã bao gồm bộ tìm thừa số tùy chọn mà thực hiện việc tìm thừa số ma trận. Theo phương án được ưu tiên, bộ giải mã nhận các ma trận được tìm thừa số trực tiếp và làm lệch các hệ số dự báo từ các ma trận được tìm thừa số này do các ma trận này có nguồn gốc từ các hệ số dự báo (so với bộ giải mã). Các phương án cho phép giảm thêm độ phức tạp của bộ giải mã.

Các phương án khác đề xuất các phương pháp tương ứng để mã hóa tín hiệu âm thanh thành dòng dữ liệu và để giải mã dòng dữ liệu thành tín hiệu âm thanh. Theo phương án bổ sung, phương pháp để mã hóa cũng như phương pháp để giải mã có thể được thực hiện hoặc ít nhất được thực hiện một phần bằng bộ xử lý ví dụ CPU của máy tính.

Mô tả vắn tắt các hình vẽ

Các phương án của sáng chế sẽ được thảo luận trong các hình vẽ đi kèm, trong đó:

Fig.1 thể hiện sơ đồ khối dạng giản lược của bộ mã hóa để mã hóa tín hiệu âm thanh theo phương án thứ nhất;

Fig.1b thể hiện lưu đồ dạng giản lược của phương pháp tương ứng để mã hóa tín hiệu âm thanh theo phương án thứ nhất;

Fig.2a thể hiện sơ đồ khối dạng giản lược của bộ giải mã để giải mã dòng dữ liệu theo phương án thứ hai;

Fig.2b thể hiện lưu đồ dạng giản lược của phương pháp tương ứng để giải mã dòng dữ liệu theo phương án thứ hai;

Fig.3a thể hiện sơ đồ dạng giản lược minh họa tỉ số tín hiệu trên nhiễu cảm nhận trung bình như hàm của các bit trên mỗi khung cho các phương pháp lượng tử hóa khác nhau; và

Fig.3b thể hiện sơ đồ dạng giản lược minh họa thời gian chạy được chuẩn hóa của các phương pháp lượng tử hóa khác nhau như hàm của các bit trên mỗi khung; và

Fig.3c thể hiện sơ đồ dạng giản lược minh họa các đặc tính của biến đổi Vandermonde.

Mô tả chi tiết sáng chế

Các phương án của sáng chế sẽ được thảo luận chi tiết sau đây bằng cách tham

chiều tới các hình vẽ kèm theo. Do đó, các số tham chiếu giống nhau sẽ được gán cho các đối tượng có chức năng giống nhau hoặc tương tự sao cho phần mô tả có thể được áp dụng hoặc hoán đổi cho nhau.

Fig.1a thể hiện bộ mã hóa 10 ở cấu hình cơ bản. Bộ mã hóa 10 bao gồm bộ dự báo 12, trong tài liệu này được thực hiện như bộ dự báo tuyến tính 12, cùng với bộ tìm thừa số 14, bộ biến đổi 16 và tăng lượng tử hóa và mã hóa 18.

Bộ dự báo tuyến tính 12 được bố trí tại đầu vào để nhận tín hiệu âm thanh AS, tốt hơn là tín hiệu âm thanh kỹ thuật số chẳng hạn như tín hiệu điều biến mã xung (pulse code modulated signal - PCM). Bộ dự báo tuyến tính 12 được ghép với bộ tìm thừa số 14 và tới đầu ra của bộ mã hóa, xem số tham chiếu DS_{LPC}/DS_{DV} qua cái gọi là LPC kênh LPC. Ngoài ra, bộ dự báo tuyến tính 12 được ghép với bộ biến đổi 16 qua cái gọi là kênh dư. Ngược lại, bộ biến đổi 16 được (bổ sung vào kênh dư) được ghép với bộ tìm thừa số 14 tại phía đầu vào của nó. Tại phía đầu ra của nó, bộ biến đổi được nối với tăng lượng tử hóa và mã hóa 18, trong đó tăng lượng tử hóa và mã hóa 18 được nối với đầu ra (xem số tham chiếu $DS_{\hat{y}}$). Hai dòng dữ liệu DS_{LPC}/DS_{DV} và $DS_{\hat{y}}$ tạo ra dòng dữ liệu DS để xuất ra ngoài.

Chức năng của bộ mã hóa 10 sẽ được thảo luận dưới đây, trong đó viện dẫn thêm đến Fig.1b mô tả phương pháp 100 để mã hóa. Như có thể thấy theo Fig.1b, phương pháp cơ bản 100 để mã hóa tín hiệu âm thanh AS vào dòng dữ liệu DS bao gồm bốn bước cơ bản 120, 140, 160 và 180 mà được thực hiện bởi các bộ phận 12, 14, 16 và 18. Trong bước thứ nhất 120, bộ dự báo tuyến tính 12 phân tích tín hiệu âm thanh AS để thu được các hệ số dự báo tuyến tính LPC. Các hệ số dự báo tuyến tính LPC mô tả đường bao phổ của tín hiệu âm thanh AS mà cho phép tổng hợp một cách cơ bản tín hiệu âm thanh sử dụng cái được gọi là hàm bộ lọc tổng hợp H sau đó. Hàm bộ lọc tổng hợp H có thể bao gồm các giá trị được gán trọng số của hàm bộ lọc tổng hợp được định rõ bởi các hệ số LPC. Các hệ số tuyến tính LPC được xuất ra tới bộ tìm thừa số 14 sử dụng LPC kênh LPC và được chuyển tiếp tới đầu ra của bộ mã hóa 10. Bộ dự báo tuyến tính 12 còn đưa tín hiệu âm thanh AS tới hàm bộ lọc phân tích H mà được định rõ bởi các hệ số dự báo tuyến tính LPC. Sự xử lý này giống với phép tổng hợp của tín hiệu âm thanh dựa trên các hệ số LPC được thực hiện bởi bộ giải mã. Kết quả của bước phụ này là tín hiệu dư x được xuất tới bộ biến đổi 16 mà không có phần

tín hiệu có thể được mô tả bằng hàm bộ lọc H . Lưu ý rằng bước này được thực hiện theo khung, tức là tín hiệu âm thanh AS có biên độ và miền thời gian được chia hoặc được lấy mẫu vào các cửa sổ thời gian (các mẫu), ví dụ có độ dài 5 ms, và được lượng tử hóa trong miền tần số.

Bước tiếp sau là biến đổi tín hiệu dư x (xem bước phương pháp 160) được thực hiện bằng bộ biến đổi 16. Bộ biến đổi 16 được tạo cấu hình để biến đổi tín hiệu dư x để thu tín hiệu dư được biến đổi y được xuất ra tới tầng lượng tử hóa và mã hóa 18. Ví dụ, sự biến đổi 160 có thể được dựa trên công thức $y = D^{1/2}Vx$ hoặc công thức $y = DVx$, trong đó các ma trận D và V được cung cấp bởi bộ tìm thừa số 14. Do đó, sự biến đổi của tín hiệu dư x được dựa trên ít nhất hai ma trận được tìm thừa số V ví dụ được gọi là ma trận Vandermonde và D được gọi là ma trận đường chéo.

Việc tìm thừa số ma trận được áp dụng có thể được lựa chọn một cách tự do, ví dụ ma trận phân tích trị riêng, việc tìm thừa số Vandermonde, phân tích Cholesky hoặc tương tự. Việc tìm thừa số Vandermonde có thể được sử dụng như việc tìm thừa số của các ma trận Toeplitz xác định dương, đối xứng, như ma các ma trận tương quan tự động, vào tích số của các ma trận Vandermonde V và V^* . Đối với ma trận tương quan tự động trong hàm mục tiêu, nó tương ứng với biến đổi Fourier rời rạc được uốn cong, mà thường được gọi là biến đổi Vandermonde. Bước 140 này của sự tìm thừa số ma trận được thực hiện bởi bộ tìm thừa số 14 và biểu diễn phần cơ bản của sáng chế, sẽ được thảo luận chi tiết sau khi thảo luận về chức năng của tầng lượng tử hóa và mã hóa 18.

Tầng lượng tử hóa và mã hóa 18 lượng tử hóa tín hiệu dư được biến đổi y , được nhận từ bộ biến đổi 16, để thu được tín hiệu dư được biến đổi được lượng tử hóa $\hat{y}$. Tín hiệu dư được lượng tử hóa được biến đổi $\hat{y}$ này được xuất ra như một phần của dòng dữ liệu $DS_{\hat{y}}$. Lưu ý, toàn bộ dòng dữ liệu DS bao gồm phần LPC, được đề cập bởi DS_{LPC}/DS_{DV} , và phần $\hat{y}$ được đề cập bởi $DS_{\hat{y}}$.

Sự lượng tử hóa của tín hiệu dư biến đổi y có thể, ví dụ, được thực hiện bằng cách sử dụng hàm mục tiêu, ví dụ, về $\eta(y) = \frac{(y^* - \hat{y})^2}{\|\hat{y}\|^2}$. Hàm mục tiêu này có, khi được so sánh với hàm mục tiêu điển hình của bộ mã hóa ACELP, độ phức tạp được giảm sao cho việc mã hóa được cải thiện một cách thuận lợi đối với hiệu suất của nó. Sự cải

thiện hiệu suất có thể được sử dụng cho các tín hiệu âm thanh mã hóa AS có độ phân giải cao hơn hoặc để giảm các nguồn được yêu cầu.

Cần lưu ý rằng tín hiệu DS_y có thể là tín hiệu được mã hóa, trong đó sự mã hóa được thực hiện bằng tăng lượng tử hóa và mã hóa 18. Do đó, theo các phương án khác, tăng lượng tử hóa và mã hóa 18 có thể bao gồm bộ mã hóa mà có thể được tạo cấu hình thành mã hóa số học. Bộ mã hóa của tăng lượng tử hóa và mã hóa 18 có thể sử dụng các bước lượng tử hóa tuyến tính (tức là khoảng cách bằng nhau) hoặc biến số, ví dụ các bước logarit, lượng tử hóa. Ngoài ra, bộ mã hóa có thể được tạo cấu hình để thực hiện mã hóa entropy (không hao hụt) trong đó độ dài mã biến đổi như hàm xác suất của các tín hiệu đầu vào suy biến AS. Do đó, để thu được độ dài mã tối ưu có thể là lựa chọn thay thế để phát hiện xác suất của các tín hiệu đầu ra được dựa trên đường bao tổng hợp và do đó được dựa trên các hệ số LPC. Do đó, tăng mã hóa được lượng tử hóa có thể cũng có đầu vào cho kênh LPC.

Dưới đây, khái quát về sự giảm độ phức tạp của hàm mục tiêu $\eta(y)$ sẽ được thảo luận. Như được đề cập ở trên, việc mã hóa được cải thiện dựa trên bước tìm thừa số ma trận 140 được thực hiện bởi bộ tìm thừa số 14. Bộ tìm thừa số 14 tìm thừa số, ví dụ, ma trận tương quan tự động R hoặc ma trận hiệp phương sai C của hàm tổng hợp bộ lọc H được định rõ bởi các hệ số dự báo tuyến tính LPC (xem kênh LPC). Kết quả của việc tìm thừa số này là hai ma trận được tìm thừa số, ví dụ, ma trận Vandermonde V và ma trận đường chéo D biểu diễn ma trận gốc H bao gồm các hệ số LPC suy biến. Do đó, các mẫu của tín hiệu dư x được giải tương quan. Theo đó, sự lượng tử hóa trực tiếp (so với bước 180) của tín hiệu dư biến đổi là sự lượng tử hóa tối ưu, mà ở đó độ phức tạp tính toán hầu như không phụ thuộc vào tốc độ bit. Khi so sánh, phương thức thông thường để tối ưu hóa bảng mã ACELP phải cân bằng giữa độ phức tạp tính toán và độ chính xác, đặc biệt ở tốc độ bit cao. Do đó, khái quát được thảo luận thực sự bắt đầu từ các quy trình ACELP thông thường.

Hàm mục tiêu thông thường của ACELP có dạng ma trận hiệp phương sai. Theo các phương thức được cải thiện, có hàm mục tiêu thay thế mà sử dụng ma trận tương quan tự động của hàm tổng hợp được gán trọng số. Các bộ mã hóa-giải mã dựa trên tỉ số tín hiệu trên nhiễu (signal to noise ratio - SNR) được tối ưu hóa ACELP nằm trong miền tổng hợp được gán trọng số theo cảm giác. Hàm mục tiêu có thể được biểu

thị như:

$$\eta(\mathbf{x}, \gamma) = \|\mathbf{H}(\mathbf{x} - \gamma\hat{\mathbf{x}})\|^2 \quad (1)$$

trong đó $\mathbf{x}$ là phần dư đích, phần dư được lượng tử hóa $\hat{\mathbf{x}}$, ma trận nhân chập $\mathbf{H}$ tương ứng với bộ lọc tổng hợp được gán trọng số và hệ số độ khuếch đại định tỉ lệ γ . Để tìm sự lượng tử hóa tối ưu $\hat{\mathbf{x}}$, phương thức tiêu chuẩn là tìm giá trị tối ưu của γ , được biểu thị bởi γ^* , tại 0 của đạo hàm $\eta(\mathbf{x}, \gamma)$. Bằng việc thêm γ^* vào phương trình (1) hàm mục tiêu mới thu được:

$$\eta(x) = \frac{(\mathbf{x}^* \mathbf{H}^* \mathbf{H} \hat{\mathbf{x}})^2}{\hat{\mathbf{x}}^* \mathbf{H}^* \mathbf{H} \hat{\mathbf{x}}} \quad (2)$$

trong đó $\mathbf{H}^*$ là phiên bản được biến đổi được liên hợp của phép tổng hợp với hàm $\mathbf{H}$.

Lưu ý rằng phương thức thông thường $\mathbf{H}$ là ma trận nhân chập tam giác dưới bình phương, mà tại đó ma trận hiệp phương sai $\mathbf{C} = \mathbf{H}^* \mathbf{H}$ là ma trận hiệp phương sai đối xứng. Sự thay thế của ma trận tam giác dưới bằng ma trận nhân chập kích thước đầy đủ, mà ở đó ma trận tương quan tự động $\mathbf{R} = \mathbf{H}^* \mathbf{H}$ là ma trận Toeplitz đối xứng, tương ứng với tương quan khác của bộ lọc tổng hợp được gán trọng số. Sự thay thế này mang đến sự giảm đáng kể và độ phức tạp, với ảnh hưởng tối thiểu về chất lượng.

