



(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)



1-0030675

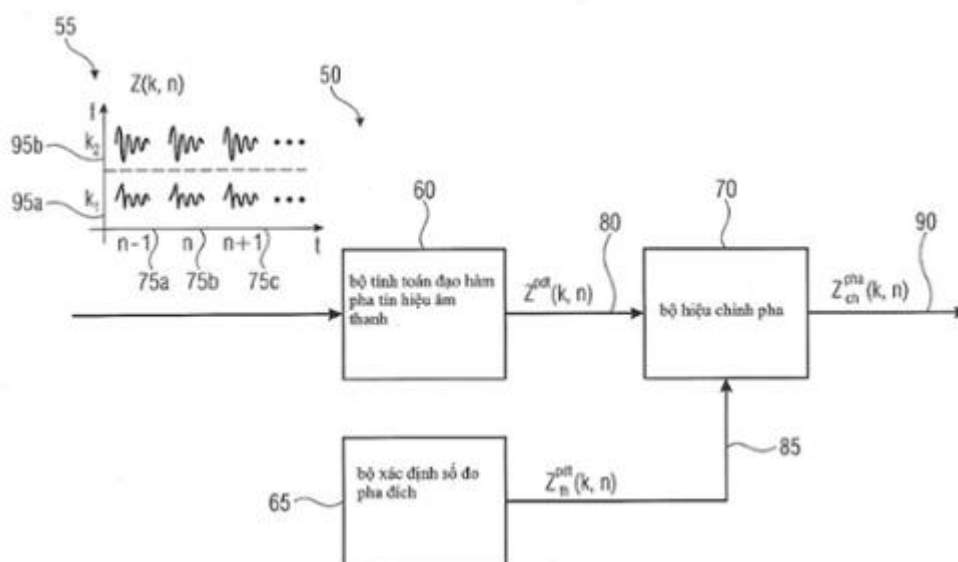
(51)⁷ G10L 21/038

(13) B

- (21) 1-2016-05144 (22) 25/06/2015
(86) PCT/EP2015/064443 25/06/2015 (87) WO 2016/001069 A1 07/01/2016
(30) 14175202.2 01/07/2014 EP; 15151478.3 16/01/2015 EP
(45) 25/01/2022 406 (43) 27/03/2017 348A
(73) FRAUNHOFER-GESELLSCHAFT ZUR FOERDERUNG DER ANGEWANDTEN
FORSCHUNG E.V. (DE)
Hansastraße 27c, 80686 Muenchen, Germany (DE)
(72) DISCH, Sascha (DE); LAITINEN, Mikko-Ville (FI); PULKKI, Ville (FI).
(74) CÔNG TY LUẬT TRÁCH NHIỆM HỮU HẠN AMBYS HÀ NỘI (AMBYS
HANOI)

(54) BỘ XỬ LÝ ÂM THANH VÀ PHƯƠNG PHÁP XỬ LÝ TÍN HIỆU ÂM THANH,
BỘ GIẢI MÃ VÀ PHƯƠNG PHÁP GIẢI MÃ TÍN HIỆU ÂM THANH, BỘ MÃ
HÓA VÀ PHƯƠNG PHÁP MÃ HÓA TÍN HIỆU ÂM THANH

(57) Sáng chế đề cập đến bộ xử lý âm thanh và phương pháp xử lý tín hiệu âm thanh, bộ giải mã và phương pháp giải mã tín hiệu âm thanh, bộ mã hóa và phương pháp mã hóa tín hiệu âm thanh. Sáng chế thể hiện bộ xử lý âm thanh (50) để xử lý tín hiệu âm thanh (55). Bộ xử lý âm thanh bao gồm bộ tính toán số đo pha tín hiệu âm thanh (60) được tạo cấu hình để tính toán số đo pha (80) của tín hiệu âm thanh cho khung thời gian (75a), bộ xác định số đo pha đích (65) để xác định số đo pha đích (85) cho khung thời gian thứ nhất (75a) đã nêu, và bộ hiệu chỉnh pha (70) được tạo cấu hình để hiệu chỉnh các pha (45) của tín hiệu âm thanh (55) cho khung thời gian (75a) sử dụng số đo pha được tính toán (80) và số đo pha đích (85) để thu tín hiệu âm thanh được xử lý (90).



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến bộ xử lý âm thanh và phương pháp xử lý tín hiệu âm thanh, bộ giải mã và phương pháp giải mã tín hiệu âm thanh, và bộ mã hóa và phương pháp mã hóa tín hiệu âm thanh. Hơn nữa, sáng chế còn đề cập đến bộ tính toán và phương pháp xác định dữ liệu hiệu chỉnh pha, tín hiệu âm thanh, và chương trình máy tính để thực hiện một trong số các phương pháp đã được đề cập trước đó. Nói cách khác, sáng chế đề cập đến sự hiệu chỉnh đạo hàm pha và sự mở rộng băng thông (bandwidth extension - BWE) cho các bộ mã hóa-giải mã âm thanh cảm quan hoặc hiệu chỉnh phổ pha của các tín hiệu mở rộng băng thông trong miền QMF trên cơ sở tầm quan trọng cảm quan.

Tình trạng kỹ thuật của sáng chế

Mã hóa âm thanh cảm quan

Việc mã hóa âm thanh cảm quan đến nay tuân theo một số chủ đề thông thường, bao gồm sử dụng việc xử lý miền thời gian/tần số, sự giảm phần dư (mã hóa entropy), và sự loại bỏ không thích hợp thông qua việc khai thác rõ rệt các hiệu quả cảm quan [1]. Thông thường, tín hiệu đầu vào được phân tích bởi giàn lọc phân tích mà chuyển đổi tín hiệu miền thời gian thành phép biểu diễn phổ (thời gian/tần số). Việc chuyển đổi thành các hệ số phổ cho phép xử lý một cách chọn lọc các thành phần tín hiệu phụ thuộc vào nội dung tần số của chúng (ví dụ, các nhạc khí khác nhau với cấu trúc họa âm cao riêng của chúng).

Đồng thời, tín hiệu đầu vào được phân tích đối với các đặc điểm cảm quan của nó, tức là, cụ thể, ngưỡng che phụ thuộc thời gian và phụ thuộc tần số được tính toán. Ngưỡng che phụ thuộc thời gian/tần số được phân phối đến bộ phận lượng tử hóa qua ngưỡng mã hóa đích dưới dạng giá trị năng lượng tuyệt đối hoặc tỉ lệ mạng che so với tín hiệu (Mask-to-Signal-Ratio - MSR) đối với mỗi băng và khung thời gian mã hóa.

Các hệ số phổ được phân phối bởi giàn lọc phân tích được lượng tử hóa để giảm

tốc độ dữ liệu cần cho việc biểu diễn tín hiệu. Bước này biểu thị sự hao hụt thông tin và đưa độ méo mã hóa (sai số, nhiễu âm) vào tín hiệu. Để tối thiểu hóa tác động khả năng của nhiễu âm mã hóa, các kích thước bước tính tử hóa được điều khiển theo các ngưỡng mã hóa đích cho mỗi băng và khung. Theo lý tưởng, nhiễu âm mã hóa được đưa vào trong mỗi băng là thấp hơn ngưỡng (mạng che) mã hóa và do đó không có sự suy giảm về âm thanh chủ quan có thể nhận thấy được (loại bỏ sự không tương quan). Việc điều khiển này của nhiễu âm lượng tử hóa qua tần số và thời gian theo các yêu cầu tâm thính học dẫn đến hiệu quả tạo hình nhiễu âm phức tạp và đó là điều khiến cho bộ mã hóa trở thành bộ mã hóa âm thanh cảm quan.

Tiếp theo, các bộ mã hóa âm thanh hiện đại thực hiện việc mã hóa entropy (ví dụ, mã hóa Huffman, mã hóa số học) trên dữ liệu phổ được lượng tử hóa. Mã hóa entropy là bước mã hóa không tổn hao, mà còn tiết kiệm tốc độ bit.

Cuối cùng, tất cả các dữ liệu phổ được mã hóa và các tham số bổ sung liên quan (thông tin phụ, như, ví dụ, các thiết lập bộ lượng tử hóa cho mỗi băng) và được gói cùng nhau thành dòng bit, mà là phép biểu diễn được giải mã cuối cùng dự định cho việc lưu trữ tập tin hoặc việc truyền.

Sự mở rộng băng thông

Trong việc mã hóa âm thanh cảm quan trên cơ sở các giàn lọc, phân chính của tốc độ bit được tiêu thụ thường được dùng cho các hệ số phổ được lượng tử hóa. Do đó, tại tốc độ bit rất thấp, không đủ bit có thể có sẵn để biểu diễn tất cả các hệ số trong sự chính xác cần thiết để đạt được sự tái tạo nguyên vẹn về cảm quan. Do đó, các yêu cầu tốc độ bit thấp thiết lập một cách hiệu quả giới hạn đến băng thông âm thanh mà có thể thu được bằng mã hóa âm thanh cảm quan. Việc mở rộng băng thông [2] loại bỏ giới hạn cơ bản tồn tại lâu dài này. Ý tưởng chủ yếu của việc mở rộng băng thông là để bù bộ mã hóa-giải mã cảm quan giới hạn băng bằng bộ xử lý tần số cao bổ sung mà truyền và khôi phục nội dung tần số cao bị khuyết ở dạng tham số thu gọn. Nội dung tần số cao có thể được tạo ra trên cơ sở phép điều biến băng phụ đơn của tín hiệu băng cơ sở, trên cơ sở các kỹ thuật sao chép như được sử dụng trong việc sao chép băng phổ (Spectral Band Replication - SBR) [3] hoặc trên cơ sở ứng dụng các kỹ thuật dịch chuyển độ cao như, ví dụ, bộ mã tiếng nói [4].

Các hiệu quả âm thanh kỹ thuật số

Việc kéo dài thời gian hoặc các hiệu quả dịch chuyển độ cao thường thu được bằng cách áp dụng các kỹ thuật miền thời gian như các kỹ thuật chồng lấp-cộng được đồng bộ hóa (synchronized overlap-add - SOLA) hoặc kỹ thuật miền tần số (bộ mã tiếng nói). Ngoài ra, các hệ thống lai đã được đề xuất áp dụng quy trình xử lý SOLA trong các băng phụ. Các bộ mã tiếng nói và các hệ thống lai thường phải chịu thành phần lạ được gọi là sự định pha [8] mà có thể được quy cho sự hao hụt của nhất quán pha dọc. Một số tài liệu công bố đề cập đến các cải thiện về chất lượng âm của các thuật toán kéo dài thời gian bằng cách duy trì sự nhất quán pha dọc mà ở đó nó quan trọng [6][7].

Các bộ mã hóa âm thanh theo tình trạng kỹ thuật [1] thường hứa hẹn chất lượng cảm quan của các tín hiệu âm thanh bằng cách bỏ qua các đặc tính pha quan trọng để được mã hóa. Đề xuất chung về hiệu chỉnh nhất quán pha trong các bộ mã hóa âm thanh cảm quan được đề cập trong [9].

Tuy nhiên, không phải tất cả các loại sai số nhất quán pha có thể được hiệu chỉnh cùng lúc và không phải tất cả các sai số nhất quán pha đều quan trọng về cảm quan. Ví dụ, trong sự mở rộng băng thông âm thanh, không rõ ràng từ tình trạng kỹ thuật, mà các sai số liên quan đến sự nhất quán pha nên được hiệu chỉnh với ưu tiên cao nhất và sai số nào có thể còn lại chỉ được hiệu chỉnh một phần hoặc, đối với tác động cảm quan không đáng kể của chúng, hoàn toàn bị bỏ qua.

Cụ thể, do sự ứng dụng sự mở rộng băng thông âm thanh [2][3][4], sự nhất quán pha qua tần số và qua thời gian thường bị suy yếu. Kết quả là âm vô giá trị mà thể hiện độ thô thính giác và có thể chứa các âm được nhận biết bổ sung mà phân ly khỏi các đối tượng thính giác trong tín hiệu ban đầu và do đó được nhận biết như đối tượng thính giác của riêng mình hơn nữa đến tín hiệu ban đầu. Hơn nữa, âm có thể cũng có vẻ đến từ khoảng cách xa, trở nên ít "ù", và do đó gợi lên một chút sự ăn khớp người nghe [5].

Do đó, cần có phương thức cải tiến.

Tài liệu tham khảo

[1] Painter, T.; Spanias, A. Perceptual coding of digital audio, Proceedings of the IEEE, 88(4), 2000; pp. 451-513.

[2] Larsen, E.; Aarts, R. Audio Bandwidth Extension: Application of psychoacoustics, signal processing and loudspeaker design, John Wiley and Sons Ltd, 2004, Chapters 5, 6.

[3] Dietz, M.; Liljeryd, L.; Kjørting, K.; Kunz, O. Spectral Band Replication, a Novel Approach in Audio Coding, 112th AES Convention, April 2002, Preprint 5553.

[4] Nagel, F.; Disch, S.; Rettelbach, N. A Phase Vocoder Driven Bandwidth Extension Method with Novel Transient Handling for Audio Codecs, 126th AES Convention, 2009.

[5] D. Griesinger 'The Relationship between Audience Engagement and the ability to Perceive Pitch, Timbre, Azimuth and Envelopment of Multiple Sources' Tonmeister Tagung 2010.

[6] D. Dorran and R. Lawlor, "Time-scale modification of music using a synchronized subband/time domain approach," IEEE International Conference on Acoustics, Speech and Signal Processing, pp. IV 225 - IV 228, Montreal, May 2004.

[7] J. Laroche, "Frequency-domain techniques for high quality voice modification," Proceedings of the International Conference on Digital Audio Effects, pp. 328-322, 2003.

[8] Laroche, J.; Dolson, M.; , "Phase-vocoder: about this phasiness business," Applications of Signal Processing to Audio and Acoustics, 1997. 1997 IEEE ASSP Workshop on, vol., no., pp.4 pp., 19-22, Oct 1997

[9] M. Dietz, L. Liljeryd, K. Kjörting, and O. Kunz, "Spectral band replication, a novel approach in audio coding," in AES 112th Convention, (Munich, Germany), May 2002.

[10] P. Ekstrand, "Bandwidth extension of audio signals by spectral band replication," in IEEE Benelux Workshop on Model based Processing and Coding of Audio, (Leuven, Belgium), November 2002.

[11] B. C. J. Moore and B. R. Glasberg, "Suggested formulae for calculating auditory-filter bandwidths and excitation patterns," J. Acoust. Soc. Am., vol. 74, pp. 750-753, September 1983.

[12] T. M. Shackleton and R. P. Carlyon, "The role of resolved and unresolved harmonics in pitch perception and frequency modulation discrimination," J. Acoust. Soc. Am., vol. 95, pp. 3529-3540, June 1994.

[13] M.-V. Laitinen, S. Disch, and V. Pulkki, "Sensitivity of human hearing to changes in phase spectrum," J. Audio Eng. Soc., vol. 61, pp. 860-877, November 2013.

[14] A. Klapuri, "Multiple fundamental frequency estimation based on harmonicity and spectral smoothness," IEEE Transactions on Speech and Audio Processing, vol. 11, November 2003.

Bản chất kỹ thuật của sáng chế

Mục đích của sáng chế là đề xuất khái niệm được cải thiện để giải mã hoặc mã hóa tín hiệu âm thanh. Mục đích này đạt được bởi đối tượng của các điểm yêu cầu bảo hộ độc lập kèm theo.

Các phương án cụ thể được định rõ trong các yêu cầu bảo hộ phụ thuộc.

Sáng chế dựa trên việc tìm ra rằng pha của tín hiệu âm thanh có thể được hiệu chỉnh theo pha đích được tính toán bởi bộ xử lý âm thanh hoặc bộ giải mã. Pha đích có thể được xem như phép biểu diễn của pha của tín hiệu âm thanh chưa được xử lý. Do đó, pha của tín hiệu âm thanh được xử lý được điều chỉnh để tốt hơn là vừa với pha của tín hiệu âm thanh chưa được xử lý. Ví dụ, có phép biểu diễn thời gian tần số của tín hiệu âm thanh, pha của tín hiệu âm thanh có thể được điều chỉnh cho các khung thời gian tiếp theo trong băng phụ, hoặc pha có thể được điều chỉnh trong khung thời gian cho các băng phụ tần số tiếp theo. Do đó, các tác giả sáng chế đã tìm ra rằng bộ tính toán phát hiện và lựa chọn một cách tự động phương pháp chính xác thích hợp nhất. Các kết quả được mô tả có thể được thực hiện theo các phương án khác nhau hoặc được thực hiện đồng thời trong bộ giải mã và/hoặc bộ mã hóa.

Các phương án thể hiện bộ xử lý âm thanh để xử lý tín hiệu âm thanh bao gồm

bộ tính toán số đo pha tín hiệu âm thanh được tạo cấu hình để tính toán số đo pha của tín hiệu âm thanh cho khung thời gian. Hơn nữa, tín hiệu âm thanh bao gồm bộ xác định số đo pha đích để xác định số đo pha đích cho khung thời gian đã nêu và bộ hiệu chỉnh pha được tạo cấu hình để hiệu chỉnh các pha của tín hiệu âm thanh cho khung thời gian sử dụng số đo pha được tính toán và số đo pha đích để thu được tín hiệu âm thanh được xử lý.

Theo các phương án khác, tín hiệu âm thanh có thể bao gồm nhiều tín hiệu băng phụ cho khung thời gian. Bộ xác định số đo pha đích được tạo cấu hình để xác định số đo pha đích thứ nhất cho tín hiệu băng phụ thứ nhất và số đo pha đích thứ hai cho tín hiệu băng phụ thứ hai. Hơn nữa, bộ tính toán số đo pha tín hiệu âm thanh xác định số đo pha thứ nhất cho tín hiệu băng phụ thứ nhất và số đo pha thứ hai cho tín hiệu băng phụ thứ hai. Bộ hiệu chỉnh pha được tạo cấu hình để hiệu chỉnh pha thứ nhất của tín hiệu băng phụ thứ nhất sử dụng số đo pha thứ nhất của tín hiệu âm thanh và số đo pha đích thứ nhất và để hiệu chỉnh pha thứ hai của tín hiệu băng phụ thứ hai sử dụng số đo pha thứ hai của tín hiệu âm thanh và số đo pha đích thứ hai. Do đó, bộ xử lý âm thanh có thể bao gồm bộ tổng hợp tín hiệu âm thanh để tổng hợp tín hiệu âm thanh được hiệu chỉnh sử dụng tín hiệu băng phụ thứ nhất được hiệu chỉnh và tín hiệu băng phụ thứ hai được hiệu chỉnh.

Theo sáng chế, bộ xử lý âm thanh được tạo cấu hình để hiệu chỉnh pha của tín hiệu âm thanh theo hướng ngang, tức là, hiệu chỉnh qua thời gian. Do đó, tín hiệu âm thanh có thể được chia nhỏ thành tập hợp các khung thời gian, trong đó pha của mỗi khung thời gian có thể được điều chỉnh theo pha đích. Pha đích có thể là phép biểu diễn tín hiệu âm thanh ban đầu, trong đó bộ xử lý âm thanh có thể là phần của bộ giải mã để giải mã tín hiệu âm thanh mà là phép biểu diễn được mã hóa của tín hiệu âm thanh ban đầu. Một cách tùy chọn, sự hiệu chỉnh pha ngang có thể được áp dụng riêng cho số lượng băng phụ của tín hiệu âm thanh, nếu tín hiệu âm thanh có sẵn trong phép biểu diễn thời gian - tần số. Sự hiệu chỉnh pha của tín hiệu âm thanh có thể được thực hiện bằng cách trừ độ lệch của đạo hàm pha theo thời gian của pha đích và pha của tín hiệu âm thanh từ pha của tín hiệu âm thanh.

Do đó, vì việc đạo hàm pha theo thời gian là tần số ($\frac{d\varphi}{dt} = f$, với φ là pha), sự

hiệu chỉnh pha được mô tả thực hiện điều chỉnh tần số cho mỗi băng phụ của tín hiệu âm thanh. Nói cách khác, sự chênh lệch của mỗi băng phụ của tín hiệu âm thanh so với tần số đích có thể được giảm để thu được chất lượng tốt hơn cho tín hiệu âm thanh.

Để xác định pha đích, bộ xác định pha đích được tạo cấu hình để thu được ước tính tần số cơ bản cho khung thời gian hiện thời và để tính toán ước tính tần số cho mỗi băng phụ trong số nhiều băng phụ của khung thời gian sử dụng ước tính tần số cơ bản cho khung thời gian. Ước tính tần số có thể được chuyển đổi thành đạo hàm pha theo thời gian sử dụng tổng số các băng phụ và tần số lấy mẫu của tín hiệu âm thanh. Trong phương án khác, bộ xử lý âm thanh bao gồm bộ xác định số đo pha đích để xác định số đo pha đích cho tín hiệu âm thanh trong khung thời gian, bộ tính toán sai số pha để tính toán sai số pha sử dụng pha của tín hiệu âm thanh và khung thời gian của số đo pha đích, và bộ hiệu chỉnh pha được tạo cấu hình để hiệu chỉnh pha của tín hiệu âm thanh và khung thời gian sử dụng sai số pha.

Theo các phương án khác, tín hiệu âm thanh có sẵn trong phép biểu diễn tần số, trong đó tín hiệu âm thanh bao gồm nhiều băng phụ cho khung thời gian. Bộ xác định số đo pha đích xác định số đo pha đích thứ nhất cho tín hiệu băng phụ thứ nhất và số đo pha đích thứ hai cho tín hiệu băng phụ thứ hai. Hơn nữa, bộ tính toán sai số pha tạo ra vectơ của các sai số pha, trong đó phần tử thứ nhất của vectơ đề cập đến độ lệch thứ nhất của pha của tín hiệu băng phụ thứ nhất và số đo pha đích thứ nhất và trong đó phần tử thứ hai của vectơ đề cập đến độ lệch thứ hai của pha của tín hiệu băng phụ thứ hai và số đo pha đích thứ hai. Ngoài ra, bộ xử lý âm thanh của phương án này bao gồm bộ tổng hợp tín hiệu âm thanh để tổng hợp tín hiệu âm thanh được hiệu chỉnh sử dụng tín hiệu băng phụ thứ nhất được hiệu chỉnh và tín hiệu băng phụ thứ hai được hiệu chỉnh. Sự hiệu chỉnh pha này tạo ra các giá trị pha được hiệu chỉnh trung bình.

Ngoài ra, nhiều băng phụ được nhóm thành băng cơ sở và tập hợp các phần ráp nối tần số, trong đó băng cơ sở bao gồm một băng phụ của tín hiệu âm thanh và tập hợp các phần ráp nối tần số bao gồm ít nhất một băng phụ của băng cơ sở tại tần số cao hơn tần số của ít nhất một băng phụ trong băng cơ sở.

Các phương án khác thể hiện bộ tính toán sai số pha được tạo cấu hình để tính toán trung bình các phần tử của vectơ của các sai số pha đề cập đến phần ráp nối thứ

nhất của số lượng thứ hai của các phần ráp nối tần số để thu được sai số pha trung bình. Bộ hiệu chỉnh pha được tạo cấu hình để hiệu chỉnh pha của tín hiệu băng phụ trong các phần ráp nối tần số thứ nhất và tiếp theo của tập hợp các phần ráp nối tần số của tín hiệu phần ráp nối sử dụng sai số pha trung bình được gán trọng số, trong đó sai số pha trung bình được chia theo chỉ số của phần ráp nối tần số để thu được tín hiệu phần ráp nối cải biên. Sự hiệu chỉnh pha này cung cấp chất lượng tốt tại các tần số giao chéo, mà là các tần số biên giữa hai phần ráp nối tần số tiếp theo.

Theo phương án khác, hai phương án được mô tả trước đó có thể được tổ hợp để thu được tín hiệu âm thanh được hiệu chỉnh bao gồm các giá trị được hiệu chỉnh pha trung bình tốt và tại các tần số giao chéo. Do đó, bộ tính toán đạo hàm pha tín hiệu âm thanh được tạo cấu hình để tính toán trung bình của các đạo hàm pha theo tần số cho băng cơ sở. Bộ hiệu chỉnh pha tính toán tín hiệu phần ráp nối cải biên khác với phần ráp nối tần số thứ nhất được tối ưu hóa bằng cách cộng trung bình của các đạo hàm pha theo tần số được gán trọng số bởi chỉ số băng phụ hiện thời vào pha của tín hiệu băng phụ với chỉ số băng phụ cao nhất trong băng cơ sở của tín hiệu âm thanh. Hơn nữa, bộ hiệu chỉnh pha có thể được tạo cấu hình để tính toán trung bình được gán trọng số của tín hiệu phần ráp nối cải biên và tín hiệu phần ráp nối cải biên khác để thu được tín hiệu phần ráp nối cải biên được tổ hợp và để cập nhật một cách đệ quy, trên cơ sở các phần ráp nối tần số, tín hiệu phần ráp nối cải biên được tổ hợp bằng cách cộng trung bình các đạo hàm pha theo tần số, được gán trọng số bởi chỉ số băng phụ của băng phụ hiện thời, vào pha của tín hiệu băng phụ với chỉ số băng phụ cao nhất trong phần ráp nối tần số trước đó của tín hiệu phần ráp nối cải biên được tổ hợp.