Bộ dự báo tuyến tính 14 có thể sử dụng cả hai, cụ thể là ma trận hiệp phương sai $\mathbf{C}$ hoặc ma trận tương quan tự động $\mathbf{R}$ cho việc tìm thừa số ma trận. Các thảo luận dưới đây được thực hiện trên giả định rằng ma trận tương quan tự động $\mathbf{R}$ được sử dụng để cải biên hàm mục tiêu bằng việc tìm thừa số ma trận phụ thuộc vào các hệ số LPC. Các ma trận Toeplitz được xác định dương đối xứng như $\mathbf{R}$ có thể được phân tích như:

$$\mathbf{R} = \mathbf{V}^* \mathbf{D} \mathbf{V} \quad (3)$$

qua một số phương pháp, gồm phân tích trị riêng. Tại đây, $\mathbf{V}^*$ là phiên bản được biến đổi được liên hợp của ma trận Vandermonde $\mathbf{V}$. Trong phương thức thông thường sử dụng ma trận hiệp phương sai $\mathbf{C}$ việc tìm thừa số khác có thể được áp dụng như phân tích giá trị đơn $\mathbf{C} = \mathbf{U} \mathbf{S} \mathbf{V}$.

Đối với ma trận tương quan tự động, việc tìm thừa số khác trong tài liệu này là việc tìm thừa số Vandermonde, mà cũng là một dạng phương trình (3) có thể được sử dụng. Việc tìm thừa số Vandermonde là một khái niệm mới cho phép tìm thừa số/biến

đổi. Ma trận Vandermonde có V với giá trị bằng $|v_k| = 1$ và:

$$V = \begin{pmatrix} 1 & v_0 & v_0^2 & \dots & v_0^{N-1} \\ 1 & v_1 & v_1^2 & \dots & v_1^{N-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & v_{N-1} & v_{N-1}^2 & \dots & v_{N-1}^{N-1} \end{pmatrix} \quad (4)$$

và D là ma trận đường chéo với các mục nhập dương hoàn toàn. Phép phân tích có thể được tính toán với độ chính xác tùy ý với độ phức tạp $O(N^3)$. Phép phân tích trực tiếp có độ phức tạp thông thường $O(N^3)$, nhưng trong tài liệu này có thể được giảm xuống $O(N^2)$ hoặc nếu việc tìm thừa số xấp xỉ là đủ, thì độ phức tạp có thể được giảm xuống $O(N \log N)$. Đối với phép phân tích được chọn, nó có thể được định rõ:

$$y = D^{1/2} V x \text{ và } \hat{y} = D^{1/2} V \hat{x} \quad (5)$$

trong đó $x = V^{-1} D^{-1/2} y$ và chèn vào phương trình (2) nó có thể thu được:

$$\eta(y) = \frac{(y^* \hat{y})^2}{\|\hat{y}\|^2} \quad (6)$$

Lưu ý trong tài liệu này, các mẫu y không tương quan với nhau, và hàm mục tiêu ở trên chỉ là tương quan được chuẩn hóa giữa phần dư đích và phần dư được lượng tử hóa. Theo đó các mẫu y có thể được lượng tử hóa một cách độc lập và nếu độ chính xác của tất cả các mẫu là bằng nhau thì sự lượng tử này sinh ra độ chính xác khả thi nhất.

Trong trường hợp việc tìm thừa số Vandermonde, do V có giá trị $|v_k| = 1$ nó tương ứng với biến đổi Fourier rời rạc được uốn cong và các phần tử của y tương ứng với các thành phần của phần dư. Ngoài ra, phép nhân của ma trận đường chéo D tương ứng với sự định tỷ lệ của các băng tần và theo đó y là phép biểu diễn miền tần số của phần dư.

Trái lại, phân tích trị riêng có sự diễn dịch vật lý chỉ khi độ dài của số gần tới vô cực, khi phân tích trị riêng và biến đổi Fourier trùng nhau. Do đó các phép phân tích trị riêng độ dài vô cực liên quan không chặt chẽ tới phép biểu diễn tần số của tín hiệu, nhưng việc tạo nhãn các thành phần tới các tần số là khó. Tuy nhiên, phép phân tích trị riêng được biết tới là cơ sở tối ưu hóa, mà ở đó nó có thể cho hiệu suất tốt nhất trong một số trường hợp.

Bắt đầu từ hai ma trận được tìm thừa số V và D bộ biến đổi 16 thực hiện sự

biến đổi 160 sao cho tín hiệu dư x được biến đổi sử dụng vectơ được định rõ bằng phương trình (5).

Giả sử x là nhiễu âm trắng không tương quan, thì các mẫu Vx cũng sẽ có trung bình năng lượng bằng nhau. Kết quả là, bộ mã hóa số học hoặc bộ mã hóa sử dụng bảng mã đại số để mã hóa các giá trị có thể được sử dụng. Tuy nhiên, sự lượng tử hóa của Vx không được tối ưu hóa đối với hàm mục tiêu do nó bỏ qua ma trận đường chéo $D^{1/2}$. Mặt khác, sự biến đổi hoàn toàn $y = D^{1/2}Vx$ gồm việc định tỷ lệ bởi ma trận đường chéo D , mà thay đổi trung bình năng lượng của các mẫu y . Để tạo bảng mã số học với các phương sai không đồng đều là không đáng kể. Do đó, có thể là tùy chọn để sử dụng bảng mã số học thay vì sự tiêu thụ bit tối ưu. Việc lập mã số học có thể được định rõ một cách chính xác như được bộc lộ trong [14].

Lưu ý rằng, nếu phép phân tích được sử dụng, như biến đổi Vandermonde hoặc biến đổi phức tạp khác, các phần thực và các phần ảo không phụ thuộc vào các biến số ngẫu nhiên. Nếu các biến thể của biến số phức tạp là σ^2 , thì các phần thực và ảo có phương sai $\sigma^2/2$. Các phép phân tích được gán giá trị thực như phép phân tích trị riêng chỉ cung cấp các giá trị thực, mà tại đó sự phân tách các phần thực và ảo là không cần thiết. Đối với hiệu suất cao hơn với sự biến đổi được gán giá trị phức tạp, các phương pháp thông thường để lập mã số học cho các giá trị phức tạp có thể được áp dụng.

Theo phương án ở trên các hệ số dự báo LPC (so với DS_{LPC}) được xuất ra như các tín hiệu LSF (các tín hiệu tần số vạch phổ), trong đó tùy chọn khác xuất ra các hệ số dự báo LPC trong các ma trận được tìm thừa số V và D (xem. DS_{DV}). Tùy chọn khác này được ngụ ý bằng các nét đứt được đánh dấu bởi V , D và chỉ ra rằng DS_{DV} là kết quả từ đầu ra của bộ tìm thừa số 14.

Do đó phương án khác của sáng chế đề cập đến dòng dữ liệu (DS) bao gồm các hệ số dự báo LPC dưới dạng hai ma trận được tìm thừa số (DS_{VD}).

Đối với Fig.2 bộ giải mã 20 và phương pháp tương ứng 200 để giải mã sẽ được thảo luận.

Fig.2a thể hiện bộ giải mã 20 bao gồm tầng giải mã 22, bộ tìm thừa số tùy chọn 24, bộ biến đổi lại 26 và tầng tổng hợp 28. Tầng giải mã 22 cùng với bộ tìm thừa số 24 được bố trí tại đầu ra của bộ giải mã 20 và do đó được tạo cấu hình để nhận dòng dữ liệu DS. Cụ thể là, phần thứ nhất của dòng dữ liệu DS, cụ thể là các hệ số dự báo tuyến

tính được cung cấp tới bộ tìm thừa số tùy chọn 24 (xem DS_{LPC}/DS_{DV}), trong đó phần thứ hai, cụ thể là tín hiệu dư biến đổi được lượng tử hóa y hoặc tín hiệu dư biến đổi được lượng tử hóa được mã hóa y được cung cấp tới tầng mã hóa 22 (xem DS_y). Tầng tổng hợp 28 được bố trí tại đầu ra của bộ giải mã 20 và được tạo cấu hình để xuất ra tín hiệu âm thanh AS' tương tự, nhưng không bằng với tín hiệu âm thanh AS .

Phép tổng hợp tín hiệu âm thanh AS' dựa trên các hệ số LPC (xem DS_{LPC}/DS_{DV}) và dựa trên tín hiệu dư x . Do đó, tầng tổng hợp 28 được nối với đầu ra để nhận tín hiệu DS_{LPC} và bộ biến đổi lại 26 cung cấp tín hiệu dư x . Bộ biến đổi lại 26 tính toán tín hiệu dư x dựa trên tín hiệu dư được biến đổi y và dựa trên ít nhất hai ma trận được tìm thừa số V và D . Do đó, bộ biến đổi lại 26 có ít nhất hai đầu vào, cụ thể là đầu vào thứ nhất để nhận V và D , ví dụ, từ bộ tìm thừa số 24, và một đầu ra để nhận tín hiệu dư được biến đổi y từ tầng giải mã.

Chức năng của bộ giải mã 20 sẽ được thảo luận cụ thể bên dưới viện dẫn tới phương pháp tương ứng 200 được minh họa bởi Fig.2b. Bộ giải mã 20 nhận dòng dữ liệu DS (từ bộ mã hóa). Tín hiệu dữ liệu DS cho phép bộ giải mã 20 tổng hợp tín hiệu âm thanh AS' , trong đó phần của dòng dữ liệu được tham chiếu bởi DS_{LPC}/DS_{DV} kích hoạt phép tổng hợp của tín hiệu cơ bản, trong đó phần được tham chiếu bởi DS_y kích hoạt phép tổng hợp của phần chi tiết của tín hiệu âm thanh AS' . Trong bước thứ nhất 220 tầng giải mã 22 giải mã tín hiệu quay trở lại DS_y và xuất ra tín hiệu dư được biến đổi y tới bộ biến đổi lại 26 (xem bước 260).

Song song hoặc nối tiếp, bộ tìm thừa số 24 thực hiện việc tìm thừa số (xem bước 240). Như được thảo luận cho bước 140, bộ tìm thừa số 24 áp dụng việc tìm thừa số ma trận tương quan tự động R hoặc ma trận hiệp phương sai C của hàm bộ lọc tổng hợp H , tức là, phép tổng hợp này được sử dụng bởi bộ giải mã 20 tương tự hoặc gần tương tự với việc tìm thừa số được mô tả trong sự mã hóa (xem phương pháp 100) và do đó, có thể phân tích trị riêng hoặc tìm thừa số Cholesky như đã thảo luận ở trên. Tại đây, hàm bộ lọc tổng hợp H được làm lệch khỏi dòng dữ liệu quay lại DS_{LPC}/DS_{DV} . Ngoài ra, bộ tìm thừa số 24 xuất ra hai ma trận được tìm thừa số V và D tới bộ biến đổi lại 26.

Dựa trên hai ma trận V và D bộ biến đổi lại 26 biến đổi lại tín hiệu dư x từ tín hiệu dư được biến đổi và xuất ra x tới tầng tổng hợp 28 (xem bước 280). Tầng tổng

hợp 28 tổng hợp tín hiệu âm thanh AS' được dựa trên tín hiệu dư x và dựa trên LPC các hệ số LPC được nhận như dòng dữ liệu DS_{LPC}/DS_{DV} . Cần lưu ý rằng tín hiệu âm thanh AS' là tương tự nhưng không bằng tín hiệu âm thanh AS do sự lượng tử hóa được thực hiện bằng bộ mã hóa 10 là không hao hụt.

Theo phương án khác, các ma trận được tìm thừa số V và D có thể được cung cấp tới bộ biến đổi lại 26 từ đối tượng khác, ví dụ, trực tiếp từ bộ mã hóa 10 (như một phần của dòng dữ liệu). Do đó, bộ tìm thừa số 24 của bộ giải mã 20 cùng với bước 240 của việc tìm thừa số ma trận là các đối tượng/bước tùy chọn và do đó được minh họa bằng các nét đứt. Trong tài liệu này, các hệ số dự báo LPC có thể là tùy chọn thay thế (trên cơ sở mà phép tổng hợp 280 được thực hiện) có thể được suy ra từ các ma trận được tìm thừa số quay lại V và D. Nói cách khác, điều này có nghĩa dòng dữ liệu DS bao gồm DS_y và các ma trận V và D (tức là DS_{DV}) thay cho DS_y và DS_{LPC} .

Các cải thiện hiệu suất của việc mã hóa được mô tả ở trên (cùng với việc giải mã) được thảo luận ở dưới tham chiếu tới các Fig.3a và Fig.3b.

Fig.3a thể hiện sơ đồ minh họa tỉ số tín hiệu trên nhiễu cảm nhận trung bình như hàm bit được sử dụng để có thể nhận được độ dài và 64 khung bằng nhau. Trên sơ đồ 5 các đường cong thể hiện năm phương thức lượng tử hóa được minh họa, trong đó hai phương thức, cụ thể là lượng tử hóa tối ưu và lượng tử hóa lặp theo cặp là các phương thức thông thường. Công thức (1) tạo ra cơ sở của sự so sánh này. Như sự so sánh hiệu suất lượng tử hóa của phương pháp giải tương quan được đề xuất với miền thời gian thông thường biểu diễn tín hiệu dư, bộ ACELP có thể được thực hiện như sau. Tín hiệu đầu được lấy mẫu lại tới 12,8kHz và bộ dự báo tuyến tính được ước lượng với cửa sổ Hamming có độ dài 32 ms, tại trung tâm mỗi khung. Phần dư dự báo sau đó được tính toán cho các khung có độ dài 5 ms, tương ứng với khung phụ của bộ mã hóa giải mã AMR-WB. Bộ dự báo dài hạn được tối ưu tại các độ trễ số nguyên giữa 32 và 150 mẫu, với sự tìm kiếm toàn diện. Giá trị tối ưu được sử dụng cho độ khuếch đại LTP mà không có sự lượng tử hóa.

Sự chỉnh tăng với bộ lọc $(1 - 0,68z^{-1})$ được áp dụng cho tín hiệu đầu vào và trong phép tổng hợp như trong AMR-WB. Sự gán trọng số theo cảm giác được áp dụng là $A(0,92z^{-1})$, trong đó $A(z)$ là bộ lọc dự báo tuyến tính.

Để đánh giá hiệu suất cần so sánh sự lượng tử hóa được đề xuất với các phương

thức thông thường (lượng tử hóa tối ưu và lượng tử hóa lập theo cặp). Các phương thức được sử dụng thường xuyên nhất chia tín hiệu dư của khung có độ dài 64 khung thành 4 rãnh được đan xen. Các phương thức được áp dụng với hai phương pháp, cụ thể là phương thức lượng tử hóa tối ưu (so bởi Opt) trong đó tất cả các tổ hợp được thử trong tìm kiếm toàn diện hoặc lượng tử hóa lập theo cặp (so với Pair) trong đó hai xung được bổ sung liên tiếp bằng việc thử chúng trên tất cả vị trí có thể.

Phương thức cũ trở nên phức tạp không thể tính toán được cho các tốc độ bit trên 15 bit mỗi khung, khi phương thức mới là điểm cực thuận. Lưu ý rằng phương thức mới phức tạp hơn trình độ kỹ thuật của các phương pháp được áp dụng trong bộ mã hóa-giải mã như AMR-WB nhưng, do đó, nó cũng có khả năng sinh ra tỉ số tín hiệu trên nhiễu tốt hơn. Các phương pháp thông thường được so sánh với các thuật toán được thảo luận ở trên cho sự lượng tử hóa.

Lượng tử hóa Vandermonde (so với Vand) biến đổi vectơ dư x bằng $y = D^{1/2}Vx$ trong đó các ma trận V và D được thu từ việc tìm thừa số Vandermonde và sự lượng tử hóa đang sử dụng bộ lập mã số học. Lượng tử hóa trị riêng (so với Eig) tương tự với lượng tử hóa Vandermonde nhưng trong tài liệu này các ma trận V và D được thu bằng các phân tích trị riêng. Ngoài ra, lượng tử hóa FFT (xem FFT) có thể được áp dụng, tức là, theo phương án khác, tổ hợp việc tạo cửa sổ sử dụng các bộ lọc tại biến đổi $y = D^{1/2} Vx$ có thể được sử dụng thay thế cho biến đổi Fourier rời rạc (discrete Fourier transformation - DFT), biến đổi cosin rời rạc (discrete cosine transformation - DCT), biến đổi cosin rời rạc được biến đổi (modified discrete cosine transformation - MDCT) hoặc các biến đổi trong các thuật toán xử lý tín hiệu. FFT (biến đổi Fourier nhanh) của tín hiệu dư được lấy ra nơi bộ lập mã số học như đối với lượng tử hóa Vandermonde được áp dụng. Phương pháp FFT rõ ràng sẽ cho một chất lượng kém vì nó cũng được biết đến rộng rãi rằng điều quan trọng là xem xét sự tương quan giữa các mẫu trong phương trình (2). Do đó, sự lượng tử hóa này là điểm tham chiếu thấp.

Sự minh họa hiệu suất của phương pháp được mô tả được minh họa bởi Fig.3a đánh giá tỉ số tín hiệu trên nhiễu theo cảm giác kéo dài trung bình và độ phức tạp của các phương pháp như được định rõ bằng các phương trình (1). Có thể thấy rõ ràng rằng, như được liệu trước, sự lượng tử hóa trong miền FFT mang tới tỉ số tín hiệu trên nhiễu kém nhất. Hiệu suất kém có thể là do thực tế là việc lượng tử hóa này không

xem xét tương quan giữa các mẫu dư. Ngoài ra, có thể khẳng định rằng sự lượng tử hóa tối ưu hóa của các tín hiệu dư miền thời gian là bằng với sự tối ưu hóa theo cặp tại 5 và 10 bit mỗi khung, do các tốc độ bit chỉ có 1 hoặc 2 xung, trong đó các phương pháp giống nhau. Đối với 15 bit mỗi khung phương pháp tối ưu tốt hơn một chút so với sự lượng tử hóa theo cặp như được liệu trước.

Tại 10 bit mỗi khung và lớn hơn, sự lượng tử hóa trong miền Vandermonde tốt hơn các lượng tử hóa miền thời gian và miền trị riêng là một bước tốt hơn miền Vandermonde. Tại 5 bit mỗi khung hiệu suất của bộ lập mã số học có khả năng giảm nhanh chóng do nó được biết đến là điểm tối ưu cho các tín hiệu thưa.

Quan sát thấy rằng phương pháp theo cặp khởi động dễ làm lệch phương pháp theo cặp trên 80 bit mỗi khung. Các thí nghiệm không chính thức thể hiện rằng xu hướng này tăng tại tốc độ bit cao hơn sao cho thậm chí FFT và các phương pháp theo cặp đạt được tỉ số tín hiệu trên nhiễu tương tự, thấp hơn nhiều so với các phương pháp trị riêng và các phương pháp Vandermonde. Trái lại, giá trị trị riêng và Vandermonde tiếp tục nhiều hơn hoặc ít hơn các hàm tuyến tính của tốc độ bit. Phương pháp trị riêng xấp xỉ không đổi bằng 0,36dB tốt hơn phương pháp Vandermonde. Giả thiết rằng ít nhất phần khác nhau này được giải thích bằng sự phân tách các phần thực và phần phức tạp trong bộ mã hóa-giải mã số học. Đối với hiệu suất tối ưu, các phần thực và các phần phức tạp cần được mã hóa được khớp nối.

Fig.3b thể hiện phép đo thời gian hoạt động của mỗi phương thức tại từng tốc độ bit để minh họa sự ước lượng độ phức tạp của các thuật toán khác nhau. Có thể thấy rằng độ phức tạp của phương thức miền thời gian tối ưu (so với Opt) đã bùng nổ tại các tốc độ bit thấp. Sự tối ưu hóa theo cặp của phần dư miền thời gian (so với Pair), lần lượt, tăng một cách tuyến tính như tốc độ hàm. Lưu ý rằng trình độ kỹ thuật của các phương pháp hạn chế độ phức tạp của các phương thức theo cặp sao cho nó trở nên bất biến ở các tốc độ bit cao mặc dù các tỉ số tín hiệu trên nhiễu cạnh tranh của các thí nghiệm được minh họa bởi Fig.3a không thể đạt được với các giới hạn này. Ngoài ra, cả hai phương thức giải tương quan (so với Eig và Vand) cùng với phương thức FFT (so với FFT) là bất biến xấp xỉ trên toàn bộ tốc độ bit. Biến đổi Vandermonde có trong việc thực hiện ở trên có độ phức tạp cao hơn 50% so với phương pháp phân tích trị riêng nhưng lý do cho điều này có thể được giải thích bằng việc sử dụng các phiên

bản được tối ưu hóa cao của phân tích trị riêng được cung cấp bởi MATLAB, trong đó việc tìm thừa số Vandermonde không phải là sự thực thi tối ưu. Tuy nhiên, quan trọng là, tại tốc độ bit là 100 bit mỗi khung, ACELP được tối ưu theo cặp phức tạp hơn 30 và 50 lần so với lần lượt là Vandermonde và phân tích trị riêng dựa trên thuật toán. Chỉ FFT nhanh hơn phương pháp phân tích trị riêng, nhưng do tỉ số tín hiệu trên nhiễu của FFT kém, nó không phải là tùy chọn khả thi.

Tóm lại, phương pháp mô tả ở trên có hai lợi ích đáng kể. Thứ nhất, bằng việc áp dụng sự lượng tử hóa trong miền cảm giác, tín hiệu cảm giác tới tỉ số tín hiệu trên nhiễu được cải thiện. Thứ hai, do tín hiệu dư được giải tương quan (đối với hàm mục tiêu) sự lượng tử hóa có thể được áp dụng một cách trực tiếp, mà không có vòng lặp phân tích-tổng hợp phức tạp cao. Theo đó độ phức tạp tính toán của phương pháp được đề xuất là hầu như bất biến đối với các tốc độ bit, trong khi phương thức thông thường trở nên ngày càng phức tạp với tốc độ bit tăng.

Phương thức được trình bày ở trên không được thực hiện hoàn toàn với các phương pháp mã hóa tiếng nói và âm thanh thông thường. Cụ thể là, sự giải tương quan của hàm mục tiêu có thể được áp dụng trong chế độ ACELP của các bộ mã hóa-giải mã như MPEG USAC hoặc AMR-WB+, mà không có sự giới hạn các công cụ khác có mặt trong bộ mã hóa giải mã. Các cách mà trong đó băng thông lỗi hoặc các phương pháp mở rộng băng thông được áp dụng là giống nhau, các cách mà trong đó dự báo dài hạn, sự nâng cao focman, sự lọc trước âm trầm, v.v. trong ACELP không cần phải được thay đổi, và các cách mà các chế độ mã hóa khác nhau này được thực thi (như ACELP và TCX) và chuyển đổi giữa các chế độ này sẽ không bị ảnh hưởng từ giải tương quan của hàm mục tiêu.

Mặt khác, rõ ràng rằng các công cụ (tức là ít nhất tất cả các sự thực hiện ACELP) mà sử dụng hàm đối tượng giống nhau (xem phương trình (1)) có thể được thay đổi công thức để tận dụng giải tương quan. Do đó, theo phương án khác, giải tương quan, ví dụ, tới sự đóng góp dự báo dài hạn có thể được áp dụng và, do đó, các thừa số khuếch đại có thể được tính toán sử dụng tín hiệu được giải tương quan.

Ngoài ra, do miền biến đổi được thể hiện là phép biểu diễn miền tần số, phương pháp truyền thống của các bộ mã hóa-giải mã tiếng nói và âm thanh miền tần số có thể được áp dụng tới miền mới theo các phương án khác. Theo phương án đặc biệt, trong

sự lượng tử hóa vạch phổ, vùng chết có thể được áp dụng cho hiệu suất tăng. Theo phương án khác sự lấp đầy âm có thể được áp dụng để tránh các lỗ phổ.

Mặc dù các phương án mã hóa ở trên (xem các Fig.1a và Fig.1b) đã được thảo luận trong ngữ cảnh của bộ mã hóa âm thanh sử dụng bộ dự báo tuyến tính, cần lưu ý rằng bộ dự báo có thể được tạo cấu hình để chứa bộ dự báo dài hạn để xác định các hệ số dự báo dài hạn mô tả tần số cơ bản của tín hiệu âm thanh AS để lọc tín hiệu âm thanh AS dựa trên hàm bộ lọc được định rõ bởi các hệ số dự báo dài hạn và đề xuất ra tín hiệu dư x để xử lý thêm. Theo phương án khác bộ dự báo có thể là tổ hợp của bộ dự báo tuyến tính và bộ dự báo dài hạn.

Rõ ràng rằng biến đổi được đề xuất có thể được áp dụng một cách dễ dàng cho các tác vụ trong xử lý tiếng nói và âm thanh ví dụ như nâng cao chất lượng tiếng nói. Thứ nhất, phương pháp dựa trên không gian con được dựa trên phân tích trị riêng hoặc phân tích giá trị đơn của tín hiệu. Do phương thức được thể hiện được dựa trên các phân tích tương tự, các phương pháp nâng cao chất lượng tiếng nói dựa trên phép phân tích không gian con có thể được làm thích ứng tới miền đề xuất theo phương án khác. Sự chênh lệch các phương pháp không gian con thông thường là khi mô hình tín hiệu, dựa trên việc dự báo tuyến tính hoặc việc tạo cửa sổ trong miền dư, được áp dụng, ví dụ như được áp dụng trong ACELP. Ngược lại, các phương pháp không gian con truyền thống áp dụng các cửa sổ chồng lấp mà được cố định theo thời gian (không thích ứng).

Thứ hai, giải tương quan dựa trên giải tương quan Vandermonde cung cấp miền tần số tương tự để được cung cấp bởi biến đổi Fourier rời rạc, biến đổi cosin hoặc các biến đổi tương tự. Do đó, bất kỳ thuật toán xử lý tiếng nói mà thường thực hiện trong miền biến đổi Fourier rời rạc, biến đổi cosin hoặc biến đổi tương tự có thể được áp dụng với các sự cải biên tối thiểu trong các miền biến đổi của phương thức được mô tả ở trên. Do đó, việc nâng cao chất lượng tiếng nói sử dụng phép trừ phổ trong miền biến đổi có thể được áp dụng, theo như các phương án nữa, sự biến đổi được đề xuất có thể được sử dụng trong việc nâng cao chất lượng tiếng nói hoặc âm thanh, ví dụ với phương pháp trừ phổ, phân tích không gian con hoặc các đạo hàm hoặc cải biên của chúng. Trong tài liệu này, các lợi ích là phương thức này sử dụng việc tạo cửa sổ tương tự như ACELP để thuật toán nâng cao chất lượng tiếng nói có thể được tích hợp

chặt chẽ vào bộ mã hóa-giải mã tiếng nói. Ngoài ra, cửa sổ ACELP có độ trễ thuật toán logarit thấp hơn so với cửa sổ được sử dụng trong phân tích không gian con thông thường. Kết quả là, do đó việc tạo cửa sổ được dựa trên mô hình tín hiệu có hiệu năng cao hơn.

Đề cập tới phương trình (5) mà được sử dụng cho bộ biến đổi 14, tức là, trong bước 140, cần lưu ý rằng việc tạo lập có thể cũng khác nhau, ví dụ, trong dạng $y = DVx$.

Theo phương án khác, bộ mã hóa 10 có thể bao gồm bộ đóng gói tại đầu ra được tạo cấu hình để đóng gói hai dòng dữ liệu DS_{LPC}/DS_{DV} và DS_y cho gói DS thông thường. Ngược lại, bộ giải mã 20 có thể bao gồm bộ giải gói được tạo cấu hình để chia dòng dữ liệu DS thành hai gói DS_{LPC}/DS_{DV} và DS_y .

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

Các tín hiệu được mã hóa theo sáng chế có thể được lưu trữ trên môi trường lưu trữ kỹ thuật số hoặc có thể được truyền trong môi trường truyền như môi trường truyền không dây hoặc môi trường truyền có dây như Internet.

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

Một số phương án theo sáng chế bao gồm vật mang dữ liệu có các tín hiệu điều

khiến 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 trong tài liệu này được thực hiện.

Thông thường, các phương án theo sáng chế có thể được thực thi như sản phẩm chương trình máy tính với mã chương trình, mã chương trình hoạt động để thực hiện một trong các phương pháp khi sản phẩm chương trình máy tính chạy trên máy tính. Mã chương trình có thể ví dụ được lưu trữ trên vật 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 trong tài liệu này, 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 các phương pháp được mô tả trong tài liệu này, 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 trong tài liệu này. Sóng mang dữ liệu, môi trường lưu trữ số hoặc môi trường đã được ghi thường là hữu hình và/hoặc không chuyển tiếp.