Để xác định pha đích, bộ xác định số đo pha đích có thể bao gồm bộ trích xuất dòng dữ liệu được tạo cấu hình để trích xuất vị trí đỉnh và tần số cơ bản của các vị trí đỉnh trong khung thời gian hiện thời của tín hiệu âm thanh từ dòng dữ liệu. Ngoài ra, bộ xác định số đo pha đích có thể bao gồm bộ phân tích tín hiệu âm thanh được tạo cấu hình để phân tích khung thời gian hiện thời để tính toán vị trí đỉnh và tần số cơ bản của các vị trí đỉnh trong khung thời gian hiện thời. Hơn nữa, bộ xác định số đo pha đích bao gồm bộ tạo phổ đích để ước tính các vị trí đỉnh khác trong khung thời gian hiện thời sử dụng vị trí đỉnh và tần số cơ bản của các vị trí đỉnh. Cụ thể, bộ tạo phổ đích có thể bao gồm bộ phát hiện đỉnh để tạo ra chuỗi xung của thời gian, bộ tạo tín

hiệu để điều chỉnh tần số của chuỗi xung theo tần số cơ bản của các vị trí đỉnh, bộ tạo vị trí xung để điều chỉnh pha của chuỗi xung theo vị trí, và bộ phân tích phổ để tạo ra phổ pha của chuỗi xung được điều chỉnh, trong đó phổ pha của tín hiệu miền thời gian là số đo pha đích. Phương án được mô tả của bộ xác định số đo pha đích là có lợi để tạo ra phổ đích cho tín hiệu âm thanh có dạng sóng với các đỉnh.

Các phương án của bộ xử lý âm thanh thứ hai mô tả sự hiệu chỉnh pha dọc. Sự hiệu chỉnh pha dọc điều chỉnh pha của tín hiệu âm thanh trong một khung thời gian qua tất cả các băng phụ. Sự điều chỉnh pha của tín hiệu âm thanh, được áp dụng độc lập đối với mỗi băng, các kết quả, sau khi tổng hợp các băng phụ của tín hiệu âm thanh, trong dạng sóng của tín hiệu âm thanh khác với tín hiệu âm thanh không được hiệu chỉnh. Do đó, ví dụ, là khả thi để phục hồi hình dạng đỉnh bị nhòe hoặc sự chuyển tiếp.

Theo phương án khác, bộ tính toán được thể hiện để xác định dữ liệu hiệu chỉnh pha cho tín hiệu âm thanh với bộ xác định độ biến thiên để xác định độ biến thiên của pha của tín hiệu âm thanh trong chế độ biến thiên thứ nhất và thứ hai, bộ so sánh độ biến thiên để so sánh độ biến thiên thứ nhất được xác định sử dụng chế độ biến thiên pha và độ biến thiên thứ hai được xác định sử dụng chế độ biến thiên thứ hai, và bộ tính toán dữ liệu hiệu chỉnh để tính toán sự hiệu chỉnh pha phù hợp với chế độ biến thiên thứ nhất hoặc chế độ biến thiên thứ hai trên cơ sở kết quả của sự so sánh.

Phương án khác thể hiện bộ xác định độ biến thiên để xác định số đo độ lệch tiêu chuẩn của đạo hàm pha theo thời gian (phase derivative over time - PDT) cho nhiều khung thời gian của tín hiệu âm thanh như độ biến thiên của pha trong chế độ biến thiên thứ nhất hoặc số đo độ lệch tiêu chuẩn của pha đạo hàm pha theo tần số (phase derivative over frequency - PDF) cho nhiều băng phụ như độ biến thiên của pha trong chế độ biến thiên thứ hai. Bộ so sánh độ biến thiên so sánh số đo của đạo hàm pha theo thời gian như chế độ biến thiên thứ nhất và số đo của đạo hàm pha theo tần số như chế độ biến thiên thứ hai cho các khung thời gian của tín hiệu âm thanh. Theo phương án khác, bộ xác định độ biến thiên được tạo cấu hình để xác định độ biến thiên của pha của tín hiệu âm thanh trong chế độ biến thiên thứ ba, trong đó chế độ biến thiên thứ ba là chế độ phát hiện chuyển tiếp. Do đó, bộ so sánh độ biến thiên so sánh

ba chế độ biến thiên và bộ tính toán dữ liệu hiệu chỉnh tính toán hiệu chỉnh pha phù hợp với chế độ biến thiên thứ nhất, chế độ biến thiên thứ hai, hoặc chế độ biến thiên thứ ba dựa trên kết quả so sánh.

Các quy tắc quyết định của bộ tính toán dữ liệu hiệu chỉnh có thể được mô tả như sau. Nếu sự chuyển tiếp được phát hiện, pha được hiệu chỉnh theo sự hiệu chỉnh pha cho các chuyển tiếp để khôi phục hình dạng của sự chuyển tiếp. Cách khác, nếu độ biến thiên thứ nhất nhỏ hơn hoặc bằng độ biến thiên thứ hai, sự hiệu chỉnh pha của chế độ biến thiên thứ nhất được áp dụng hoặc, nếu độ biến thiên thứ hai lớn hơn độ biến thiên thứ nhất, sự hiệu chỉnh pha phù hợp với chế độ biến thiên thứ hai được áp dụng. Nếu sự vắng mặt của sự chuyển tiếp được phát hiện và nếu cả độ biến thiên thứ nhất và thứ hai vượt quá giá trị ngưỡng, không có chế độ nào trong số các chế độ hiệu chỉnh pha được áp dụng.

Bộ tính toán có thể được tạo cấu hình để phân tích tín hiệu âm thanh, ví dụ, trong giai đoạn mã hóa âm thanh, để xác định chế độ hiệu chỉnh pha tốt nhất và để tính toán các tham số liên quan cho chế độ hiệu chỉnh pha được xác định. Trong giai đoạn giải mã, các tham số có thể được sử dụng để thu được tín hiệu âm thanh được giải mã mà có chất lượng tốt hơn so với các tín hiệu âm thanh được giải mã sử dụng các bộ mã hóa-giải mã theo tình trạng kỹ thuật. Phải lưu ý rằng bộ tính toán phát hiện một cách độc lập chế độ hiệu chỉnh đúng cho mỗi khung thời gian của tín hiệu âm thanh.

Các phương án thể hiện bộ giải mã để giải mã tín hiệu âm thanh với bộ tạo phổ đích thứ nhất để tạo ra phổ đích cho khung thời gian thứ nhất của tín hiệu thứ hai của tín hiệu âm thanh sử dụng dữ liệu hiệu chỉnh thứ nhất và bộ hiệu chỉnh pha thứ nhất để hiệu chỉnh pha của tín hiệu băng phụ trong khung thời gian thứ nhất của tín hiệu âm thanh được xác định với thuật toán hiệu chỉnh pha, trong đó sự hiệu chỉnh được thực hiện bằng cách giảm chênh lệch giữa số đo của tín hiệu băng phụ trong khung thời gian thứ nhất của tín hiệu âm thanh và phổ đích. Ngoài ra, bộ giải mã bao gồm bộ tính toán tín hiệu băng phụ âm thanh để tính toán tín hiệu băng phụ âm thanh cho khung thời gian thứ nhất sử dụng pha được hiệu chỉnh cho khung thời gian và để tính toán tín hiệu băng phụ âm thanh cho khung thời gian thứ hai khác với khung thời gian thứ nhất sử dụng số đo của tín hiệu băng phụ trong khung thời gian thứ hai hoặc sử dụng sự

tính toán pha được hiệu chỉnh phù hợp với thuật toán được hiệu chỉnh pha khác khác với thuật toán hiệu chỉnh pha.

Theo các phương án khác, bộ giải mã bao gồm bộ tạo phổ đích thứ hai và thứ ba tương đương với việc tạo phổ đích thứ nhất và bộ hiệu chỉnh pha thứ hai và thứ ba tương đương với bộ hiệu chỉnh pha thứ nhất. Do đó, bộ hiệu chỉnh pha thứ nhất có thể thực hiện sự hiệu chỉnh pha ngang, bộ hiệu chỉnh pha thứ hai có thể thực hiện hiệu chỉnh pha dọc, và bộ hiệu chỉnh pha thứ ba có thể thực hiện các chuyển tiếp hiệu chỉnh pha. Theo phương án khác, bộ giải mã bao gồm bộ giải mã lỗi được tạo cấu hình để giải mã tín hiệu âm thanh trong khung thời gian với số lượng giảm các băng phụ so với tín hiệu âm thanh. Hơn nữa, bộ giải mã có thể bao gồm bộ ráp nối để ráp nối tập hợp các băng phụ của tín hiệu âm thanh được giải mã lỗi với số lượng băng phụ giảm, trong đó tập hợp các băng phụ tạo ra phần ráp nối thứ nhất, với các băng phụ khác trong khung thời gian, liền kề với số lượng các băng phụ giảm, để thu được tín hiệu âm thanh với số lượng các băng phụ đều đều. Hơn nữa, bộ giải mã có thể bao gồm bộ xử lý cường độ để xử lý các giá trị cường độ của tín hiệu băng phụ âm thanh trong khung thời gian và bộ tổng hợp tín hiệu âm thanh để tổng hợp các tín hiệu băng phụ âm thanh hoặc cường độ của các tín hiệu băng phụ âm thanh được xử lý để thu được tín hiệu âm thanh được giải mã được tổng hợp. Phương án này có thể thiết lập bộ giải mã cho việc mở rộng băng thông bao gồm sự hiệu chỉnh pha của tín hiệu âm thanh được giải mã.

Theo đó, bộ mã hóa để mã hóa tín hiệu âm thanh bao gồm bộ xác định pha để xác định pha của tín hiệu âm thanh, bộ tính toán để xác định dữ liệu hiệu chỉnh pha cho tín hiệu âm thanh trên cơ sở pha được xác định của tín hiệu âm thanh, bộ mã hóa lỗi được tạo cấu hình để mã hóa lỗi của tín hiệu âm thanh để thu được tín hiệu âm thanh được mã hóa lỗi có số lượng các băng phụ giảm so với tín hiệu âm thanh, bộ trích xuất tham số được tạo cấu hình để trích xuất các tham số của tín hiệu âm thanh để thu được phép biểu diễn tham số độ phân giải thấp cho tập hợp thứ hai gồm các băng phụ không được bao gồm trong tín hiệu âm thanh được mã hóa lỗi, và bộ tạo tín hiệu âm thanh để tạo ra tín hiệu âm thanh bao gồm các tham số, tín hiệu âm thanh được mã hóa lỗi, và dữ liệu hiệu chỉnh pha có thể tạo ra bộ mã hóa cho sự mở rộng băng thông.

Tất cả các phương án được mô tả trước đó có thể được nhìn toàn bộ hoặc trong tổ hợp, ví dụ trong bộ mã hóa và/hoặc bộ giải mã cho sự mở rộng băng thông với sự hiệu chỉnh pha của tín hiệu âm thanh được giải mã. Cách khác, cũng khả thi để xem tất cả các phương án được mô tả độc lập mà không liên quan đến nhau.

Mô tả vắn tắt các hình vẽ

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

Fig.1a thể hiện phổ cường độ của tín hiệu đàn viôlông trong phép biểu diễn thời gian tần số;

Fig.1b thể hiện phổ pha tương ứng với phổ cường độ trên Fig.1a;

Fig.1c thể hiện phổ cường độ của tín hiệu kèn trombon trong miền QMF trong phép biểu diễn thời gian tần số;

Fig.1d thể hiện phổ pha tương ứng với phổ cường độ trên Fig.1c;

Fig.2 thể hiện sơ đồ thời gian tần số bao gồm các ô thời gian tần số (ví dụ, các ngăn QMF, các ngăn giàn bộ lọc gương cầu phương), được định rõ bởi khung thời gian và băng phụ;

Fig.3a thể hiện sơ đồ tần số minh họa của tín hiệu âm thanh, trong đó cường độ của tần số được biểu thị trên mười băng phụ khác nhau;

Fig.3b thể hiện phép biểu diễn tần số minh họa của tín hiệu âm thanh sau khi thu nhận, ví dụ, trong suốt quy trình giải mã tại bước trung gian;

Fig.3c thể hiện phép biểu diễn tần số minh họa của tín hiệu âm thanh được khôi phục $Z(k, n)$;

Fig.4a thể hiện phổ cường độ của tín hiệu đàn viôlông trong miền QMF sử dụng SBR sao chép trực tiếp trong phép biểu diễn thời gian tần số;

Fig.4b thể hiện phổ pha tương ứng với phổ cường độ trên Fig.4a;

Fig.4c thể hiện phổ cường độ của tín hiệu kèn trombon trong miền QMF sử dụng SBR sao chép trực tiếp trong phép biểu diễn thời gian tần số;

Fig.4d thể hiện phổ pha tương ứng với phổ cường độ trên Fig.4c;

Fig.5 thể hiện phép biểu diễn miền thời gian của ngân QMF đơn với các giá trị pha khác nhau;

Fig.6 thể hiện phép biểu diễn miền thời gian và miền tần số của tín hiệu, mà có một băng khác 0 và pha thay đổi với giá trị cố định, $\pi/4$ (phía trên) và $3\pi/4$ (phía dưới);

Fig.7 thể hiện phép biểu diễn miền thời gian và miền tần số của tín hiệu, mà có một băng khác 0 và pha thay đổi một cách ngẫu nhiên;