Phương án khác của các phương pháp theo sáng chế là, dòng dữ liệu hoặc chuỗi tín hiệu thể hiện chương trình máy tính để thực hiện một trong các phương pháp đã được mô tả trong tài liệu này. Ví dụ, có thể cấu hình dòng dữ liệu hoặc chuỗi tín hiệu để được truyền thông qua sự kết nối truyền thông dữ liệu, ví dụ thông qua 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 hoặc được làm thích ứng để thực hiện một trong số các phương pháp được mô tả trong tài liệu này.

Phương án khác gồm có máy tính có chương trình máy tính được lắp trên máy tính để thực hiện một trong các phương pháp được mô tả trong tài liệu này.

Phương án nữa theo sáng chế bao gồm thiết bị hoặc hệ thống được tạo cấu hình để truyền (ví dụ, bằng điện tử hoặc quang học) chương trình máy tính để thực hiện một trong các phương pháp được mô tả trong tài liệu này đến bộ nhận. Ví dụ, bộ nhận có thể là máy tính, thiết bị di động, thiết bị nhớ hoặc thiết bị tương tự. Thiết bị hoặc hệ

thống có thể, ví dụ, bao gồm máy chủ tệp tin để chuyển chương trình máy tính đến bộ nhận.

Theo một số phương án, thiết bị logic lập trình được (ví dụ, mảng cổng lập trình được dạng trường) có thể được sử dụng để thực hiện một số hoặc tất cả các chức năng của các phương pháp được mô tả trong tài liệu này. 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 trong tài liệu này. Thông thường, các phương pháp ưu tiên được thực hiện bởi thiết bị phần cứng bất kỳ.

Các chỉ dẫn được mô tả ở trên sẽ được thảo luận phía dưới với cách diễn đạt khác và một số chi tiết có thể trợ giúp để làm sáng tỏ nền tảng của sáng chế. Biến đổi Vandermonde được biểu diễn gần đây như biến đổi tần số thời gian mà, khác với biến đổi Fourier rời rạc, cũng giải tương quan tín hiệu. Mặc dù giải tương quan xấp xỉ hoặc tiệm cận được cung cấp bởi Fourier là đủ trong nhiều trường hợp, hiệu suất của nó là không thích hợp trong các ứng dụng mà dùng các cửa sổ ngắn. Do đó, biến đổi Vandermonde sẽ hữu ích trong các ứng dụng xử lý tiếng nói và âm thanh, mà sử dụng các cửa sổ phân tích ngắn do tín hiệu đầu vào biến thiên nhanh chóng qua thời gian. Các ứng dụng thường được sử dụng trên thiết bị di động với dung lượng tính toán hạn chế, trong đó các tính toán hiệu quả là hết sức quan trọng.

Tuy nhiên, việc thực hiện biến đổi Vandermonde hóa ra là nỗ lực đáng kể: nó yêu cầu các công cụ bằng số tiên tiến mà hiệu suất của nó được tối ưu cho độ phức tạp và độ chính xác. Sự đóng góp này cung cấp giải pháp cơ sở cho tác vụ này bao gồm việc đánh giá hiệu suất. Các thuật ngữ đại diện-các biến đổi tần số thời gian, giải tương quan, ma trận Vandermonde, ma trận Toeplitz, biến đổi Fourier rời rạc uốn cong.

Biến đổi Fourier rời rạc là một trong những công cụ cơ bản nhất trong xử lý tín hiệu kỹ thuật số. Nó cung cấp phép biểu diễn được thúc đẩy tự nhiên của tín hiệu đầu vào ở dạng các thành phần tần số. Do biến đổi Fourier nhanh (Fast Fourier Transform - FFT) tính toán biến đổi Fourier rời rạc cũng với độ phức tạp tính toán thấp $O(N \log N)$, nó trở thành một trong những công cụ quan trọng nhất trong xử lý âm thanh kỹ thuật số.

Mặc dù được khen ngợi, nhưng biến đổi Fourier rời rạc có thiếu sót: Nó không giải tương quan các thành phần tín hiệu hoàn toàn (nhiều ví dụ, xem Phần 4). Chỉ khi

độ dài biến đổi hội tụ về vô cực khiến các thành phần trở nên trực giao. Giải tương quan xấp xỉ này có trong rất nhiều ứng dụng đủ tốt. Tuy nhiên, các ứng dụng mà dùng các biến đổi tương đối nhỏ như các thuật toán xử lý tiếng nói và âm thanh, độ chính xác của sự xấp xỉ này hạn chế toàn bộ hiệu quả của các thuật toán. Ví dụ, tiêu chuẩn lập mã tiếng nói WB-AMR dùng các cửa sổ có độ dài $N = 64$. Thực tế cho thấy hiệu suất của biến đổi Fourier rời rạc trong trường hợp này là không đủ và kết quả là các bộ mã hóa-giải mã tiếng nói chính sử dụng sự mã hóa miền thời gian.

Fig.3c thể hiện các đặc tính của biến đổi Vandermonde, đường đậm 51 minh họa phổ Fourier (không uốn cong) của tín hiệu và đường 52, 53 và 54 là đáp ứng của bộ lọc thông băng của ba tần số được lựa chọn, được lọc với tín hiệu đầu vào. Kích thước việc tìm thừa số Vandermonde là 64.

Có nhiều biến đổi tự nhiên mà cung cấp giải tương quan của tín hiệu đầu vào, như biến đổi Karhunen-Loève (Karhunen-Loève transform - KLT). Tuy nhiên, các thành phần KLT là các thực thể trừu tượng mà không có diễn dịch vật lý đơn giản như biến đổi Fourier. Mặt khác, miền được thúc đẩy tự nhiên, cho phép thực hiện đơn giản các tiêu chí thúc đẩy tự nhiên vào các phương pháp xử lý. Do đó, biến đổi mà cung cấp cả diễn dịch vật lý và giải tương quan được mong muốn.

Biến đổi được biểu diễn gần đây, được gọi là biến đổi Vandermonde, mà có cả hai đặc tính được ưu tiên. Điều này dựa trên sự phân giải ma trận Hermitian Toeplitz thành tích của ma trận đường chéo và ma trận Vandermonde. Việc tìm thừa số này thực sự được biết đến như sự tham số hóa Carathéodory của các ma trận hiệp phương sai và tương tự như việc tìm thừa số Vandermonde của các ma trận Hankel.

Đối với các trường hợp đặc biệt của các ma trận Hermitian Toeplitz xác định dương, việc tìm thừa số Vandermonde sẽ tương ứng với biến đổi Fourier rời rạc tần số được uốn cong. Nói cách khác, biến đổi thời gian - tần số mà cung cấp các thành phần tín hiệu được lấy mẫu tại các tần số mà không được nhất thiết được phân bố một cách đồng nhất. Do đó biến đổi Vandermonde cung cấp các thuộc tính được mong muốn: giải tương quan và diễn giải vật lý.

Trong khi sự tồn tại và các thuộc tính của biến đổi Vandermonde được mô tả theo phép phân tích, mục đích của công việc hiện thời, trước hết, để thu thập và cung cấp tư liệu về các thuật toán thực tiễn tồn tại cho các biến đổi Vandermonde. Các

phương pháp này xuất hiện trong nhiều lĩnh vực khác nhau, bao gồm đại số bằng số, phân tích bằng số, nhận diện hệ thống, phân tích thời gian - tần số và xử lý tín hiệu, mà người ta thường khó tìm thấy. Do đó tài liệu này xem xét các phương pháp mà cung cấp các nền tảng kết nối cho việc phân tích và thảo luận các kết quả. Thứ hai, các mẫu đại số được cung cấp như cơ sở cho sự đánh giá khác về hiệu suất của các phương pháp khác nhau.

Phần này cung cấp giới thiệu ngắn gọn về các biến đổi Vandermonde. Để thúc đẩy toàn diện và thảo luận về các ứng dụng, đề cập đến:

Ma trận Vandermonde V được định rõ bởi các đại lượng vô hướng v_k như:

$$V = \begin{bmatrix} 1 & v_0 & v_0^2 & \cdots & v_0^{N-1} \\ 1 & v_1 & v_1^2 & \cdots & v_1^{N-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & v_{N-1} & v_{N-1}^2 & \cdots & v_{N-1}^{N-1} \end{bmatrix} \quad (1z)$$

Nó là ma trận hạng đầy đủ nếu các đại lượng vô hướng v_k khác biệt ($v_k \neq v_h$ cho $k \neq h$) và nghịch đảo của nó có công thức tường minh.

Ma trận Toeplitz T được định rõ bởi các đại lượng vô hướng τ_k như:

$$T = \begin{bmatrix} \tau_0 & \tau_1 & \cdots & \tau_{N-1} \\ \tau_1 & \tau_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \tau_1 \\ \tau_{N-1} & \cdots & \tau_1 & \tau_0 \end{bmatrix} \quad (2z)$$

Nếu T là xác định dương, thì nó có thể được tìm thừa số như:

$$T = V^* \Lambda V, \quad (3z)$$

trong đó Λ là ma trận đường chéo với các chỉ mục thực và dương hoàn toàn $\lambda_{kk} > 0$ và các chuỗi lũy thừa V đều nằm trên vòng tròn đơn vị $v_k = \exp(i\beta_k)$. Dạng này cũng được biết đến như sự tham số hóa Carathéodory của ma trận Toeplitz.

Hai cách sử dụng biến đổi Vandermonde được trình bày trong tài liệu này: như biến đổi giải tương quan hoặc như sự thay thế cho ma trận nhân chập. Xem xét tín hiệu x mà có ma trận tương quan tự động $E[xx^*] = R_x$. Do ma trận tương quan tự động là xác định dương, đối xứng và Toeplitz, nó có thể được tìm thừa số như $R = V^* \Lambda V$. Theo đó nếu áp dụng biến đổi:

$$y_d = V^{-*} x \quad (4z)$$

trong đó V^{-*} là Hermitian nghịch đảo của V , thì ma trận tương quan tự động của y_d là:

$$R_y = E[y_d y_d^*] = V^{-*} E[xx^*] V^{-1} = V^{-*} R_x V^{-1} = V^{-*} V^* \Lambda V V^{-1} = \Lambda \quad (5z)$$

Do đó tín hiệu biến đổi y_d là không tương quan tự động. Biến đổi nghịch đảo là

$$\mathbf{x} = \mathbf{V}^* \mathbf{y}_d. \quad (6z)$$

Như mô tả phỏng đoán, có thể nói rằng biến đổi thuận chiều $\mathbf{V}^{-*}$ chứa trong hàng k th của nó bộ lọc mà băng thông của nó tại tần số $-\beta_k$ và đầu ra bằng đúng cho $\mathbf{x}$ có năng lượng thấp. Cụ thể là, hình dạng phổ của đầu ra gần với bộ lọc AR với cực đơn trên vòng tròn đơn vị. Lưu ý rằng do dàn bộ lọc thích ứng tín hiệu, đầu ra của bộ lọc được xem xét hơn là đáp ứng tần số của các hàm cơ bản.

Biến đổi ngược chiều $\mathbf{V}^*$ lần lượt có các chuỗi lũy thừa trong các cột của nó, sao cho $\mathbf{x}$ là tổng được gán trọng số của các chuỗi lũy thừa. Nói cách khác, biến đổi là biến đổi thời gian - tần số được uốn. Fig.3c mô tả phổ Fourier (không được uốn) rời rạc của tín hiệu đầu vào $\mathbf{x}$ và các đáp ứng tần số của các hàng được lựa chọn của $\mathbf{V}^{-*}$.

Biến đổi Vandermonde để đánh giá tín hiệu trong miền cuốn chap có thể được xây dựng như sau. Để $\mathbf{C}$ là ma trận nhân chap và $\mathbf{x}$ là tín hiệu đầu vào. Xem xét trường hợp trong đó đối tượng là để đánh giá tín hiệu được nhân chap $\mathbf{y}_c = \mathbf{C}\mathbf{x}$. Sự đánh giá này xuất hiện, ví dụ, trong bộ mã hóa-giải mã tiếng nói sử dụng ACELP, trong đó năng lượng lỗi lượng tử hóa được đánh giá trong miền cảm giác và trong đó sự ánh xạ tới miền cảm giác được mô tả bằng bộ lọc.

Năng lượng $\mathbf{y}_c$ là:

$$\|\mathbf{y}_c\|^2 = \|\mathbf{C}\mathbf{x}\|^2 = \mathbf{x}^* \mathbf{C}^* \mathbf{C} \mathbf{x} = \mathbf{x}^* \mathbf{R}_c \mathbf{x} = \mathbf{x}^* \mathbf{V}^* \mathbf{\Lambda} \mathbf{V} \mathbf{x} = \|\mathbf{\Lambda}^{1/2} \mathbf{V} \mathbf{x}\|^2 \quad (7z)$$

Do đó, năng lượng $\mathbf{y}_c$ bằng với năng lượng của tín hiệu được định tỉ lệ và được biến đổi:

$$\mathbf{y}_v = \mathbf{\Lambda}^{1/2} \mathbf{V} \mathbf{x} \quad (8z)$$

Năng lượng tín hiệu có thể được đánh giá một cách tương đương trong miền cuốn chap hoặc trong miền được biến đổi. $\|\mathbf{y}_c\|^2 = \|\mathbf{y}_v\|^2$. Biến đổi nghịch đảo là hiển nhiên:

$$\mathbf{x} = \mathbf{V}^{-1} \mathbf{\Lambda}^{-1/2} \mathbf{y}_v. \quad (9z)$$

Biến đổi thuận chiều $\mathbf{V}$ có các chuỗi lũy thừa trong các hàng của nó, trong đó nó là biến đổi Fourier được uốn cong. Nghịch đảo của nó $\mathbf{V}^{-1}$ có các bộ lọc trong các cột của nó, với các băng thông tại β_k . Trong dạng này đáp ứng tần số của giàn bộ lọc bằng với biến đổi Fourier rời rạc. Nó chỉ là biến đổi nghịch đảo mà sử dụng những gì

thường được xem như các thành phần răng cưa để kích hoạt sự khôi phục hoàn hảo.

Để sử dụng các biến đổi Vandermonde, cần các thuật toán hiệu quả để xác định cũng như áp dụng các biến đổi. Trong phần này các thuật toán khả dụng sẽ được thảo luận. Bắt đầu với ứng dụng các biến đổi do nó là tác vụ không phức tạp.

Các phép nhân với V and V^* là đơn giản và có thể được thực hiện trong $O(N^2)$.

Để giảm các yêu cầu lưu trữ, các thuật toán được thể hiện trong đó các lũy thừa v_k^h cần được đánh giá một cách rõ ràng đối với $h > 1$. Cụ thể là, nếu $y = Vx$ và các phần tử của x là ξ_k , thì các phần tử η_k của y có thể được xác định với phép truy toán:

$$\begin{cases} \tau_{h,0} = \xi_{N-1} \\ \tau_{h,k} = \xi_{N-1-k} + v_h \tau_{h,k-1}, & 1 \leq k \leq N \\ \eta_h = \tau_{h,N-1}. \end{cases} \quad (10z)$$

Ở đây, $\tau_{h,k}$ là đại lượng vô hướng tạm thời, mà chỉ giá trị hiện thời cần được lưu trữ. Toàn bộ phép truy toán có N bước cho N thành phần, trong đó toàn bộ độ phức tạp là $O(N^2)$ lưu trữ không đổi. Thuật toán tương tự có thể được viết một cách dễ dàng cho $y = V^*x$.

Phép nhân với các ma trận Vandermonde nghịch đảo V^{-1} and V^{*-1} là tác vụ phức tạp hơn nhưng may mắn là các phương pháp tương đối hiệu quả đã khả dụng từ phần tổng quát. Các thuật toán là đơn giản để thực hiện và cho cả $x = V^{-1}y$ và $x = V^{*-1}y$ độ phức tạp là $O(N^2)$ và lưu trữ tuyến tính $O(N)$. Tuy nhiên, thuật toán gồm phép chia tại mỗi bước, mà có trong nhiều cấu trúc chi phí không đổi cao.

Mặc dù các thuật toán ở trên cho các phép nhân bởi các nghịch đảo được trích xuất theo cảm giác phân tích, việc thực hiện trong thực tế là không ổn định về số lượng đối với N lớn. Theo kinh nghiệm, sự tính toán với các ma trận lên tới kích thước $N \sim 64$ đôi khi khả thi, nhưng ngoài các bộ kết xuất không ổn định số học các thuật toán này là vô ích. Giải pháp thực tế là thứ tự Leja của các nghiệm v_k là bằng với phép khử Gaussian với chốt bán phần. Ý tưởng chính đằng sau thứ tự Leja là sắp xếp lại các nghiệm theo cách mà khoảng cách của nghiệm v_k tới phần tử trước của nó là $0 \dots (k-1)$ được tối đa hóa. Bằng việc kết xuất các mẫu số xuất hiện trong thuật toán được tối đa hóa và các giá trị biến số trung gian được tối thiểu hóa, trong đó sự đóng góp của các sai số cắt được tối thiểu hóa. Sự thực hiện thứ tự Leja là đơn giản và có thể đạt được với độ phức tạp $O(N^2)$ và bộ nhớ $O(N)$.

Sau đó trở ngại cuối cùng là thu được việc tìm thừa số, mà là, các nghiệm v_k và khi cần thiết, các giá trị đường chéo λ_{kk} . Do biết được rằng các nghiệm có thể được thu được bằng việc giải:

$$Ra = [1 \ 1 \ \dots \ 1]^T, \quad (11z)$$

trong đó có các phần tử α_k . thì $v_0 = 1$ và các nghiệm còn lại là $v_1 \dots v_N$ là các nghiệm của $A(z) = \sum_{k=0}^{N-1} \alpha_k z^{-k}$. Có thể thể hiện dễ dàng rằng điều này tương đương với việc giải quyết hệ thống Hankel:

$$\begin{bmatrix} \tau_{N-1} & \dots & \tau_1 & \tau_0 \\ \vdots & \ddots & \tau_0 & \tau_1 \\ \tau_1 & \ddots & \ddots & \vdots \\ \tau_0 & \tau_1 & \dots & \tau_{N-1} \end{bmatrix} \begin{bmatrix} \hat{a}_1 \\ \hat{a}_2 \\ \vdots \\ \hat{a}_N \end{bmatrix} = - \begin{bmatrix} \tau_1 \\ \tau_2 \\ \vdots \\ \tau_N \end{bmatrix} \quad (12z)$$

trong đó $\tau_N = \frac{1}{\alpha_0} \sum_{k=1}^{N-1} \alpha_{k+1} \tau_{N-k}$. Các nghiệm v_k là các nghiệm của $\hat{A}(z) = 1 + \sum_{k=1}^N \hat{a}_k z^{-k}$.

Do việc tìm thừa số của phương trình hệ thống Toeplitz gốc 11z tương đương với phương trình 12z, có thể sử dụng thuật toán nhanh cho việc tìm thừa số của các ma trận Hankel. Thuật toán này trả về ma trận ba đường chéo mà các giá trị riêng của nó tương ứng với các nghiệm của $\hat{A}(z)$. Các giá trị riêng sau đó có thể thu được trong $O(N^2)$ bằng việc áp dụng thuật toán LR, hoặc trong $O(N^3)$ bằng thuật toán QR không đối xứng tiêu chuẩn. Các nghiệm thu được theo các này là các giá trị gần đúng, trong đó chúng có thể nằm ngoài xa vòng tròn đơn vị. Nó hữu ích để chuẩn hóa giá trị tuyệt đối của các nghiệm để thống nhất, và lọc với 2 hoặc 3 lần lặp theo phương pháp Newton. Quy trình hoàn thành có phí tính toán là $O(N^2)$.

Bước cuối cùng trong việc tìm thừa số là để thu các giá trị đường chéo Λ . Thấy rằng:

$$Re = V^* \Lambda V e = V^* \lambda \quad (13z)$$

trong đó $e = [1 \ 0 \ \dots \ 0]^T$ và λ là vectơ chứa các giá trị đường chéo Λ . Nói cách khác, bằng việc tính toán:

$$\lambda = V^{-*}(Re), \quad (14z)$$

Thu được các giá trị đường chéo λ_{kk} . Nghịch đảo này có thể được tính toán với các phương pháp được thảo luận ở trên, trong đó các giá trị đường chéo được thu với độ phức tạp $O(N^2)$.

Tóm lại, các bước được yêu cầu cho việc tìm thừa số của ma trận R là:

1. Giải phương trình 11z cho việc sử dụng Levinson-Durbin hoặc các phương pháp cổ điển khác.

2. Mở rộng chuỗi tương quan bằng $r_N = \frac{1}{\alpha_0} \sum_{k=1}^{N-1} \alpha_{k+1} r_{N-k}$.

3. Áp dụng thuật toán ba đường chéo của chuỗi r_k .

4. Giải các giá trị riêng v_k sử dụng LR hoặc thuật toán QR đối xứng.

5. Lọc các vị trí nghiệm bằng việc định tỉ lệ v_k để hợp nhất và một số phép lặp theo phương pháp Newton.

6. Xác định các giá trị đường chéo λ_{kk} sử dụng phương trình 14z.

Bắt đầu với các ví dụ số học để minh họa các khái niệm được sử dụng. Ở đây, ma trận C là ma trận nhân chập đáp ứng bộ lọc không quan trọng $1 + z^{-1}$, ma trận R là ma trận tương quan tự động, ma trận V ma trận Vandermonde tương ứng được thu được với thuật toán trong phần 3, ma trận F là ma trận biến đổi Fourier rời rạc và các ma trận Λ_V và Λ_F minh họa độ chính xác chéo hóa của hai biến đổi. Do đó có thể định rõ:

$$\mathbf{C} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}, \quad \mathbf{R} = \mathbf{C}\mathbf{C}^* = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix} \quad (15z)$$

$$\mathbf{V} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & i & -1 \\ 1 & -1 & -1 \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & e^{+i\pi/3} & e^{+i\pi/3} \\ 1 & e^{+i\pi/3} & e^{-i\pi/3} \end{bmatrix}$$

trong đó có thể đánh giá sự chéo hóa với:

$$\Lambda_V = |\mathbf{V}^{-*}\mathbf{R}\mathbf{V}^{-1}| = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.5 & 0 \\ 0 & 0 & 0.5 \end{bmatrix} \quad (16z)$$

$$\Lambda_F = |\mathbf{F}^{-*}\mathbf{R}\mathbf{F}^{-1}| = \begin{bmatrix} 1.11 & 0.111 & 0.111 \\ 0.111 & 0.444 & 0.222 \\ 0.111 & 0.222 & 0.444 \end{bmatrix}$$

Ở đây có thể thấy rằng với biến đổi Vandermonde thu được ma trận đường chéo hoàn hảo Λ_V . Sự biểu diễn biến đổi Fourier rời rạc là không tối ưu hóa, do các giá trị ngoài đường chéo là rõ ràng khác 0. Như một biện pháp thực hiện, có thể tính toán tỉ số của các tổng tuyệt đối của các giá trị ngoài và trong đường chéo, mà bằng 0 cho việc tìm thừa số Vandermonde và 0,444 cho biến đổi Fourier.

Sau đó có thể tiến hành đánh giá sự thực hiện được mô tả trong phần 3. Mỗi thuật toán trong MATLAB được thực hiện với mục đích cung cấp cơ sở thực hiện mà

các công việc trong tương lai có thể so sánh và để tìm ra các tắc nghẽn thực hiện có thể xảy ra. Sự thực hiện độ phức tạp và độ chính xác sẽ được xem xét.

Để xác định sự thực hiện tìm thừa số, việc tìm thừa số Vandermonde biến đổi Fourier rời rạc và biến đổi Karhunen-Loève sẽ được so sánh, sau đó được áp dụng với phân tích trị riêng. Việc tìm thừa số Vandermonde được áp dụng sử dụng hai phương pháp, thứ nhất, thuật toán được mô tả trong mục này (V_1), và thứ hai phương thức được mô tả trong việc sử dụng hàm tìm nghiệm được cài đặt sẵn được cung cấp bởi MATLAB (V_2). Do hàm MATLAB này là thuật toán chung được tinh chỉnh, các kết quả chính xác được mong đợi thu được nhưng với độ phức tạp cao hơn các thuật toán được cài đặt sẵn có mục đích.

Do dữ liệu cho tất cả thí nghiệm, tập hợp các mẫu tiếng nói, âm thanh và âm thanh được trộn được sử dụng trong việc đánh giá tiêu chuẩn MPEG USAC với tốc độ lấy mẫu 12.8 kHz. Các mẫu âm thanh được tạo cửa sổ với các cửa sổ Hamming tới độ dài mong muốn và tương quan tự động của chúng được tính toán. Để đảm bảo các ma trận tương quan tự động là xác định dương, đường chéo chính được nhân với $(1+10^{-5})$.

Đối với các biện pháp thực hiện, độ phức tạp tính toán được sử dụng về thời gian hoạt động chuẩn hóa và độ chính xác về độ gần $\tilde{\mathbf{A}} = \mathbf{V}^{-*}\mathbf{R}\mathbf{V}^{-1}$ tới ma trận đường chéo, được đo bởi tỉ số tổng tuyệt đối của các thành phần ngoài đường chéo và trong đường chéo. Các kết quả được thống kê trong bảng 1 và bảng 2.

Bảng 1. Độ phức tạp của các thuật toán tìm thừa số cho các độ dài cửa sổ khác nhau N về thời gian hoạt động được chuẩn hóa.

N	16	32	64	128	256	512
V_1	1,00	3,02	10,13	35,96	131,80	496,91
V_2	1,00	2,10	8,77	90,61	634,17	4056,62
KLT	1,00	4,33	8,93	30,59	109,53	419,76

Bảng 2. Độ chính xác của các thuật toán tìm thừa số cho các độ dài cửa sổ N về $\log_{10}$ của tỉ số của tổng tuyệt đối của các giá trị ngoài đường chéo và trong đường chéo $\tilde{\mathbf{A}} = \mathbf{V}^{-*}\mathbf{R}\mathbf{V}^{-1}$

N	16	32	64	128	256	512
FFT	-0,22	-0,16	-0,13	-0,11	-0,08	-0,07
V ₁	-2,36	-2,14	-1,93	-1,72	-1,26	-0,97
V ₂	-13,00	-13,56	-13,11	-12,67	-12,14	-11,56
KLT	-14,56	-14,24	-14,07	-13,89	-13,65	-13,23

Lưu ý rằng ở đây không phải là hợp lý để so sánh thời gian chạy giữa các thuật toán, chỉ có sự gia tăng trong độ phức tạp như là một hàm của kích thước khung, do các hàm MATLAB được cài đặt sẵn đã được thực thi trong nhiều ngôn ngữ khác với các thuật toán này. Có thể thấy rằng độ phức tạp của thuật toán được đề xuất V₁ tăng với tỉ lệ so sánh như KLT, khi thuật toán sử dụng các hàm tìm nghiệm của MATLAB V₂ tăng hơn. Độ chính xác của thuật toán tìm thừa số được đề xuất V₁ là không tối ưu. Tuy nhiên, do hàm tìm nghiệm của MATLAB V₂ sinh ra độ chính xác như KLT, những cải thiện này được kết luận rằng khả thi bởi các cải thiện thuật toán.

Thí nghiệm thứ hai là ứng dụng các biến đổi để xác định độ chính xác và độ phức tạp. Thứ nhất, sử dụng các phương trình 4z và 9z, mà độ phức tạp của nó được liệt kê trong bảng 3. Tại đây có thể thấy rằng phép nhân ma trận KLT và phép giải cài đặt sẵn của các hệ thống ma trận của MATLAB V₂ có mức tăng gần như nhau trong độ phức tạp, trong khi các phương pháp được đề xuất cho các phương trình 4z và 9z có độ tăng nhỏ hơn nhiều. FFT nhanh hơn tất cả các phương pháp khác một cách tự nhiên.

Cuối cùng, để thu độ chính xác của các giải pháp Vandermonde, áp dụng các biến đổi xuôi chiều và thuận chiều theo trình tự. Các khoảng cách Euclidean giữa các vector ban đầu và các vector được khôi phục được liệt kê trong bảng 4. Có thể quan sát rằng, thứ nhất, các thuật toán FFT và KLT là, như được kỳ vọng, là chính xác nhất, do chúng dựa trên các biến đổi trực tâm. Thứ hai, có thể thấy rằng độ chính xác của thuật toán V₁ được đề xuất thấp hơn một ít so với phép giải MATLAB V₂, nhưng cả hai thuật toán cung cấp độ chính xác đủ.