Fig.8 minh họa hiệu quả được mô tả liên quan đến Fig.6 trong phép biểu diễn thời gian tần số của bốn khung thời gian và bốn băng phụ tần số, trong đó chỉ ba băng phụ bao gồm tần số khác 0;

Fig.9 thể hiện phép biểu diễn miền thời gian và miền tần số của tín hiệu, mà có một khung thời gian khác 0 và pha thay đổi với giá trị cố định, $\pi/4$ (phía trên) và $3\pi/4$ (phía dưới);

Fig.10 thể hiện phép biểu diễn miền thời gian và miền tần số của tín hiệu, mà có một khung thời gian khác 0 và pha thay đổi một cách ngẫu nhiên;

Fig.11 thể hiện sơ đồ thời gian tần số tương tự với sơ đồ thời gian tần số được thể hiện trên Fig.8, trong đó chỉ khung thứ ba bao gồm tần số khác 0;

Fig.12a thể hiện đạo hàm pha theo thời gian của tín hiệu đàn viôlông trong miền QMF trong phép biểu diễn thời gian - tần số;

Fig.12b thể hiện tần số đạo hàm pha tương ứng với đạo hàm pha theo thời gian được thể hiện trên Fig.12a;

Fig.12c thể hiện đạo hàm pha theo thời gian của tín hiệu kèn trombon trong miền QMF trong phép biểu diễn thời gian - tần số;

Fig.12d thể hiện đạo hàm pha theo tần số của đạo hàm pha theo thời gian tương ứng được thể hiện trên Fig.12c;

Fig.13a thể hiện đạo hàm pha theo thời gian của tín hiệu đàn viôlông trong miền QMF sử dụng SBR sao chép trực tiếp trong phép biểu diễn thời gian tần số;

Fig.13b thể hiện đạo hàm pha theo tần số tương ứng với đạo hàm pha theo thời gian được thể hiện trên Fig.13a;

Fig.13c thể hiện đạo hàm pha theo thời gian của tín hiệu kèn trombon trong miền QMF sử dụng SBR sao chép trực tiếp trong phép biểu diễn thời gian tần số;

Fig.13d thể hiện đạo hàm pha theo tần số tương ứng với đạo hàm pha theo thời gian được thể hiện trên Fig.13c;

Fig.14a thể hiện dưới dạng giản đồ bốn pha của, ví dụ, các khung thời gian tiếp theo hoặc các băng phụ tần số, ở hình tròn đơn vị;

Fig.14b thể hiện các pha được minh họa trong Fig.14a sau khi xử lý SBR và, dưới dạng đường nét đứt, các pha được hiệu chỉnh;

Fig.15 thể hiện sơ đồ khối dạng giản lược của bộ xử lý âm thanh 50;

Fig.16 thể hiện bộ xử lý âm thanh trong sơ đồ khối dạng giản lược theo phương án khác;

Fig.17 thể hiện sai số được làm nhẵn trong PDT của tín hiệu đàn viôlông trong miền QMF sử dụng SBR sao chép trực tiếp trong phép biểu diễn thời gian tần số;

Fig.18a thể hiện sai số trong PDT của tín hiệu đàn viôlông trong miền QMF cho SBR được hiệu chỉnh trong phép biểu diễn thời gian tần số;

Fig.18b thể hiện đạo hàm pha theo thời gian tương ứng với sai số được thể hiện trên Fig.18a;

Fig.19 thể hiện sơ đồ khối dạng giản lược của bộ giải mã;

Fig.20 thể hiện sơ đồ khối dạng giản lược của bộ mã hóa;

Fig.21 thể hiện sơ đồ khối dạng giản lược của dòng dữ liệu mà có thể là tín hiệu âm thanh;

Fig.22 thể hiện dòng dữ liệu trên Fig.21 theo phương án khác;

Fig.23 thể hiện sơ đồ khối dạng giản lược của phương pháp xử lý tín hiệu âm thanh;

Fig.24 thể hiện sơ đồ khối dạng giản lược của phương pháp giải mã tín hiệu âm

thanh;

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

Fig.26 thể hiện sơ đồ khối dạng giản lược của bộ xử lý âm thanh theo phương án khác;

Fig.27 thể hiện sơ đồ khối dạng giản lược của bộ xử lý âm thanh theo phương án được ưu tiên;

Fig.28a thể hiện sơ đồ khối dạng giản lược của bộ hiệu chỉnh pha trong bộ xử lý âm thanh minh họa dòng tín hiệu cụ thể hơn;

Fig.28b thể hiện các bước hiệu chỉnh pha từ quan điểm khác so với các hình vẽ từ Fig.26 đến Fig.28a;

Fig.29 thể hiện sơ đồ khối dạng giản lược của bộ xác định số đo pha đích trong bộ xử lý âm thanh minh họa bộ xác định số đo pha đích cụ thể hơn;

Fig.30 thể hiện sơ đồ khối dạng giản lược của bộ tạo phổ đích trong bộ xử lý âm thanh minh họa bộ tạo phổ đích cụ thể hơn;

Fig.31 thể hiện sơ đồ khối dạng giản lược của bộ giải mã;

Fig.32 thể hiện sơ đồ khối dạng giản lược của bộ mã hóa;

Fig.33 thể hiện sơ đồ khối dạng giản lược của dòng dữ liệu mà có thể là tín hiệu âm thanh;

Fig.34 thể hiện sơ đồ khối dạng giản lược của phương pháp xử lý tín hiệu âm thanh;

Fig.35 thể hiện sơ đồ khối dạng giản lược của phương pháp giải mã tín hiệu âm thanh;

Fig.36 thể hiện sơ đồ khối dạng giản lược của phương pháp giải mã tín hiệu âm thanh;

Fig.37 thể hiện sai số trong phổ pha của tín hiệu kèn trombon trong miền QMF sử dụng SBR sao chép trực tiếp trong phép biểu diễn thời gian tần số;

Fig.38a thể hiện sai số trong phổ pha của tín hiệu kèn trombon trong miền QMF sử dụng SBR được hiệu chỉnh trong phép biểu diễn thời gian tần số;

Fig.38b thể hiện đạo hàm pha theo tần số tương ứng với sai số được thể hiện trên Fig.38a;

Fig.39 thể hiện sơ đồ khối dạng giản lược của bộ tính toán;

Fig.40 thể hiện sơ đồ khối dạng giản lược của bộ tính toán minh họa dòng tín hiệu trong bộ xác định độ biến thiên cụ thể hơn;

Fig.41 thể hiện sơ đồ khối dạng giản lược của bộ tính toán theo phương án khác;

Fig.42 thể hiện sơ đồ khối dạng giản lược của phương pháp xác định dữ liệu hiệu chỉnh pha cho tín hiệu âm thanh;

Fig.43a thể hiện đạo hàm tiêu chuẩn của đạo hàm pha theo thời gian của tín hiệu đàn viôlông trong miền QMF trong phép biểu diễn thời gian - tần số;

Fig.43b thể hiện đạo hàm tiêu chuẩn của đạo hàm pha theo tần số tương ứng với độ lệch tiêu chuẩn của đạo hàm pha theo thời gian được thể hiện trên Fig.43a;

Fig.43c thể hiện đạo hàm tiêu chuẩn của đạo hàm pha theo thời gian của tín hiệu kèn trombon trong miền QMF trong phép biểu diễn thời gian - tần số;

Fig.43d thể hiện đạo hàm tiêu chuẩn của đạo hàm pha theo tần số tương ứng với độ lệch tiêu chuẩn của đạo hàm pha theo thời gian được thể hiện trên Fig.43c;

Fig.44a thể hiện cường độ của tín hiệu đàn viôlông + tiếng vĩ tay trong miền QMF trong phép biểu diễn thời gian-tần số;

Fig.44b thể hiện phổ pha tương ứng với phổ cường độ được thể hiện trên Fig.44a;

Fig.45a thể hiện đạo hàm pha theo thời gian của tín hiệu đàn viôlông + tiếng vĩ tay trong miền QMF trong phép biểu diễn thời gian - tần số;

Fig.45b thể hiện đạo hàm pha theo tần số tương ứng với đạo hàm pha theo thời gian được thể hiện trên Fig.45a;

Fig.46a thể hiện đạo hàm pha theo thời gian của tín hiệu đàn viôlông + tiếng vỗ tay trong miền QMF sử dụng SBR được hiệu chỉnh trong phép biểu diễn thời gian - tần số;

Fig.46b thể hiện đạo hàm pha theo tần số tương ứng với đạo hàm pha theo thời gian được thể hiện trên Fig.46a;

Fig.47 thể hiện các tần số của băng QMF trong phép biểu diễn thời gian tần số;

Fig.48a thể hiện các tần số của SBR sao chép trực tiếp các băng QMF so với các tần số ban đầu được thể hiện trong phép biểu diễn thời gian - tần số;

Fig.48b thể hiện các tần số của băng QMF sử dụng SBR được hiệu chỉnh so với các tần số ban đầu trong phép biểu diễn thời gian - tần số;

Fig.49 thể hiện các tần số được ước tính của các sóng hài so với các tần số của các băng QMF của tín hiệu ban đầu trong phép biểu diễn thời gian - tần số;

Fig.50a thể hiện sai số trong đạo hàm pha theo thời gian của tín hiệu đàn viôlông trong miền QMF sử dụng SBR được hiệu chỉnh với dữ liệu hiệu chỉnh được nén trong phép biểu diễn thời gian tần số;

Fig.50b thể hiện đạo hàm pha theo thời gian tương ứng với sai số của đạo hàm pha theo thời gian được thể hiện trên Fig.50a;

Fig.51a thể hiện dạng sóng của tín hiệu kèn trombon trong sơ đồ thời gian;

Fig.51b biểu diễn tín hiệu miền thời gian tương ứng với tín hiệu trombon trên Fig.51a mà chỉ chứa các đỉnh được ước tính; trong đó các vị trí của các đỉnh đã thu được sử dụng siêu dữ liệu được truyền;

Fig.52a thể hiện sai số trong phổ pha của tín hiệu kèn trombon trong miền QMF sử dụng SBR được hiệu chỉnh với dữ liệu hiệu chỉnh được nén trong phép biểu diễn thời gian tần số;

Fig.52b thể hiện đạo hàm pha theo tần số tương ứng với sai số trong phổ pha được thể hiện trên Fig.52a;

Fig.53 thể hiện sơ đồ khối dạng giản lược của bộ giải mã;

Fig.54 thể hiện sơ đồ khối dạng giản lược theo phương án khác được ưu tiên;

Fig.55 thể hiện sơ đồ khối dạng giản lược của bộ giải mã theo phương án khác;

Fig.56 thể hiện sơ đồ khối dạng giản lược của bộ mã hóa;

Fig.57 thể hiện sơ đồ khối của bộ tính toán mà có thể được sử dụng trong bộ mã hóa được thể hiện trên Fig.56;

Fig.58 thể hiện sơ đồ khối dạng giản lược của phương pháp giải mã tín hiệu âm thanh; và

Fig.59 thể hiện sơ đồ khối dạng giản lược của phương pháp mã hóa tín hiệu âm thanh.

Mô tả chi tiết sáng chế

Sau đây, các phương án của sáng chế sẽ được mô tả chi tiết hơn. Các phần tử được thể hiện trên các hình vẽ tương ứng có chức năng giống nhau hoặc tương tự sẽ kết hợp với các kí hiệu tham chiếu giống nhau.

Các phương án của sáng chế sẽ được mô tả đối với việc xử lý tín hiệu cụ thể. Do đó, các hình vẽ từ Fig.1 đến Fig.14 mô tả việc xử lý tín hiệu được áp dụng cho tín hiệu âm thanh. Thậm chí các phương án được mô tả đối với việc xử lý tín hiệu đặc biệt này, sáng chế không chỉ giới hạn ở việc xử lý này và cũng có thể còn được áp dụng cho nhiều sơ đồ xử lý khác. Hơn nữa, các hình vẽ từ Fig.15 đến Fig.25 thể hiện các phương án của bộ xử lý âm thanh mà có thể được sử dụng cho sự hiệu chỉnh pha ngang của tín hiệu âm thanh. Các hình vẽ từ Fig.26 đến Fig.38 thể hiện các phương án của bộ xử lý âm thanh mà có thể được sử dụng cho sự hiệu chỉnh pha dọc của tín hiệu âm thanh. Hơn nữa, các hình vẽ từ Fig.39 đến Fig.52 thể hiện các phương án của bộ tính toán để xác định dữ liệu hiệu chỉnh pha cho tín hiệu âm thanh. Bộ tính toán có thể phân tích tín hiệu âm thanh và xác định bộ xử lý âm thanh nào trong số các bộ xử lý âm thanh được đề cập trước đó được áp dụng hoặc, nếu không có bộ xử lý âm thanh nào phù hợp cho tín hiệu âm thanh, để không áp dụng bộ nào trong số các bộ xử lý âm thanh cho tín hiệu âm thanh. Các hình vẽ từ Fig.53 đến 59 thể hiện các phương án của bộ giải mã và bộ mã hóa mà có thể bao gồm bộ xử lý thứ hai và bộ tính toán.