Sự thực thi chi tiết của các biến đổi thời gian - tần số giải tương quan sử dụng việc tìm thừa số Vandermonde được biểu diễn với mục đích xem lại các thuật toán khả thi cũng như cung cấp cơ sở thực hiện cho để phát triển thêm trong khi các thuật toán

trên nguyên tắc khả thi từ công việc trước đó, nó chỉ ra rằng để có được hệ thống để chạy các yêu cầu.

Bảng 3. Độ phức tạp của các phép giải Vandermonde cho các độ dài của số khác nhau N về thời gian hoạt động được chuẩn hóa. V_1^{-*} và V_1^{-1} biểu thị nghiệm của các phương trình 4x và 9z với các thuật toán được đề xuất lần lượt

N	16	32	64	128	256	512
FFT	1,00	1,13	1,31	1,99	2,96	3,82
V_1^{-*}	1,00	2,00	4,30	10,17	24,52	68,56
V_1^{-1}	1,00	1,99	4,26	10,14	24,64	69,49
V2	1,00	1,86	7,57	23,16	78,44	284,80
KLT	1,00	1,31	5,37	8,55	46,25	289,30

Bảng 4. Độ chính xác của các biến đổi xuôi chiều và ngược chiều như được đo bởi $\log_{10} (\|\mathbf{x} - \hat{\mathbf{x}}\|^2 / \|\mathbf{x}\|^2)$, trong đó $\mathbf{x}$ và $\hat{\mathbf{x}}$ là các vectơ gốc và được khôi phục

N	16	32	64	128	256	512
FFT	-15,82	-15,71	-15,66	-15,62	-15,58	-15,55
V_1^{-*}	-14,62	-14,07	-13,43	-12,89	-12,40	-12,11
V_1^{-1}	-15,15	-14,84	-14,51	-14,14	-13,78	-13,42
V2	-15,38	-15,22	-15,00	-14,80	-14,67	-14,52
KLT	-14,98	-14,85	-14,78	-14,70	-14,61	-14,51

nỗ lực đáng kể. Thách thức chính là độ chính xác số học và độ phức tạp tính toán. Thí nghiệm khẳng định rằng các phương pháp khả thi với độ phức tạp $O(N^2)$, mặc dù thu được độ phức tạp đồng thời với sự ổn định số học là thách thức. Tuy nhiên, do những thực hiện MATLAB tổng quát cung cấp lời giải chính xác, việc thu được độ chính xác cao có thể được khẳng định là khả thi với việc điều chỉnh thực hiện khác.

Tóm lại, thí nghiệm cho thấy rằng lời giải Vandermonde, các thuật toán được đề xuất có độ chính xác tốt và độ phức tạp thấp vừa đủ. Đối với việc tìm thừa số, các việc tìm thừa số được xây dựng theo mục đích đưa ra sự tương quan tốt hơn so với FFT với

độ phức tạp hợp lý, nhưng chính xác là có chỗ để cải thiện. Các thực hiện gắn liền của MATLAB đưa ra độ chính xác thỏa đáng, mà đưa đến kết luận rằng các thuật toán $O(N^2)$ chính xác có thể được thực thi.

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 trong tài liệu này sẽ rõ ràng với những người có trình độ trung bình trong lĩnh vực kỹ thuật. Do đó, sáng chế chỉ bị giới hạn bởi phạm vi của các điểm yêu cầu bảo hộ sắp đưa ra dưới đây và không bị giới hạn bởi các chi tiết cụ thể được biểu diễn bằng cách mô tả và giải thích của các phương án trong tài liệu này.

YÊU CẦU BẢO HỘ

1. Bộ mã hóa (10) để mã hóa tín hiệu âm thanh (AS) vào dòng dữ liệu (DS), bao gồm:

bộ dự báo (12) được tạo cấu hình để phân tích tín hiệu âm thanh (AS) để thu các hệ số dự báo (LPC) mô tả đường bao phổ của tín hiệu âm thanh (AS) hoặc tần số cơ bản của tín hiệu âm thanh (AS) và đưa tín hiệu âm thanh (AS) tới hàm bộ lọc phân tích (H) phụ thuộc vào các hệ số dự báo (LPC) để xuất ra tín hiệu dư (x) của tín hiệu âm thanh (AS);

bộ tìm thừa số (14) được tạo cấu hình để áp dụng sự tìm thừa số ma trận trên ma trận tương quan tự động hoặc ma trận hiệp phương sai (R, C) của hàm bộ lọc tổng hợp (H) được định rõ bởi các hệ số dự báo (LPC) để thu các ma trận được tìm thừa số (V, D);

bộ biến đổi (16) được tạo cấu hình để biến đổi tín hiệu dư (x) dựa trên các ma trận được tìm thừa số (V, D) để thu được tín hiệu dư được biến đổi (y); và

tăng lượng tử hóa và mã hóa (18) được tạo cấu hình để lượng tử hóa tín hiệu dư được biến đổi (y) để thu tín hiệu dư được biến đổi được lượng tử hóa ($\hat{y}$) và bao gồm bộ mã hóa entropy có đầu vào cho các hệ số dự báo (LPC) và được tạo cấu hình để mã hóa entropy tín hiệu dư được biến đổi được lượng tử hóa ($\hat{y}$) với việc phát hiện xác suất dựa trên các hệ số dự báo (LPC) để thu tín hiệu dư được biến đổi được lượng tử hóa được mã hóa ($\hat{y}$).

2. Bộ mã hóa (10) theo điểm 1, trong đó hàm bộ lọc tổng hợp (H) được định rõ bởi ma trận (H) bao gồm các giá trị được gán trọng số của hàm bộ lọc tổng hợp (H).

3. Bộ mã hóa (10) theo điểm 1 hoặc điểm 2, trong đó bộ tìm thừa số (14) tính toán ma trận tương quan tự động hoặc ma trận hiệp phương sai (R, C) dựa trên tích số của các phiên bản được biến đổi được liên hợp của hàm bộ lọc tổng hợp (H^*) và phiên bản thông thường của hàm bộ lọc tổng hợp (H).

4. Bộ mã hóa (10) theo một điểm trong số các điểm từ 1 đến 3, trong đó bộ tìm thừa số (14) tìm thừa số của ma trận tương quan tự động và ma trận hiệp phương sai (R, C) dựa trên công thức $C = V^*DV$ hoặc dựa trên công thức $R = V^*DV$;

trong đó V là ma trận Vandermonde, V^* là phiên bản được biến đổi được liên hợp của ma trận Vandermonde và D là ma trận đường chéo với các mục nhập dương

hoàn toàn.

5. Bộ mã hóa (10) theo điểm 4, trong đó bộ tìm thừa số (14) được tạo cấu hình để thực hiện phép tìm thừa số Vandermonde.
6. Bộ mã hóa (10) theo một trong số các điểm từ điểm 1 đến điểm 5, trong đó bộ tìm thừa số (14) được tạo cấu hình để thực hiện phân tích trị riêng và/hoặc việc tìm thừa số Cholesky.
7. Bộ mã hóa (10) theo một trong số các điểm 4 hoặc 5, trong đó bộ biến đổi (16) biến đổi tín hiệu dư (x) dựa trên công thức $y = D^{1/2}Vx$ hoặc dựa trên công thức $y = DVx$.
8. Bộ mã hóa (10) theo một trong số các điểm từ 1 đến 7, trong đó tầng lượng tử hóa và mã hóa (18) lượng tử hóa tín hiệu dư được biến đổi ($\hat{y}$) để thu tín hiệu dư được biến đổi được lượng tử hóa ($\hat{y}$) dựa trên hàm mục tiêu $\eta(y) = \frac{(y * \hat{y})^2}{\|\hat{y}\|^2}$.
9. Bộ mã hóa (10) theo một trong số các điểm từ 1 đến 8, trong đó tầng lượng tử hóa và mã hóa (18) bao gồm phương thức để tối ưu hóa việc lượng tử hóa bằng cách áp dụng lấp đầy nhiễu âm để cung cấp phép biểu diễn phổ lấp đầy nhiễu âm của tín hiệu âm thanh (AS), tín hiệu dư (x) hoặc tín hiệu dư được biến đổi ($\hat{y}$) và hoặc bằng cách tối ưu hóa tín hiệu dư được biến đổi được lượng tử hóa ($\hat{y}$) liên quan đến các vùng chết hoặc liên quan đến các tham số lượng tử hóa khác.
10. Bộ mã hóa (10) theo một trong số các điểm từ 1 đến 9, trong đó sự biến đổi của tín hiệu dư (x) là sự biến đổi từ miền thời gian của tín hiệu dư (x) tới miền giống tần số của tín hiệu dư được biến đổi (y).
11. Bộ mã hóa (10) theo một trong số các điểm từ 1 đến 10, trong đó tầng lượng tử hóa và mã hóa bao gồm bộ mã hóa được tạo cấu hình để thực hiện việc mã hóa của tín hiệu dư được biến đổi được lượng tử hóa ($\hat{y}$) để thu được tín hiệu dư được biến đổi được lượng tử hóa được mã hóa ($\hat{y}'$).
12. Bộ mã hóa (10) theo điểm 11 trong đó việc mã hóa được thực hiện bởi bộ mã hóa không thuộc nhóm bao gồm việc lập mã số học.
13. Bộ mã hóa (10) theo điểm 11 hoặc điểm 12, trong đó bộ mã hóa (10) còn bao gồm bộ đóng gói được tạo cấu hình để đóng gói tín hiệu dư được biến đổi được lượng

tử hóa được mã hóa ($\hat{y}$) và các hệ số dự báo (LPC) tới dòng dữ liệu (DS) để được xuất ra bằng bộ mã hóa (10).

14. Bộ mã hóa (10) theo một trong số các điểm từ điểm 1 đến điểm 13, trong đó bộ dự báo (12) bao gồm bộ dự báo tuyến tính (và/hoặc bộ dự báo dài hạn).

15. Phương pháp (100) để mã hóa tín hiệu âm thanh (AS) thành dòng dữ liệu (DS), phương pháp bao gồm:

phân tích (120) tín hiệu âm thanh (AS) để thu được các hệ số dự báo (LPC) mô tả đường bao phổ của tín hiệu âm thanh (AS) hoặc tần số cơ bản của tín hiệu âm thanh (AS) và đưa tín hiệu âm thanh (AS) tới hàm bộ lọc phân tích (H) phụ thuộc vào các hệ số dự báo (LPC) để xuất ra tín hiệu dư (x) của tín hiệu âm thanh (AS);

áp dụng (140) việc tìm thừa số ma trận trên ma trận tương quan tự động hoặc ma trận hiệp phương sai (R, C) của hàm bộ lọc tổng hợp (H) được định rõ bởi các hệ số dự báo (LPC) để thu các ma trận được tìm thừa số (V, D);

biến đổi (160) tín hiệu dư (x) dựa trên các ma trận được tìm thừa số (V, D) để thu tín hiệu dư được biến đổi (y); và

lượng tử hóa và mã hóa (180) tín hiệu dư được biến đổi (y) để thu tín hiệu dư được biến đổi được lượng tử hóa ($\hat{y}$) và sự mã hóa entropy sử dụng các hệ số dự báo (LPC) tín hiệu dư được biến đổi được lượng tử hóa ($\hat{y}$) với việc phát hiện xác suất dựa trên các hệ số dự báo (LPC) để thu tín hiệu dư được biến đổi được lượng tử hóa được mã hóa ($\hat{y}$).

16. Phương pháp (100) theo điểm 15 trong đó phương pháp đã nêu được sử dụng thay thế cho biến đổi Fourier rời rạc, biến đổi cosin rời rạc, biến đổi cosin rời rạc cải biên hoặc biến đổi khác trong các thuật toán xử lý tín hiệu.

17. Bộ giải mã (20) giải mã dòng dữ liệu (DS) thành tín hiệu âm thanh (AS'), bao gồm:

tăng giải mã (22) được tạo cấu hình để xuất ra tín hiệu dư được biến đổi (y) dựa trên tín hiệu dư được biến đổi được lượng tử hóa được mã hóa trở lại ($\hat{y}$) sử dụng sự mã hóa entropy với việc phát hiện xác suất dựa trên các hệ số dự báo (LPC) mô tả đường bao phổ của tín hiệu âm thanh (AS) hoặc tần số cơ bản của tín hiệu âm thanh (AS);

bộ biến đổi lại (26) được tạo cấu hình để biến đổi lại tín hiệu dư (x) từ tín hiệu dư được biến đổi (y) dựa trên các ma trận được tìm thừa số (V, D) biểu diễn kết quả của việc tìm thừa số ma trận của ma trận tương quan tự động hoặc ma trận hiệp phương sai (R, C) của hàm bộ lọc tổng hợp (H) được định rõ bằng các hệ số dự báo (LPC); và

tầng tổng hợp (28) được tạo cấu hình để tổng hợp tín hiệu âm thanh (AS') dựa trên tín hiệu dư (x) bằng việc sử dụng hàm bộ lọc tổng hợp (H) được định rõ bằng các hệ số dự báo (LPC).

18. Bộ giải mã (20) theo điểm 17 trong đó bộ giải mã (20) bao gồm bộ tìm thừa số (24) được tạo cấu hình để áp dụng việc tìm thừa số ma trận trên ma trận tương quan tự động hoặc ma trận hiệp phương sai (R, C) của hàm bộ lọc tổng hợp (H) được định rõ bởi các hệ số dự báo trở lại (LPC) để thu được các ma trận được tìm thừa số (V, D).

19. Bộ giải mã (20) theo điểm 17, trong đó bộ giải mã (20) bao gồm bộ tạo các hệ số dự báo được tạo cấu hình để làm lệch các hệ số dự báo (LPC) dựa trên các ma trận được tìm thừa số trở lại (V, D).

20. Bộ giải mã (20) theo một trong số các điểm từ 17 đến 19, trong đó tầng giải mã (22) thực hiện việc giải mã dựa trên các quy tắc mã hoá đã biết và/hoặc các tham số mã hóa được làm lệch từ các quy tắc lập mã trở lại và/hoặc tham số lập mã.

21. Phương pháp (200) để giải mã dòng dữ liệu (DS) thành tín hiệu âm thanh (AS'), phương pháp bao gồm các bước:

xuất ra (220) tín hiệu dư được biến đổi (y) dựa trên tín hiệu dư được biến đổi được lượng tử hóa được mã hóa trở lại ($\hat{y}$) sử dụng việc giải mã entropy với việc phát hiện xác suất dựa trên các hệ số dự báo (LPC) mô tả các đường bao phổ của tín hiệu âm thanh (AS) hoặc tần số cơ bản của tín hiệu âm thanh (AS);

áp dụng (240) việc tìm thừa số ma trận trên ma trận tương quan tự động hoặc ma trận hiệp phương sai (R, C) của hàm bộ lọc tổng hợp (H) được định rõ bằng các hệ số dự báo (LPC), mô tả (240) đường bao phổ của tín hiệu âm thanh (AS) hoặc tần số cơ bản của tín hiệu âm thanh (AS) để thu các ma trận được tìm thừa số (V, D);

biến đổi lại (260) tín hiệu dư (x) từ tín hiệu dư được biến đổi lại (y) dựa trên các ma trận được tìm thừa số (V, D); và

tổng hợp (280) tín hiệu âm thanh (AS') dựa trên tín hiệu dư (x) bằng cách sử

dụng hàm bộ lọc tổng hợp (H) được định rõ bằng các hệ số dự báo (LPC).

22. Vật ghi lưu trữ kỹ thuật số có thể đọc được bằng máy tính lưu trữ trên đó chương trình máy tính có mã chương trình để thực hiện, khi chạy trên máy tính, phương pháp (100) theo điểm 15.

23. Vật ghi lưu trữ kỹ thuật số có thể đọc được bằng máy tính lưu trữ trên đó chương trình máy tính có mã chương trình để thực hiện, khi chạy trên máy tính, phương pháp (200) theo điểm 21.

24. Vật ghi lưu trữ kỹ thuật số có thể đọc được bằng máy tính lưu trữ trên đó dòng dữ liệu (DS) bao gồm tín hiệu âm thanh (AS), bao gồm:

phần thứ nhất (DSVD) bao gồm các ma trận được tìm thừa số (V, D), kết quả từ việc tìm thừa số ma trận trên ma trận tương quan tự động hoặc ma trận hiệp phương sai (R, C) của hàm bộ lọc tổng hợp (H) được định rõ bởi các hệ số dự báo (LPC), và các hệ số dự báo (LPC), mô tả đường bao phổ của tín hiệu âm thanh (AS) hoặc tần số cơ bản của tín hiệu âm thanh (AS); và

phần thứ hai DS_y bao gồm tín hiệu dư (x) của tín hiệu âm thanh (AS), sau đó đưa tín hiệu âm thanh (AS) tới hàm bộ lọc phân tích (H) phụ thuộc vào các hệ số dự báo (LPC), dưới dạng tín hiệu dư được biến đổi được lượng tử hóa được mã hóa ($\hat{y}$) được thu bằng cách mã hóa entropy sử dụng các hệ số dự báo (LPC) tín hiệu dư được biến đổi được lượng tử hóa ($\hat{y}$) với việc phát hiện xác suất dựa trên các hệ số dự báo (LPC).