1. Giới thiệu

Việc mã hóa âm thanh cảm quan phát triển nhanh chóng như dòng chính cho

phép công nghệ kỹ thuật số với tất cả các loại ứng dụng mà cung cấp âm thanh và đa phương tiện đến người dùng sử dụng sự truyền hoặc lưu trữ các kênh với dung lượng được giới hạn. Các bộ mã hóa-giải mã âm thanh cảm quan hiện đại được yêu cầu truyền tải chất lượng âm thanh thỏa đáng với tốc độ bit ngày càng thấp. Kết quả là, nó phải chịu các thành phần lạ mã hóa nhất định mà hầu như có thể chịu được bởi số đông người nghe. Mở rộng băng thông (Bandwidth Extension - BWE) âm thanh là kỹ thuật để mở rộng một cách nhân tạo phạm vi tần số của bộ mã hóa âm thanh bằng sự tịnh tiến hoặc sự chuyển vị phổ của các phân tín hiệu băng thấp được truyền vào băng cao đến mức đưa vào các thành phần lạ nhất định.

Các tác giả sáng chế tìm ra rằng một số thành phần lạ này liên quan đến sự thay đổi đạo hàm pha trong băng cao được mở rộng một cách nhân tạo. Một trong số các thành phần lạ này là sự thay đổi đạo hàm pha theo tần số (cũng xem sự nhất quán pha "dọc") [8]. Sự bảo toàn của đạo hàm pha nêu trên quan trọng về cảm quan đối với các tín hiệu âm có chuỗi xung như dạng sóng miền thời gian và tần số cơ bản hơi thấp. Các thành phần lạ liên quan đến sự thay đổi đạo hàm pha dọc tương ứng với độ tán cục bộ của năng lượng theo thời gian và thường được tìm thấy trong các tín hiệu âm thanh mà đã được xử lý bởi các kỹ thuật BWE. Thành phần lạ khác là sự thay đổi của đạo hàm pha theo thời gian (cũng xem sự nhất quán pha "ngang") mà quan trọng về cảm quan đối với các tín hiệu âm giàu họa âm cao của bất kỳ tần số cơ bản nào. Các thành phần lạ liên quan đến sự thay đổi đạo hàm pha ngang tương ứng với độ lệch tần số cục bộ theo độ cao và thường được tìm thấy trong các tín hiệu âm thanh mà đã được xử lý bằng các kỹ thuật BWE.

Sáng chế đề cập đến các phương tiện để tái điều chỉnh đạo hàm pha dọc hoặc ngang của các tín hiệu như vậy khi đặc tính này đã được hứa hẹn bởi sự ứng dụng của cái gọi là sự mở rộng băng thông (bandwidth extension - BWE) âm thanh. Các phương tiện khác được cung cấp để quyết định liệu sự khôi phục của đạo hàm pha là có lợi về cảm quan và liệu việc điều chỉnh đạo hàm pha dọc hoặc ngang có được ưu tiên về cảm quan không.

Các phương pháp mở rộng băng thông, như sao chép băng phổ (spectral band replication - SBR) [9], thường được sử dụng trong các bộ mã hóa-giải mã tốc độ bit

thấp. Chúng cho phép truyền chỉ vùng tần số thấp tương đối hẹp dọc theo thông tin tham số về các băng cao hơn. Vì tốc độ bit của thông tin tham số là nhỏ, có thể thu được cải thiện đáng kể về hiệu suất mã hóa.

Thông thường, tín hiệu cho các băng cao hơn thu được đơn giản bằng cách sao chép nó từ vùng tần số thấp được truyền. Việc xử lý thường được thực hiện trong miền giàn bộ lọc gương cầu phương (quadrature-mirror-filter-bank - QMF) được điều biến phức [10], mà cũng được giả định dưới đây. Tín hiệu được sao chép được xử lý bằng cách nhân phổ cường độ của nó với các độ khuếch đại thích hợp dựa trên các tham số được truyền. Mục đích là để thu được phổ cường độ tương tự như phổ cường độ của tín hiệu ban đầu. Ngược lại, phổ pha của tín hiệu được sao chép thường không được xử lý tất cả với nhau, nhưng, thay vào đó, phổ pha được sao chép được sử dụng trực tiếp.

Các kết quả cảm quan của việc sử dụng trực tiếp phổ pha được sao chép được nghiên cứu sau đây. Dựa trên các hiệu quả được quan sát, hai chuẩn đo để phát hiện các hiệu quả đáng kể nhất về cảm quan được gợi ý. Hơn nữa, các phương pháp làm thế nào để hiệu chỉnh phổ pha dựa trên chúng được gợi ý. Cuối cùng, các chiến lược để tối thiểu hóa lượng tham số được truyền để thực hiện sự hiệu chỉnh được gợi ý.

Sáng chế đề cập đến việc tìm ra rằng sự bảo toàn và khôi phục của đạo hàm pha có thể bù đắp các thành phần lạ nổi bật được cảm ứng bởi các kỹ thuật mở rộng băng thông (bandwidth extension - BWE) âm thanh. Ví dụ, các tín hiệu điển hình, trong đó sự bảo toàn đạo hàm pha là quan trọng, là các âm với nội dung họa âm cao giàu sóng hài, như tiếng nói hữu âm, các nhạc khí bằng đồng hoặc các đàn dây cung.

Sáng chế còn đề xuất phương tiện để quyết định - cho khung tín hiệu đã cho - sự khôi phục của đạo hàm pha có lợi về cảm quan và liệu việc điều chỉnh đạo hàm pha dọc hoặc ngang có được ưu tiên về cảm quan.

Sáng chế dẫn dắt thiết bị và phương pháp cho sự hiệu chỉnh đạo hàm pha trong các bộ mã hóa-giải mã âm thanh sử dụng các kỹ thuật BWE với các khía cạnh sau:

1. Lượng tử hóa của "sự quan trọng" của hiệu chỉnh đạo hàm pha
2. Ưu tiên hóa phụ thuộc tín hiệu của hiệu chỉnh đạo hàm pha ("tần số")

dọc hoặc hiệu chỉnh đạo hàm pha ("thời gian") ngang

3. Chuyển đổi phụ thuộc tín hiệu của hướng hiệu chỉnh ("tần số" hoặc "thời gian")
4. Chế độ hiệu chỉnh đạo hàm pha dọc riêng cho các chuyển tiếp
5. Thu được các tham số ổn định cho hiệu chỉnh nhẵn
6. Định dạng truyền thông tin phụ gọn của các tham số hiệu chỉnh

2. Phép biểu diễn các tín hiệu của miền QMF

Tín hiệu miền thời gian $x(m)$, trong đó m là thời gian rời rạc, có thể được biểu diễn trong miền thời gian - tần số, ví dụ, sử dụng giàn bộ lọc gương cầu phương được điều biến phức (QMF). Tín hiệu kết quả $X(k, n)$, trong đó k là chỉ số băng và n chỉ số khung theo thời gian. QMF của 64 băng và tần số lấy mẫu f_s của 48kHz được giả định để trực quan hóa và các phương án. Do đó, băng thông f_{BW} của mỗi băng là 375Hz và kích thước bước nhảy theo thời gian t_{hop} (17 trên Fig.2) là 1,33 miligiây (millisecond - ms). Tuy nhiên, việc xử lý không chỉ giới hạn ở phép biến đổi như vậy. Cách khác, phép biến đổi cosin rời rạc cải biên (Modified Discrete Cosine Transform - MDCT) hoặc phép biến đổi Fourier rời rạc (Discrete Fourier Transform - DFT) thay vào đó có thể được sử dụng.

Tín hiệu kết quả $X(k, n)$, trong đó k là chỉ số băng và n là chỉ số khung theo thời gian. $X(k, n)$ là tín hiệu phức. Do đó, nó có thể được biểu diễn sử dụng cường độ $X^{mag}(k, n)$ và các phần tử pha $X^{pha}(k, n)$ với j là số phức:

$$X(k, n) = X^{mag}(k, n)e^{jX^{pha}(k, n)}. \quad (1)$$

Các tín hiệu âm thanh được biểu diễn hầu hết sử dụng $X^{mag}(k, n)$ và $X^{pha}(k, n)$ (xem Fig.1 cho hai ví dụ).

Fig.1a thể hiện phổ cường độ $X^{mag}(k, n)$ của tín hiệu đàn viôlông, trong đó Fig.1b thể hiện phổ pha tương ứng $X^{pha}(k, n)$, cả hai trong miền QMF. Hơn nữa, Fig.1c thể hiện phổ cường độ $X^{mag}(k, n)$ của tín hiệu kèn trombon, trong đó Fig.1d thể hiện phổ pha tương ứng lần nữa trong miền QMF. Đối với các phổ cường độ trên Fig.1a và Fig.1c, gradien màu biểu thị cường độ từ đỏ = 0dB đến xanh da trời = -80dB.

Hơn nữa, đối với các phổ trên Fig.1b và Fig.1d, gradien màu biểu thị pha từ đỏ = π sang xanh da trời = $-\pi$.

3. Dữ liệu âm thanh

Dữ liệu âm thanh được sử dụng để thể hiện hiệu quả của việc xử lý âm thanh được mô tả được gọi là "kèn trombon" cho tín hiệu âm thanh của kèn trombon, "đàn viôlông" cho tín hiệu âm thanh của đàn viôlông, và "đàn viôlông + tiếng vỗ tay" cho tín hiệu đàn viôlông với tiếng vỗ tay được thêm vào ở giữa.

4. Phép toán cơ bản của SBR.

Fig.2 thể hiện sơ đồ thời gian tần số 5 bao gồm các ô thời gian tần số 10 (ví dụ, các ngăn QMF, các ngăn giàn bộ lọc gương cầu phương), được định rõ bởi khung thời gian 15 và băng phụ 20. Tín hiệu âm thanh có thể được biến đổi thành phép biểu diễn thời gian tần số sử dụng phép biến đổi giàn bộ lọc gương cầu phương (Quadrature Mirror Filter bank - QMF), phép biến đổi cosin rời rạc cải biên (Modified Discrete Cosine Transform - MDCT), hoặc phép biến đổi Fourier rời rạc. Sự phân chia tín hiệu âm thanh trong các khung thời gian có thể bao gồm các phần chồng lấp của tín hiệu âm thanh. Trong phần phía dưới trên Fig.1, sự chồng lấp đơn của các khung thời gian 15 được thể hiện, trong đó tối đa hai khung thời gian chồng lấp cùng lúc. Hơn nữa, tức là, nếu cần nhiều phần dư, tín hiệu âm thanh có thể được chia cũng sử dụng sự đa chồng lấp. Trong thuật toán đa chồng lấp, ba hoặc nhiều hơn ba khung thời gian có thể bao gồm phần chung của tín hiệu âm thanh tại thời điểm nhất định. Khoảng thời gian của sự chồng lấp là kích thước bước nhảy t_{hop} 17.

Giả sử tín hiệu $X(k, n)$, tín hiệu được mở rộng băng thông (bandwidth-extended - BWE) $Z(k, n)$ thu được từ tín hiệu đầu vào $X(k, n)$ bằng cách sao chép các phần nhất định của băng tần số thấp được truyền. Thuật toán SBR bắt đầu bằng cách lựa chọn vùng tần số sẽ được truyền. Trong ví dụ này, các băng từ 1 đến 7 được lựa chọn:

$$\forall 1 \leq k \leq 7 : X_{trans}(k, n) = X(k, n) . \quad (2)$$

Lượng băng sẽ được truyền phụ thuộc vào tốc độ bit mong muốn. Các hình vẽ và các phương trình được tạo ra sử dụng 7 băng, và từ 5 đến 11 băng được sử dụng cho dữ liệu âm thanh tương ứng. Do đó, các tần số giao chéo giữa vùng tần số được

truyền và các băng cao hơn lần lượt nằm trong khoảng từ 1875 đến 4125Hz. Các băng trên vùng này không được truyền chút nào, nhưng thay vào đó, siêu dữ liệu tham số được tạo ra để mô tả chúng. $X_{\text{trans}}(k, n)$ được mã hóa và được truyền. Vì mục đích dễ hiểu, giả sử rằng việc mã hóa không cải biên tín hiệu bằng bất kì cách nào, thậm chí cần phải nhìn ra rằng việc xử lý khác 0 bị giới hạn ở trường hợp được giả định.

Ở đầu nhận, vùng tần số được truyền được sử dụng trực tiếp cho các tần số tương ứng.