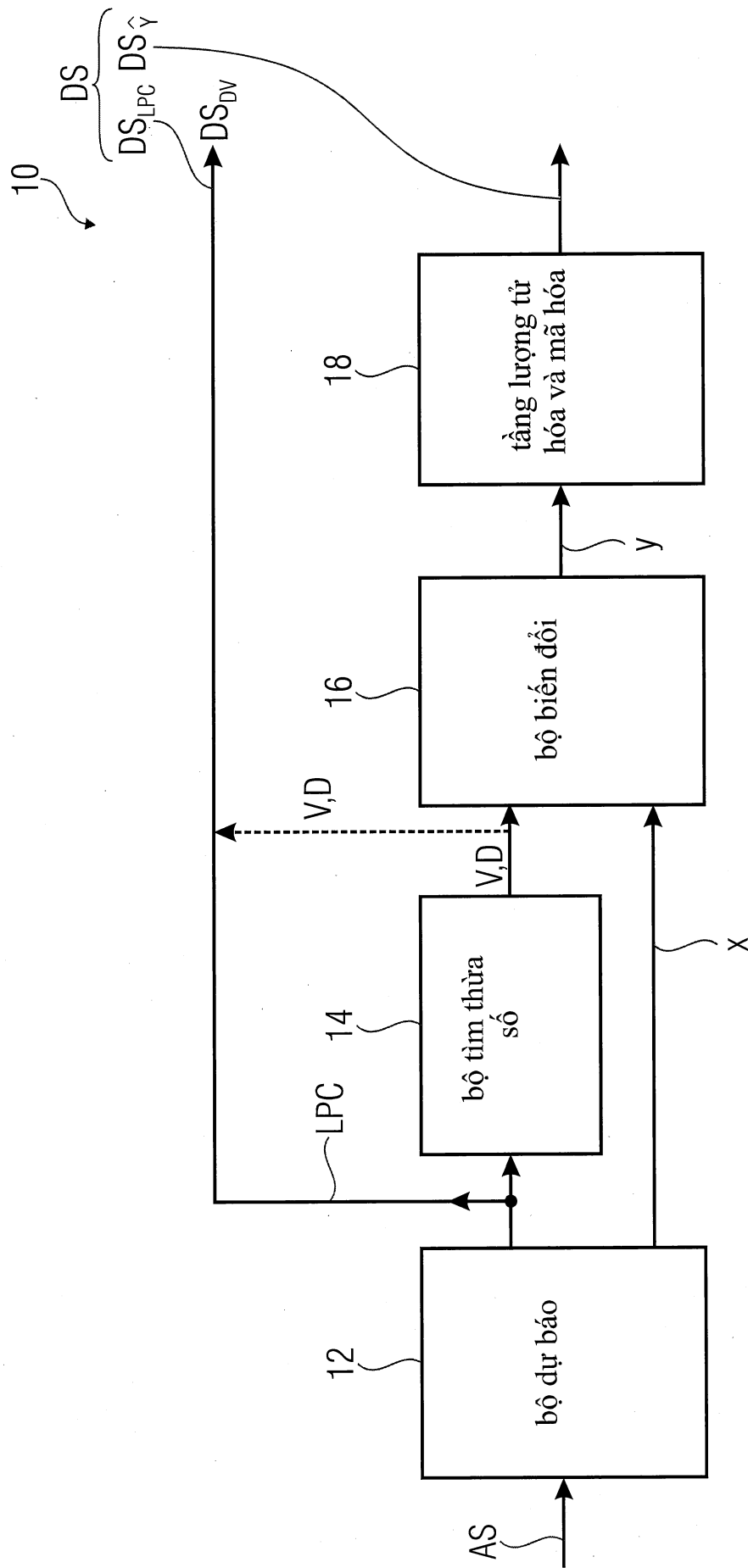


FIG 1A

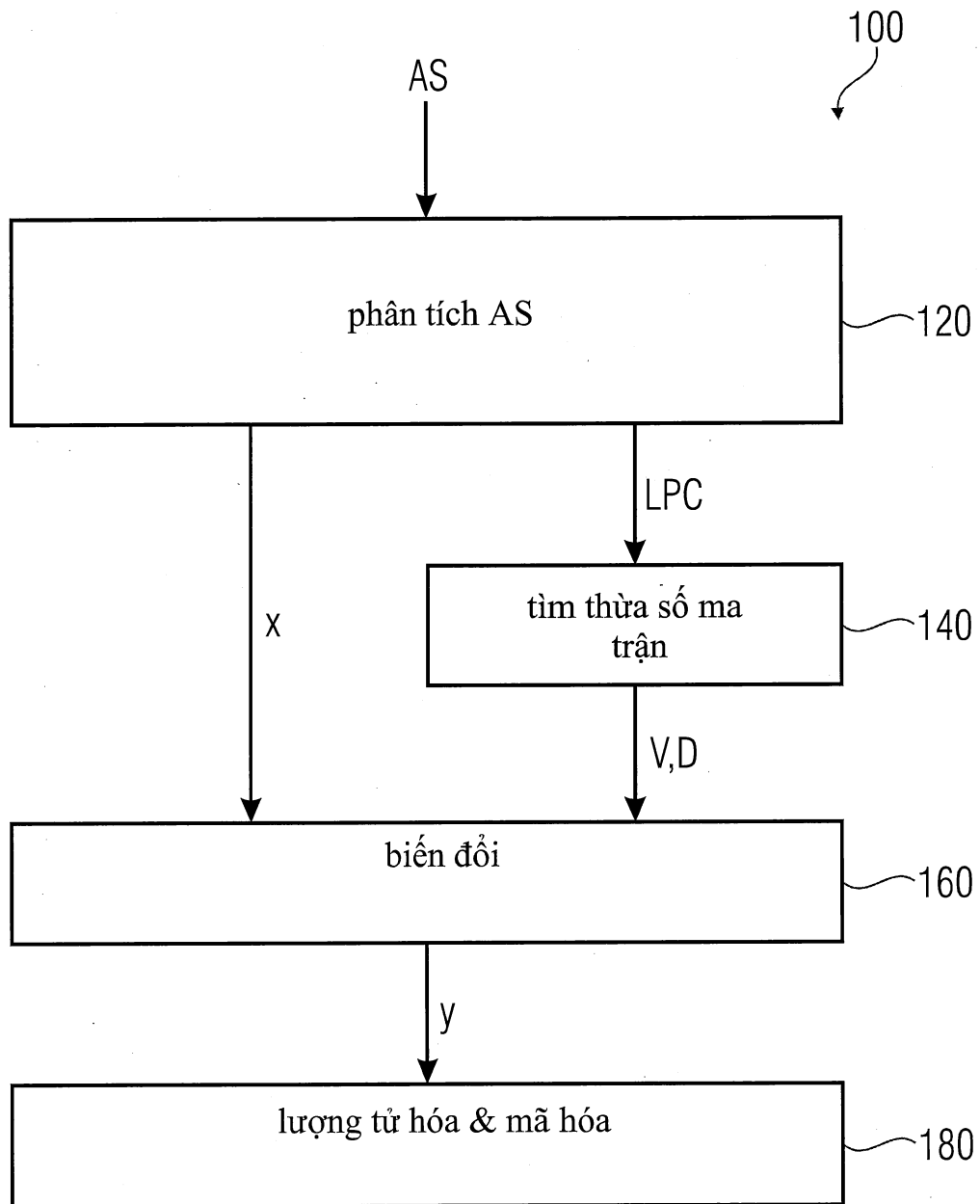


FIG 1B

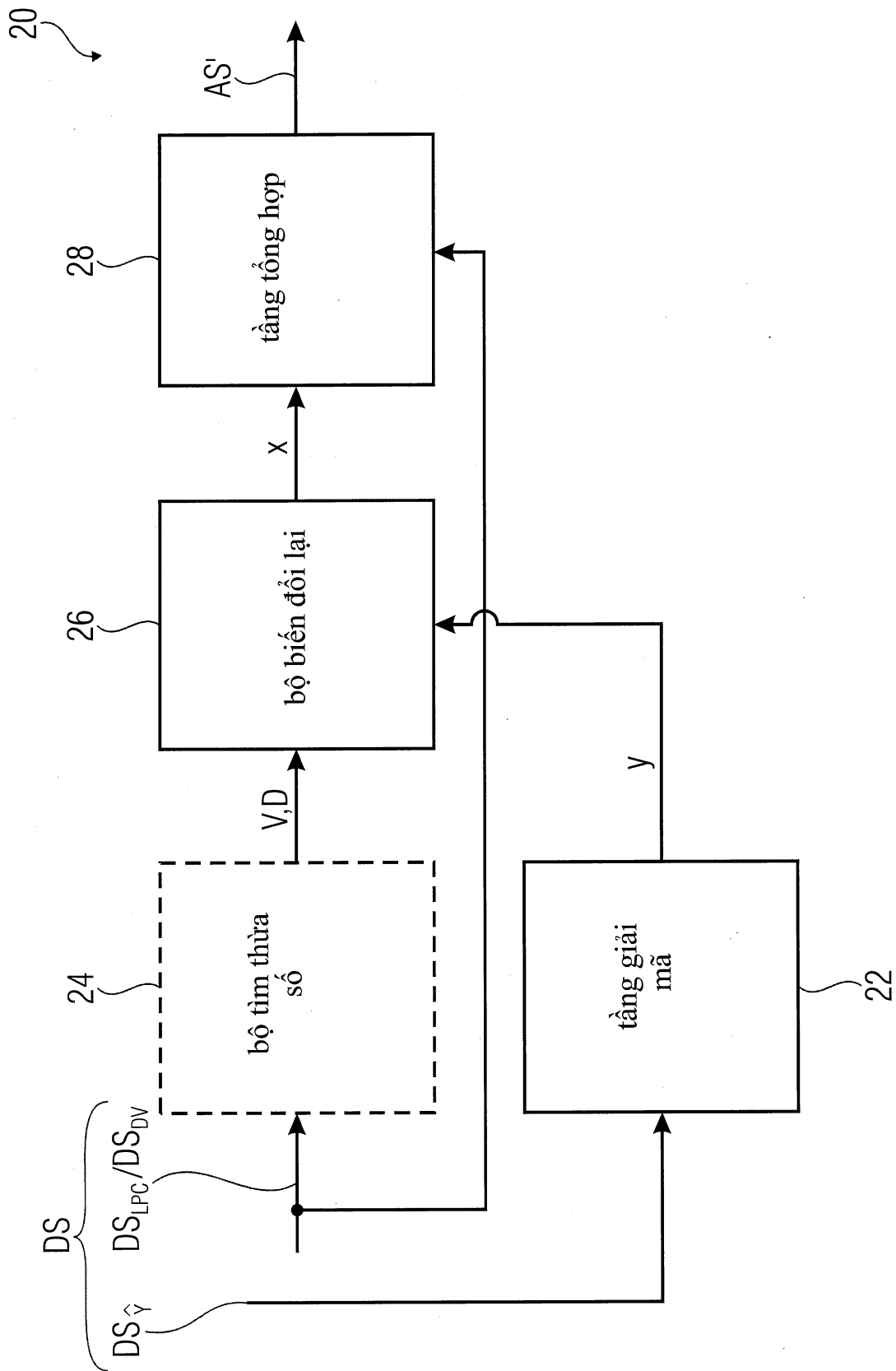


FIG 2A

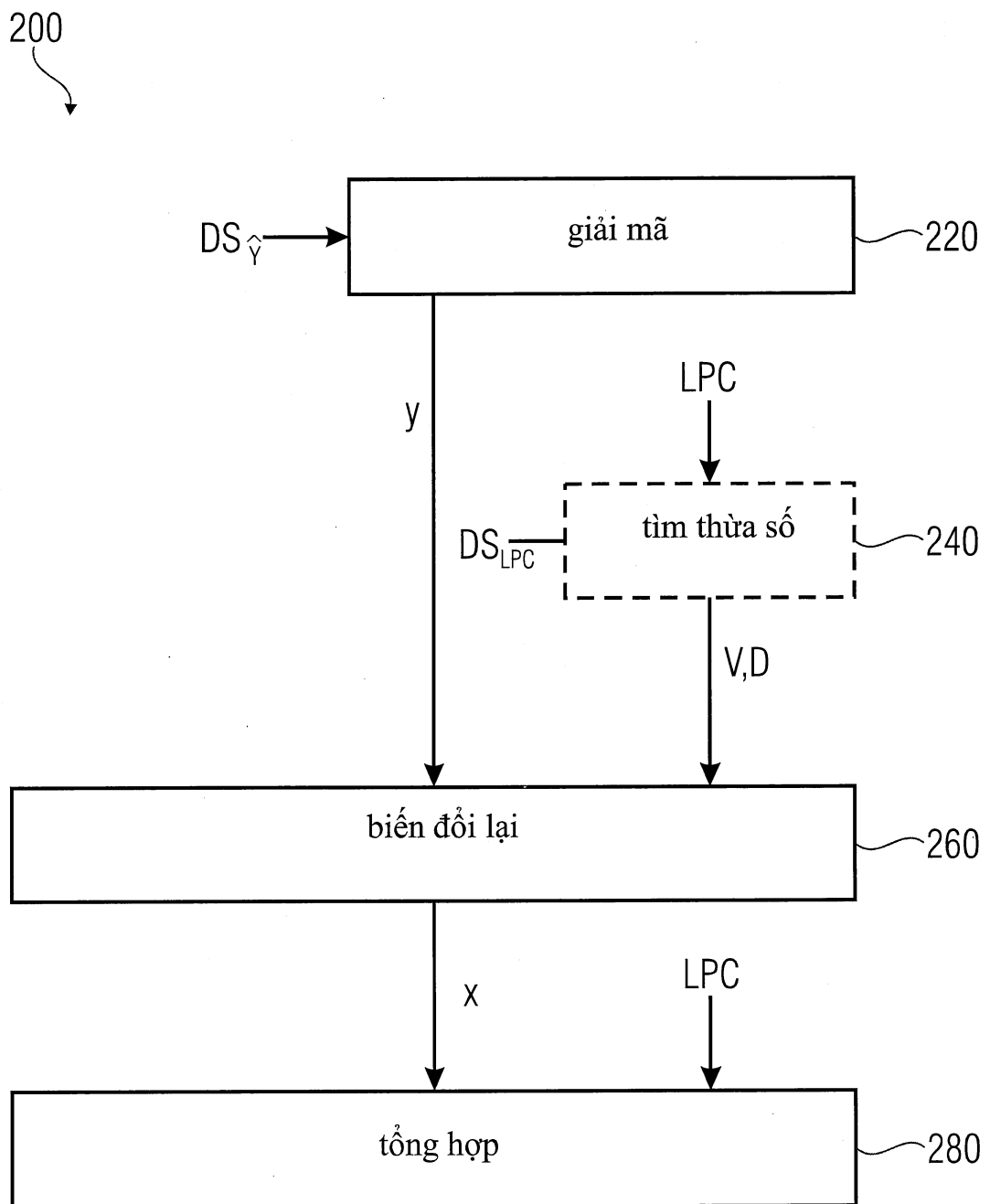


FIG 2B

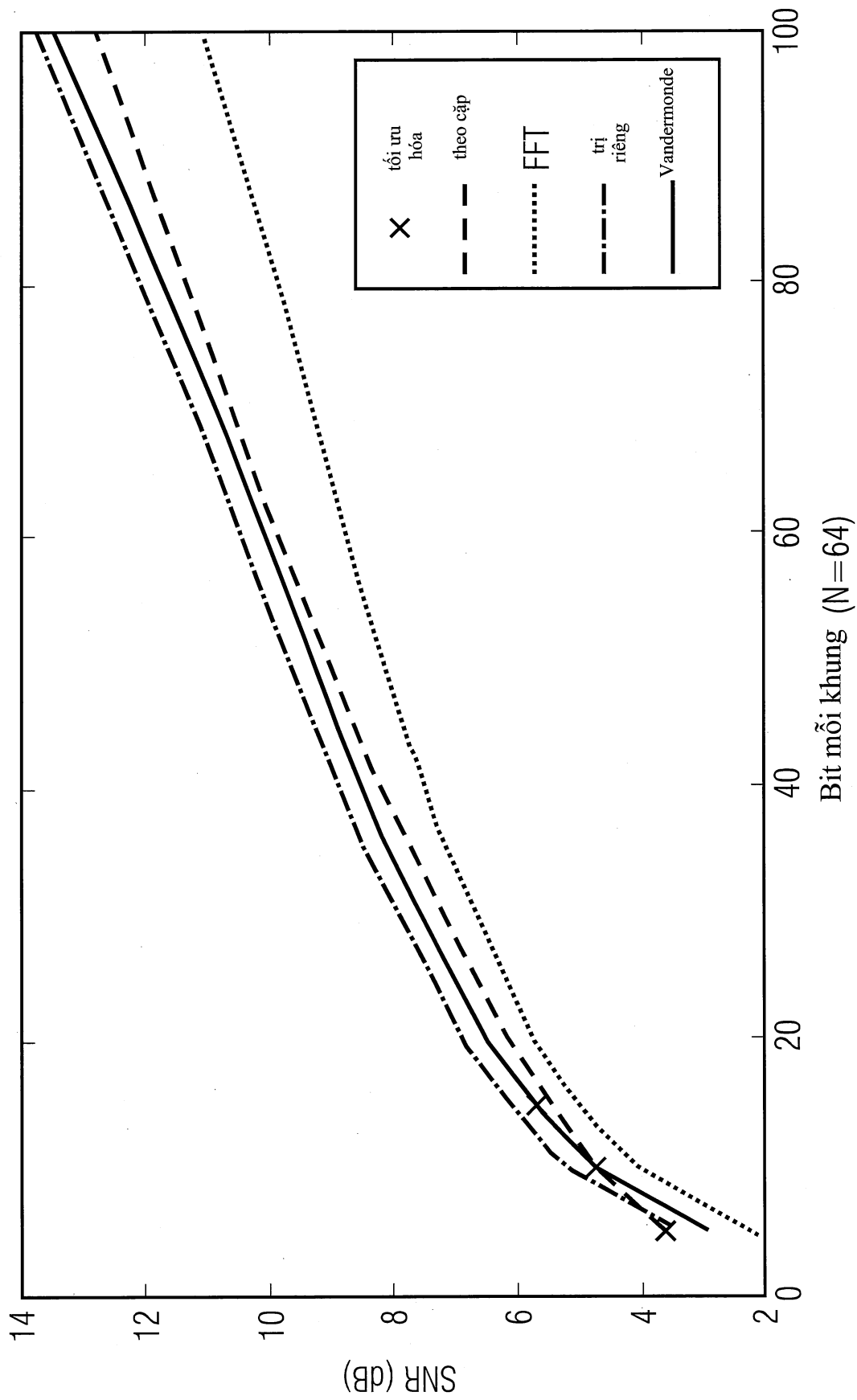


FIG 3A

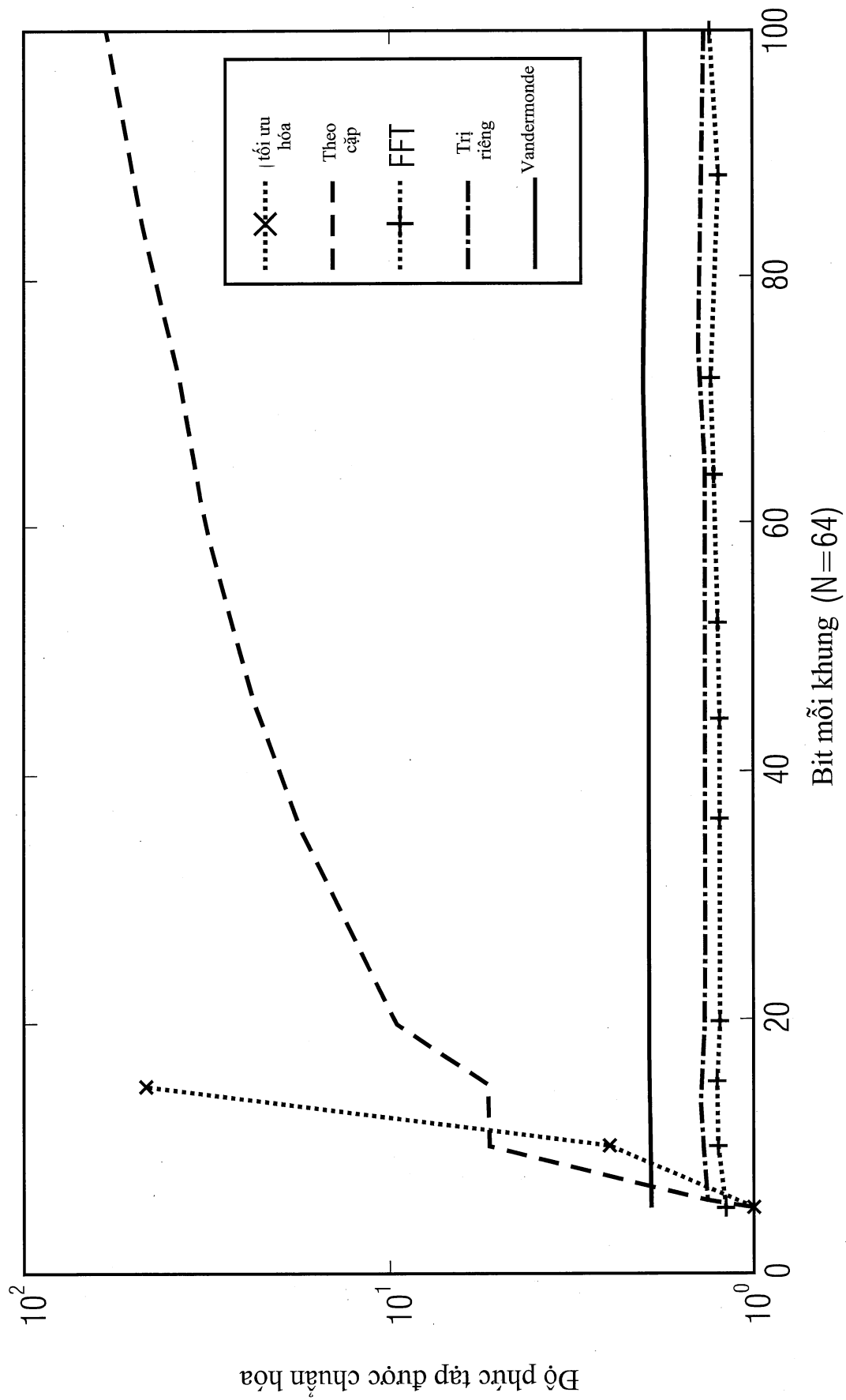


FIG 3B

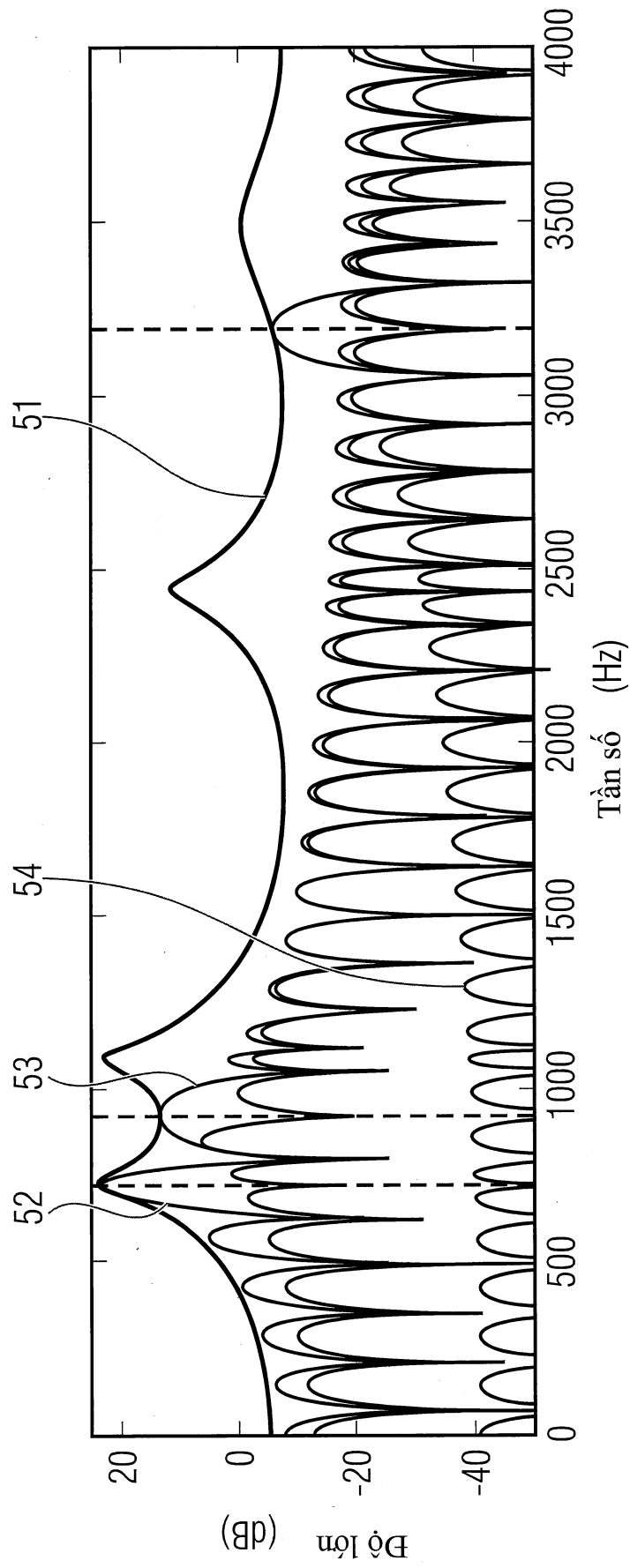


FIG 3C