Đối với các băng cao hơn, tín hiệu có thể được tạo ra làm thế nào đó sử dụng tín hiệu được truyền. Một phương thức để đơn giản sao chép tín hiệu được truyền đến các tần số cao hơn. Phiên bản cải biên một chút được sử dụng ở đây. Đầu tiên, tín hiệu băng cơ sở được lựa chọn. Có thể là toàn bộ tín hiệu được truyền, nhưng trong phương án này, băng thứ nhất được loại bỏ. Lý do cho điều này là phổ pha được thông báo là không đều cho băng thứ nhất trong nhiều trường hợp. Do đó, băng cơ sở sẽ được sao chép được định nghĩa như:

$$\forall 1 \leq k \leq 6 : X_{\text{base}}(k, n) = X_{\text{trans}}(k + 1, n) . \quad (3)$$

Các băng thông khác có thể cũng được sử dụng cho các tín hiệu được truyền và các tín hiệu băng cơ sở. Sử dụng tín hiệu băng cơ sở, các tín hiệu thô cho các tần số cao hơn được tạo ra:

$$Y_{\text{raw}}(k, n, i) = X_{\text{base}}(k, n), \quad (4)$$

trong đó $Y_{\text{raw}}(k, n, i)$ là tín hiệu QMF phức cho phần ráp nối tần số i . Các tín hiệu phần ráp nối tần số thô được thao tác theo siêu dữ liệu được truyền bằng cách nhân chúng với các độ khuếch đại $g(k, n, i)$:

$$Y(k, n, i) = Y_{\text{raw}}(k, n, i)g(k, n, i). \quad (5)$$

Nên lưu ý rằng các độ khuếch đại được tạo giá trị thực, và do đó, chỉ phổ cường độ bị ảnh hưởng và do đó được thích ứng đến giá trị đích mong muốn. Các phương thức đã biết thể hiện làm thế nào thu được các độ khuếch đại. Pha đích duy trì không được hiệu chỉnh trong các phương án đã biết nêu trên.

Tín hiệu cuối cùng sẽ được tái tạo thu được bằng cách nối chuỗi các tín hiệu được truyền và các tín hiệu ráp nối để kéo dài một cách liền mạch băng thông để thu

được tín hiệu BWE của băng thông mong muốn. Trong phương án này, $i = 7$ được giả định.

$$\begin{aligned} Z(k, n) &= X_{\text{trans}}(k, n), \\ Z(k + 6i + 1, n) &= Y(k, n, i). \end{aligned} \quad (6)$$

Fig.3 thể hiện các tín hiệu được mô tả trong phép biểu diễn đồ họa. Fig.3a thể hiện sơ đồ tần số minh họa của tín hiệu âm thanh, trong đó cường độ của tần số được biểu thị qua mười băng phụ khác nhau. Bảy băng phụ thứ nhất phản ánh các băng được truyền $X_{\text{trans}}(k, n)$ 25. Băng cơ sở $X_{\text{base}}(k, n)$ 30 được suy ra từ đó bằng cách chọn các băng phụ từ thứ hai đến thứ bảy. Fig.3a biểu diễn tín hiệu âm thanh ban đầu, tức là, tín hiệu âm thanh trước sự truyền hoặc sự mã hóa. Fig.3b thể hiện phép biểu diễn tần số minh họa của tín hiệu âm thanh sau khi thu nhận, ví dụ, trong suốt quy trình giải mã tại bước trung gian. Phổ tần số của tín hiệu âm thanh bao gồm các băng được truyền 25 và bảy tín hiệu băng cơ sở 30 được sao chép đến các băng phụ cao hơn của phổ tần số tạo ra tín hiệu âm thanh 32 bao gồm các tần số cao hơn các tần số trong băng cơ sở. Tín hiệu băng cơ sở hoàn chỉnh cũng được đề cập đến như phần ráp nối tần số. Fig.3c biểu diễn tín hiệu âm thanh được khôi phục $Z(k, n)$ 35. So với Fig.3b, các phần ráp nối của các tín hiệu băng cơ sở được nhân riêng với thừa số khuếch đại. Do đó, phổ tần số của tín hiệu âm thanh bao gồm phổ tần số chính 25 và các phần ráp nối được hiệu chỉnh cường độ $Y(k, n, 1)$ 40. Phương pháp ráp nối này được đề cập đến như sự ráp nối sao chép trực tiếp. Sự ráp nối sao chép trực tiếp được sử dụng minh họa để mô tả sáng chế, thậm chí sáng chế không chỉ giới hạn ở thuật toán ráp nối như vậy. Thuật toán ráp nối khác mà có thể được sử dụng là, ví dụ, thuật toán ráp nối sóng hài.

Giả sử rằng phép biểu diễn theo tham số của các băng cao hơn là hoàn hảo, tức là, phổ cường độ của tín hiệu được khôi phục giống với phổ đó của tín hiệu ban đầu:

$$Z^{\text{mag}}(k, n) = X^{\text{mag}}(k, n). \quad (7)$$

Tuy nhiên, nên lưu ý rằng phổ pha không được hiệu chỉnh bằng bất kì cách nào bởi thuật toán, vậy nên nó không chính xác thậm chí nếu thuật toán được làm việc một cách hiệu quả. Do đó, các phương án cho thấy làm thế nào để thích ứng bổ sung và hiệu chỉnh phổ pha của $Z(k, n)$ cho giá trị đích sao cho sự phát triển chất lượng cảm quan thu được. Trên các phương án, sự hiệu chỉnh có thể được thực hiện sử dụng ba

chế độ xử lý khác nhau, "ngang", "dọc", và "chuyển tiếp". Các chế độ này được thảo luận một cách riêng lẻ dưới đây.

$Z^{\text{mag}}(k, n)$ và $Z^{\text{pha}}(k, n)$ được mô tả trên Fig.4 cho các tín hiệu đàn viôlông và kèn trombon. Fig.4 thể hiện dưới dạng minh họa các phổ của tín hiệu âm thanh được khôi phục 35 sử dụng sự sao chép băng thông phổ (spectral bandwidth replication - SBR) với sự ráp nối sao chép trực tiếp. Phổ cường độ $Z^{\text{mag}}(k, n)$ của tín hiệu viôlông được thể hiện trên Fig.4a, trong đó Fig.4b thể hiện phổ pha tương ứng $Z^{\text{pha}}(k, n)$. Fig.4c và Fig.4d thể hiện phổ tương ứng cho tín hiệu kèn trombon. Tất cả các tín hiệu được thể hiện trong miền QMF. Như được thể hiện trên Fig.1, gradien màu biểu thị cường độ từ đỏ = 0dB đến xanh da trời = -80dB, và pha từ đỏ = π sang xanh xanh da trời = $-\pi$. Có thể thấy rằng các phổ pha của chúng là khác so với các phổ của các tín hiệu ban đầu (xem Fig.1). Do SBR, đàn viôlông được nhận biết để chứa sự phi điều hòa và kèn trombon để chứa các nhiễu âm điều biến tại các tần số giao chéo. Tuy nhiên, các đồ thị pha có vẻ hơi ngẫu nhiên, và nó rất khó để diễn đạt chúng khác nhau như thế nào, và các hiệu quả cảm quan của các chênh lệch là như thế nào. Hơn nữa, việc gửi dữ liệu hiệu chỉnh cho loại dữ liệu ngẫu nhiên này là không khả thi trong các ứng dụng mã hóa mà yêu cầu tốc độ bit thấp. Do đó, việc hiểu các hiệu quả cảm quan của phổ pha và tìm ra các chuẩn đo để mô tả chúng là cần thiết. Các chủ đề này được thảo luận trong các phần sau.

5. Ý nghĩa của phổ pha trong miền QMF

Thường thường người ta nghĩ rằng chỉ số của băng định rõ tần số của thành phần âm đơn, chỉ số xác định mức của nó, và pha định rõ "sự định thời" của nó. Tuy nhiên, băng thông của băng QMF là tương đối lớn, và dữ liệu được lấy mẫu quá. Do đó, sự tương tác giữa các ô thời gian - tần số (tức là, các ô QMF) thực tế định rõ tất cả các đặc tính này.

Phép biểu diễn miền thời gian của ngăn QMF đơn với ba giá trị pha khác nhau, tức là $X^{\text{mag}}(3,1) = 1$ và $X^{\text{pha}}(3,1) = 0, \pi/2$, hoặc π được mô tả trên Fig.5. Kết quả là chức năng tương tự đồng bộ với độ dài là 13,3ms. Hình dạng chính xác của chức năng được định rõ bởi tham số pha.

Xem xét trường hợp trong đó chỉ một băng khác 0 đối với tất cả các khung theo

thời gian, tức là,

$$\forall n \in \mathbb{N} : X^{\text{mag}}(3, n) = 1. \quad (8)$$

Bằng cách thay đổi pha giữa các khung theo thời gian với giá trị cố định α , tức là,

$$X^{\text{pha}}(k, n) = X^{\text{pha}}(k, n - 1) + \alpha, \quad (9)$$

đường hình sin được tạo ra. Tín hiệu thu được (tức là, tín hiệu miền thời gian sau phép biến đổi QMF nghịch đảo) được biểu diễn trên Fig.6 với các giá trị là $\alpha = \pi/4$ (phía trên) và $3\pi/4$ (phía dưới). Có thể thấy rằng tần số của đường hình sin bị ảnh hưởng bởi sự thay đổi pha. Miền tần số được thể hiện ở phần bên phải, trong đó miền thời gian của tín hiệu được thể hiện ở phần bên trái trên Fig.6.

Tương ứng, nếu pha được lựa chọn ngẫu nhiên, kết quả là nhiễu âm băng hẹp (xem Fig.7). Do đó, có thể nói rằng pha của ngăn QMF điều khiển nội dung tần số bên trong băng tương ứng.

Fig.8 minh họa hiệu quả được mô tả liên quan đến Fig.6 trong phép biểu diễn thời gian tần số của bốn khung thời gian và bốn băng phụ tần số, trong đó chỉ ba băng phụ bao gồm tần số khác 0. Các kết quả này trong tín hiệu miền tần số từ Fig.6, được biểu diễn dưới dạng giản đồ ở phần bên phải trên Fig.8, và trong phép biểu diễn miền thời gian trên Fig.6 được biểu diễn dưới dạng giản đồ ở phần phía dưới trên Fig.8.

Xem xét trường hợp trong đó chỉ một khung theo thời gian là khác 0 cho tất cả các băng, tức là,

$$\forall k \in \mathbb{N} : X^{\text{mag}}(k, 3) = 1. \quad (10)$$

Bằng cách thay đổi pha giữa các băng với giá trị cố định α , tức là,

$$X^{\text{pha}}(k, n) = X^{\text{pha}}(k - 1, n) + \alpha, \quad (11)$$

sự chuyển tiếp được tạo ra. Tín hiệu thu được (tức là, tín hiệu miền thời gian sau phép biến đổi QMF nghịch đảo) được biểu diễn trên Fig.9 với các giá trị là $\alpha = \pi/4$ (phía trên) và $3\pi/4$ (phía dưới). Có thể thấy rằng vị trí theo thời gian của sự chuyển tiếp bị ảnh hưởng bởi sự thay đổi pha. Miền tần số được thể hiện ở phần bên phải trên Fig.9, trong đó miền thời gian của tín hiệu được thể hiện ở phần bên trái trên Fig.9.

Fig.9.

Tương ứng, nếu pha được lựa chọn ngẫu nhiên, kết quả là khối tín hiệu truyền nhiều âm ngắn (xem Fig.10). Do đó, có thể nói rằng pha của ngân QMF cũng điều khiển các vị trí theo thời gian của các sóng hài bên trong khung theo thời gian tương ứng.

Fig.11 thể hiện sơ đồ thời gian tần số tương ứng với sơ đồ thời gian tần số được thể hiện trên Fig.8. Trên Fig.11, chỉ khung thời gian thứ ba bao gồm các giá trị khác 0 có sự dịch chuyển thời gian là $\pi/4$ từ một băng phụ đến băng phụ khác. Được biến đổi thành miền tần số, thu được tín hiệu miền tần số từ phía bên phải trên Fig.9, được biểu diễn dưới dạng giản đồ trên phía bên phải trên Fig.11. Giản đồ của phép biểu diễn miền thời gian của phần bên trái trên Fig.9 được thể hiện tại phần phía dưới trên Fig.11. Tín hiệu này là kết quả của việc biến đổi miền thời gian tần số thành tín hiệu miền thời gian.

6. Các số đo để mô tả về cảm quan các đặc tính liên quan của phổ pha

Như được thảo luận trong phần 4, phổ pha bản thân nó trông hơi rối, và khó có thể nhìn thấy trực tiếp hiệu quả của nó lên cảm quan là gì. Phần 5 được biểu diễn bởi hai hiệu quả mà có thể được gây ra bằng cách thao tác phổ pha trong miền QMF: (a) sự thay đổi pha không đổi theo thời gian tạo ra đường hình sin và lượng thay đổi pha điều khiển tần số của đường hình sin, và (b) sự thay đổi pha không đổi theo tần số tạo ra sự chuyển tiếp và lượng thay đổi pha điều khiển vị trí theo thời gian của sự chuyển tiếp.

Tần số và vị trí theo thời gian của một phần hiển nhiên đáng kể với cảm quan của con người, vậy nên việc phát hiện các đặc tính này khả năng có lợi. Chúng có thể được ước tính bằng cách tính toán đạo hàm pha theo thời gian (derivative over time - PDT):

$$X^{pdt}(k, n) = X^{pha}(k, n + 1) - X^{pha}(k, n) \quad (12)$$

và bằng cách tính toán đạo hàm pha theo tần số (derivative over frequency - PDF):

$$X^{pdf}(k, n) = X^{pha}(k + 1, n) - X^{pha}(k, n). \quad (13)$$

$X^{\text{pdt}}(k, n)$ liên quan đến tần số và $X^{\text{pdf}}(k, n)$ liên quan đến vị trí theo thời gian của một phần. Do các đặc tính của phép phân tích QMF (làm thế nào các pha của các bộ điều biến của các khung theo thời gian liên kế khớp tại vị trí chuyển tiếp), π được cộng vào các khung theo thời gian đều của $X^{\text{pdf}}(k, n)$ trên các hình vẽ cho các mục đích hiển thị hóa để tạo ra các đường cong nhẵn.

Tiếp theo là kiểm tra xem các số đo này trông như thế nào đối với các tín hiệu ví dụ của chúng ta. Fig.12 thể hiện các đạo hàm cho các tín hiệu của đàn viôlông và kèn trombon. Cụ thể, Fig.12a thể hiện đạo hàm pha theo thời gian $X^{\text{pdt}}(k, n)$ của tín hiệu âm thanh đàn viôlông ban đầu, tức là chưa được xử lý trong miền QMF. Fig.12b thể hiện đạo hàm pha theo tần số $X^{\text{pdf}}(k, n)$. Fig.12c và Fig.12d thể hiện lần lượt đạo hàm pha theo thời gian và đạo hàm pha theo tần số cho tín hiệu kèn trombon. Gradient màu biểu thị các giá trị pha từ đỏ = π đến xanh da trời = $-\pi$. Đối với đàn viôlông, phổ cường độ về cơ bản là nhiều âm cho đến khoảng 0,13 giây (xem Fig.1) và do đó, các đạo hàm cũng bị nhiễu. Bắt đầu từ khoảng 0,13 giây X^{pdt} xuất hiện để có các giá trị tương đối ổn định theo thời gian. Điều này sẽ có nghĩa là tín hiệu chứa các đường hình sin mạnh, tương đối ổn định. Các tần số của các đường hình sin này được xác định bởi các giá trị X^{pdt} . Ngược lại, đồ thị X^{pdf} có vẻ tương đối nhiễu, vậy nên dữ liệu liên quan được tìm ra cho đàn viôlông sử dụng nó.

Đối với kèn trombon, X^{pdt} tương đối nhiễu. Ngược lại, X^{pdf} có vẻ có giá trị giống nhau tại tất cả các tần số. Trong thực tế, điều này có nghĩa là tất cả các thành phần sóng hài được căn chỉnh theo thời gian tạo ra tín hiệu tương tự chuyển tiếp. Các vị trí theo thời gian của các chuyển tiếp được xác định bởi các giá trị X^{pdf} .

Các đạo hàm giống nhau cũng có thể được tính toán cho tín hiệu được xử lý SBR $Z(k, n)$ (xem Fig.13). Các hình vẽ từ Fig.13a đến Fig.13d liên quan trực tiếp đến các hình vẽ từ Fig.12a đến Fig.12d, được suy ra bằng cách sử dụng thuật toán SBR sao chép trực tiếp được mô tả trước đó. Vì phổ pha đơn giản được sao chép từ băng cơ sở đến các phần rập nối cao hơn, các PDT của các phần rập nối tần số giống hệt với các PDT của băng cơ sở. Do đó, đối với đàn viôlông, PDT tương đối nhẵn theo thời gian tạo ra các đường hình sin ổn định, như trong trường hợp của tín hiệu ban đầu. Tuy nhiên, các giá trị Z^{pdt} khác với các giá trị với tín hiệu ban đầu X^{pdt} , mà khiến các

đường hình sin được tạo ra có các tần số khác hơn so với trong tín hiệu ban đầu. Hiệu quả cảm quan của điều này được thảo luận trong phần 7.

Tương ứng, PDF của các phân rập nổi tần số mặt khác giống với băng cơ sở, nhưng tại các tần số giao chéo, PDF, trong thực tế, là ngẫu nhiên. Khi giao chéo, PDF thực tế được tính toán giữa giá trị pha cuối cùng và pha thứ nhất của phân rập nổi tần số, tức là:

$$Z^{pdt}(7, n) = Z^{pha}(8, n) - Z^{pha}(7, n) = Y^{pha}(1, n, i) - Y^{pha}(6, n, i) \quad (14)$$

Các giá trị này phụ thuộc vào PDF thực tế và tần số giao chéo, và chúng không khớp với các giá trị của tín hiệu ban đầu.

Đối với kèn trombon, các giá trị PDF của tín hiệu được sao chép là đúng riêng ra khỏi các tần số giao chéo. Do đó, các vị trí theo thời gian của hầu hết các sóng hài là trong các vị trí đúng, nhưng các sóng hài tại các tần số giao chéo trên thực tế ở tại các vị trí ngẫu nhiên. Hiệu quả cảm quan của điều này được thảo luận trong phần 7.

7. Cảm quan của con người về các sai số pha

Các âm có thể áng chừng được chia thành hai loại: sóng hài và các tín hiệu tương tự nhiễu âm. Các tín hiệu tương tự nhiễu âm có, đã có bởi định nghĩa, các đặc tính pha nhiễu. Do đó, các sai số pha được gây ra bởi SBR được giả định không đáng kể về cảm quan với chúng. Thay vào đó, nó tập trung vào các tín hiệu sóng hài. Hầu hết các nhạc cụ âm nhạc, và cũng cả tiếng nói, tạo ra cấu trúc sóng hài cho tín hiệu, tức là, âm chứa các thành phần đường hình sin mạnh được đặt cách trong tần số bởi tần số cơ bản.

Việc nghe của con người thường được giả định là vận hành như thể nó chứa giàn các bộ lọc thông băng chồng lấp, được đề cập đến như các bộ lọc thính giác. Do đó, việc nghe có thể được giả định là giải quyết các âm phức để các âm theo phần bên trong bộ lọc thính giác được phân tích như một đối tượng. Chiều rộng của các bộ lọc này có thể được làm gần đúng để theo sau băng thông hình chữ nhật tương đương (equivalent rectangular bandwidth - ERB) [11], mà được xác định theo:

$$ERB = 24,7(4,37 f_c + 1), \quad (15)$$

trong đó f_c là tần số trung tâm của băng (theo kHz). Như được thảo luận trong

phần 4, tần số giao chéo giữa băng cơ sở và các phần ráp nối SBR là khoảng 3kHz. Tại các tần số này, ERB là khoảng 350Hz. Băng thông của băng QMF thực tế tương đối gần với nó, 375Hz. Do đó, băng thông của các băng QMF có thể được giả định là theo sau ERB tại các tần số được quan tâm.

Hai đặc tính của âm mà có thể sai do phổ pha lỗi được quan sát trong phần 6: tần số và sự định thời của thành phần riêng phần. Tập trung vào tần số, câu hỏi là, liệu việc nghe của con người nhận biết các tần số của các sóng hài riêng nào? Nếu có thể, vậy độ lệch tần số bị gây ra bởi SBR nên được hiệu chỉnh, và nếu không, sau đó sự hiệu chỉnh không được yêu cầu.

Khái niệm của các sóng hài được giải quyết và chưa được giải quyết [12] có thể được sử dụng để làm rõ chủ đề này. Nếu chỉ có một sóng hài bên trong ERB, sóng hài được gọi là đã được giải quyết. Thường được giả định rằng việc nghe của con người xử lý các sóng hài được giải quyết riêng và, do đó, nhạy cảm với tần số của chúng. Trên thực tế, việc thay đổi tần số của các sóng hài được giải quyết được nhận biết để gây ra sự phi điều hòa.

Tương ứng, nếu nhiều sóng hài bên trong ERB, các sóng hài được gọi là chưa được giải quyết. Việc nghe của con người được cho là không xử lý các sóng hài này một cách riêng lẻ, nhưng thay vào đó, hiệu quả chung được thấy bởi hệ thống thính giác. Kết quả là tín hiệu tuần hoàn và độ dài của chu kỳ được xác định bằng khoảng cách của các sóng hài. Sự cảm quan độ cao liên quan đến chiều dài của chu kỳ, vậy nên việc nghe của con người được cho là nhạy cảm với nó. Dù sao thì, nếu tất cả các sóng hài bên trong phần ráp nối tần số trong SBR được dịch chuyển cùng một lượng, khoảng cách giữa các sóng hài, và do đó độ cao được nhận biết, duy trì giống nhau. Vì vậy, trong trường hợp các sóng hài chưa được giải quyết, việc nghe của con người không nhận biết các độ lệch tần số như sự phi điều hòa.

Các sai số liên quan đến sự định thời được gây ra bởi SBR được xem xét tiếp theo. Bằng cách định thời, vị trí theo thời gian, hoặc pha, của thành phần sóng hài là có ý nghĩa. Điều này nên không bị nhầm với pha của ngăn QMF. Sự cảm quan về các sai số liên quan đến sự định thời được nghiên cứu chi tiết trong [13]. Các tác giả sáng chế đã quan sát được rằng đối với hầu hết các tín hiệu việc nghe của con người không nhạy

cảm với sự định thời, hoặc pha, của các thành phần sóng hài. Tuy nhiên, có các tín hiệu nhất định mà với chúng, việc nghe của con người rất nhạy cảm với sự định thời của các phần. Các tín hiệu gồm có, ví dụ, các âm của kèn trombon và kèn trumpet và tiếng nói. Với các tín hiệu này, góc pha nhất định cùng xảy ra tại cùng thời điểm ngay cùng lúc với tất cả các sóng hài. Tốc độ phóng điện thuộc thần kinh của các băng thính giác khác nhau được mô phỏng trong [13]. Các tác giả sáng chế đã tìm ra rằng với các tín hiệu nhạy cảm pha này, tốc độ phóng điện thuộc thần kinh được tạo ra đạt đỉnh tại tất cả các băng thính giác và tất cả các đỉnh được căn chỉnh theo thời gian. Việc thay đổi pha của thậm chí sóng hài đơn có thể thay đổi việc đạt được đỉnh của tốc độ phóng điện thuộc thần kinh với các tín hiệu này. Theo các kết quả của bài kiểm tra nghe chính thức, việc nghe của con người nhạy cảm với điều này [13]. Các hiệu quả được tạo ra là sự cảm quan của thành phần đường hình sin được cộng vào hoặc nhiễu âm bằng hẹp tại các tần số trong đó pha được điều chỉnh.

Ngoài ra, người ta đã tìm ra rằng độ nhạy với các hiệu quả liên quan đến sự định thời phụ thuộc vào tần số cơ bản của âm sóng hài [13]. Tần số cơ bản càng thấp, các hiệu quả được nhận biết càng lớn. Nếu tần số cơ bản là trên 800Hz, hệ thống thính giác không nhạy cảm với tất cả các hiệu quả liên quan đến sự định thời.

Do đó, nếu tần số cơ bản là thấp và nếu pha của các sóng hài được căn chỉnh theo tần số (mà có nghĩa là các vị trí theo thời gian của các sóng hài được căn chỉnh), các thay đổi về sự định thời, hoặc nói cách khác pha, của các sóng hài có thể được nhận biết bởi việc nghe của con người. Nếu tần số cơ bản là cao và/hoặc pha của các sóng hài không được căn chỉnh theo tần số, việc nghe của con người không nhạy cảm để thay đổi trong sự định thời của các sóng hài.

8. Các phương pháp hiệu chỉnh

Trong phần 7, được lưu ý rằng con người nhạy cảm với các sai số trong các tần số của các sóng hài được giải quyết. Ngoài ra, con người nhạy cảm với các sai số trong các vị trí theo thời gian của các sóng hài nếu tần số cơ bản thấp và nếu các sóng hài được căn chỉnh theo tần số. SBR có thể gây ra cả hai sai số này, như được thảo luận trong phần 6, vậy nên chất lượng được nhận biết có thể được cải thiện bằng cách hiệu chỉnh chúng. Các phương pháp để làm như vậy được gợi ý trong phần này.

Fig.14 minh họa dưới dạng giản đồ ý tưởng cơ bản của các phương pháp hiệu chỉnh. Fig.14a thể hiện dưới dạng giản đồ bốn pha 45a đến 45d của, ví dụ, các khung thời gian tiếp theo hoặc các băng phụ tần số, ở hình tròn đơn vị. Các pha 45a đến 45d được đặt cách đều nhau 90° . Fig.14b thể hiện các pha sau xử lý SBR và, dưới dạng đường nét đứt, các pha được hiệu chỉnh. Pha 45a trước việc xử lý có thể được dịch chuyển đến góc pha 45a'. Điều tương tự áp dụng cho các pha từ 45b đến 45d. Thể hiện rằng sự chênh lệch giữa các pha sau khi xử lý, tức là, đạo hàm pha, có thể bị sai lệch sau khi xử lý SBR. Ví dụ, sự chênh lệch giữa các pha 45a' và 45b' là 110° sau xử lý SBR, mà là 90° trước khi xử lý. Các phương pháp hiệu chỉnh sẽ thay đổi các giá trị pha 45b' đến giá trị pha mới 45b'' để truy tìm đạo hàm pha cũ 90° . Sự hiệu chỉnh tương tự được áp dụng cho các pha 45d' và 45d''.

8.1 Hiệu chỉnh các sai số tần số - Sự hiệu chỉnh đạo hàm pha ngang

Như được thảo luận trong phần 7, con người có thể nhận biết sai số trong tần số của sóng hài hầu hết khi chỉ có sóng hài bên trong một ERB. Hơn nữa, băng thông của băng QMF có thể được sử dụng để ước tính ERB tại sự giao chéo thứ nhất. Do đó, vì tần số phải được hiệu chỉnh chỉ khi có một sóng hài bên trong một băng. Điều này rất thuận tiện, vì phần 5 đã thể hiện điều đó, nếu có một sóng hài mỗi băng, các giá trị PDT được tạo ra là ổn định, hoặc thay đổi từ từ theo thời gian, và có khả năng được hiệu chỉnh sử dụng tốc độ bit thấp.

Fig.15 thể hiện bộ xử lý âm thanh 50 để xử lý tín hiệu âm thanh 55. Bộ xử lý âm thanh 50 bao gồm bộ tính toán số đo pha tín hiệu âm thanh 60, bộ xác định số đo pha đích 65 và bộ hiệu chỉnh pha 70. Bộ tính toán số đo pha tín hiệu âm thanh 60 được tạo cấu hình để tính toán số đo pha 80 của tín hiệu âm thanh 55 cho khung thời gian 75. Bộ xác định số đo pha đích 65 được tạo cấu hình để xác định số đo pha đích 85 cho khung thời gian 75 đã nêu. Hơn nữa, bộ hiệu chỉnh pha được tạo cấu hình để hiệu chỉnh các pha 45 của tín hiệu âm thanh 55 cho khung thời gian 75 sử dụng số đo pha được tính toán 80 và số đo pha đích 85 để thu được tín hiệu âm thanh được xử lý 90. Một cách tùy chọn, tín hiệu âm thanh 55 bao gồm nhiều tín hiệu băng phụ 95 cho khung thời gian 75. Các phương án khác của bộ xử lý âm thanh 50 được mô tả đối với Fig.16. Theo phương án, bộ xác định số đo pha đích được tạo cấu hình để xác định số

đo pha đích thứ nhất 85a và số đo pha đích thứ hai 85b cho tín hiệu băng phụ thứ hai 95b. Theo đó, bộ tính toán số đo pha tín hiệu âm thanh 60 được tạo cấu hình để xác định số đo pha thứ nhất 80a cho tín hiệu băng phụ thứ nhất 95a và số đo pha thứ hai 80b cho tín hiệu băng phụ thứ hai 95b. Bộ hiệu chỉnh pha được tạo cấu hình để hiệu chỉnh pha 45a của tín hiệu băng phụ thứ nhất 95a sử dụng số đo pha thứ nhất 80a của tín hiệu âm thanh 65 và số đo pha đích thứ nhất 85a và để hiệu chỉnh pha thứ hai 45b của tín hiệu băng phụ thứ hai 95b sử dụng số đo pha thứ hai 80b của tín hiệu âm thanh 55 và số đo pha đích thứ hai 85b. Do đó, bộ xử lý âm thanh 50 bao gồm bộ tổng hợp tín hiệu âm thanh 100 để tổng hợp tín hiệu âm thanh được xử lý 90 sử dụng tín hiệu băng phụ thứ nhất được xử lý 95a và tín hiệu băng phụ thứ hai được xử lý 95b. Theo các phương án khác, số đo pha 80 là đạo hàm pha theo thời gian. Do đó, bộ tính toán số đo pha tín hiệu âm thanh 60 có thể tính toán, cho mỗi băng phụ 95 của nhiều băng phụ, đạo hàm pha của giá trị pha 45 của khung thời gian hiện thời 75b và giá trị pha của khung thời gian tương lai 75c. Theo đó, bộ hiệu chỉnh pha 70 có thể được tính toán, để mỗi băng phụ 95 của nhiều băng phụ của khung thời gian hiện thời 75b, độ lệch giữa đạo hàm pha đích 85 và đạo hàm pha theo thời gian 80, trong đó sự hiệu chỉnh được thực hiện bởi bộ hiệu chỉnh pha 70 được thực hiện sử dụng độ lệch.

Các phương án thể hiện bộ hiệu chỉnh pha 70 được tạo cấu hình để hiệu chỉnh các tín hiệu băng phụ 95 của các băng phụ khác nhau của tín hiệu âm thanh 55 nằm trong khung thời gian 75, để các tần số của các tín hiệu băng phụ được hiệu chỉnh 95 có các giá trị tần số được phân bố một cách hài hòa đến tần số cơ bản của tín hiệu âm thanh 55. Tần số cơ bản là tần số thấp nhất xảy ra trong tín hiệu âm thanh 55, hoặc nói cách khác, các sóng hài thứ nhất của tín hiệu âm thanh 55.

Hơn nữa, bộ hiệu chỉnh pha 70 được tạo cấu hình để làm nhẵn độ lệch 105 cho mỗi băng phụ 95 của nhiều băng phụ trên khung thời gian trước, khung thời gian hiện thời, và khung thời gian tương lai 75a đến 75c và được tạo cấu hình để giảm các thay đổi nhanh chóng của độ lệch 105 trong băng phụ 95. Theo các phương án khác, việc làm nhẵn là trung bình được gán trọng số, trong đó bộ hiệu chỉnh pha 70 được tạo cấu hình để tính toán trung bình được gán trọng số qua các khung trước, hiện thời, và tương lai 75a đến 75c, được gán trọng số bởi cường độ của tín hiệu âm thanh 55 trong khung trước, hiện thời, và tương lai 75a đến 75c.

Các phương án thể hiện các bước xử lý được mô tả trước đó dựa trên véctor. Do đó, bộ hiệu chỉnh pha 70 được tạo cấu hình để tạo ra véctor của các độ lệch 105, trong đó phần tử thứ nhất của véctor đề cập đến độ lệch thứ nhất 105a cho băng phụ thứ nhất 95a của nhiều băng phụ và phần tử thứ hai của véctor đề cập đến độ lệch thứ hai 105b cho băng phụ thứ hai 95b của nhiều băng phụ từ khung thời gian trước 75a đến khung thời gian hiện thời 75b. Hơn nữa, bộ hiệu chỉnh pha 70 có thể áp dụng véctor của các độ lệch 105 vào pha 45 của tín hiệu âm thanh 55, trong đó phần tử thứ nhất của véctor được áp dụng cho pha 45a của tín hiệu âm thanh 55 trong băng phụ thứ nhất 95a của nhiều băng phụ của tín hiệu âm thanh 55 và phần tử thứ hai của véctor được áp dụng cho pha 45b của tín hiệu âm thanh 55 trong băng phụ thứ hai 95b của nhiều băng phụ của tín hiệu âm thanh 55.

Từ quan điểm khác, có thể nêu rằng toàn bộ quy trình trong bộ xử lý âm thanh 50 được dựa trên cơ sở véctor, trong đó mỗi véctor biểu diễn khung thời gian 75, trong đó mỗi băng phụ 95 của nhiều băng phụ bao gồm phần tử của véctor. Các phương án khác tập trung vào bộ xác định số đo pha đích mà được tạo cấu hình để thu được ước tính tần số cơ bản 85b cho khung thời gian hiện thời 75b, trong đó bộ xác định số đo pha đích 65 được tạo cấu hình để tính toán số đo tần số 85 cho mỗi băng phụ trong số nhiều băng phụ cho khung thời gian 75 sử dụng số đo tần số 85 cho khung thời gian 75. Hơn nữa, bộ xác định số đo pha đích 65 có thể chuyển đổi các ước tính tần số 85 cho mỗi băng phụ 95 trong số nhiều băng phụ thành đạo hàm pha theo thời gian sử dụng tổng số các băng 95 và tần số lấy mẫu của tín hiệu âm thanh 55. Để làm rõ, nên lưu ý rằng đầu ra 85 của bộ xác định số đo pha đích 65 có thể là ước tính tần số hoặc đạo hàm pha theo thời gian, phụ thuộc vào phương án. Do đó, trong một phương án, ước tính tần số đã bao gồm định dạng đúng để còn xử lý trong bộ hiệu chỉnh pha 70, trong đó trong phương án khác, ước tính tần số phải được chuyển đổi thành định dạng thích hợp, mà có thể là đạo hàm pha theo thời gian.

Theo đó, bộ xác định số đo pha đích 65 cũng có thể được xem như dựa trên cơ sở véctor. Do đó, bộ xác định số đo pha đích 65 có thể thực hiện véctor của các ước tính tần số 85 cho mỗi băng phụ 95 trong số nhiều băng phụ, trong đó phần tử thứ nhất của véctor đề cập đến ước tính tần số 85a cho băng phụ thứ nhất 95a và phần tử thứ hai của véctor đề cập đến ước tính tần số 85b cho băng phụ thứ hai 95b. Ngoài ra, bộ xác định

số đo pha đích 65 có thể tính toán ước tính tần số 85 sử dụng các bội số của tần số cơ bản, trong đó ước tính tần số 85 của băng phụ hiện thời 95 là bội số của tần số cơ bản mà gần nhất với tâm của băng phụ 95, hoặc trong đó ước tính tần số 85 của băng phụ hiện thời là tần số biên của băng phụ hiện thời 95 nếu không có bội số nào trong số các bội số của tần số cơ bản nằm trong băng phụ hiện thời 95.

Nói cách khác, thuật toán được gợi ý để hiệu chỉnh các sai số trong các tần số của các sóng hài sử dụng bộ xử lý âm thanh 50 hoạt động như sau. Thứ nhất, PDT được tính toán và tín hiệu được xử lý SBR Z^{pdt} . $Z^{\text{pdt}}(k, n) = Z^{\text{pha}}(k, n + 1) - Z^{\text{pha}}(k, n)$. Sự chênh lệch giữa nó và PDT đích cho sự hiệu chỉnh ngang được tính toán tiếp theo:

$$D^{\text{pdt}}(k, n) = Z^{\text{pdt}}(k, n) - Z_{\text{th}}^{\text{pdt}}(k, n). \quad (16a)$$

Tại mức này, PDT đích có thể được giả thiết là bằng với PDT của đầu vào của tín hiệu đầu vào:

$$Z_{\text{th}}^{\text{pdt}}(k, n) = X^{\text{pdt}}(k, n). \quad (16b)$$

Sau đó, nó sẽ được biểu diễn làm thế nào PDT đích có thể thu được với tốc độ bit thấp.

Giá trị này (tức là, giá trị sai số 105) được làm nhẵn theo thời gian sử dụng của số Hann $W(l)$. Độ dài thích hợp là, ví dụ, 41 mẫu trong miền QMF (tương ứng với khoảng của 55ms). Sự làm nhẵn được gán trọng số bởi cường độ của các ô thời gian - tần số tương ứng:

$$D_{\text{sm}}^{\text{pdt}}(k, n) = \text{circmean}\{D^{\text{pdt}}(k, n + l), W(l)Z^{\text{mag}}(k, n + l)\}, -20 \leq l \leq 20, \quad (17)$$

trong đó $\text{circmean}\{a, b\}$ biểu thị sự tính toán trung bình lặp cho các giá trị góc a được gán trọng số bằng các giá trị b . Sai số được làm nhẵn trong PDT $D_{\text{sm}}^{\text{pdt}}(k, n)$ được biểu thị trên Fig.17 cho tín hiệu đàn viôlông trong miền QMF sử dụng SBR sao chép trực tiếp. Gradient màu biểu thị các giá trị pha từ đỏ $= \pi$ đến xanh da trời $= -\pi$.

Tiếp theo, ma trận điều biến được tạo ra để cải biên phổ pha để thu được PDT mong muốn:

$$Q^{\text{pha}}(k, n+1) = Q^{\text{pha}}(k, n) - D_{\text{sm}}^{\text{pdt}}(k, n). \quad (18)$$

Phổ pha được xử lý sử dụng ma trận này:

$$Z_{\text{ch}}^{\text{pha}}(k, n) = Z^{\text{pha}}(k, n) + Q^{\text{pha}}(k, n). \quad (19)$$

Fig.18a thể hiện sai số trong đạo hàm pha theo thời gian (PDT) $D_{\text{sm}}^{\text{pdt}}(k, n)$ của tín hiệu đàn viôlong trong miền QMF cho SBR được hiệu chỉnh. Fig.18b thể hiện đạo hàm pha theo thời gian tương ứng $Z_{\text{ch}}^{\text{pdt}}(k, n)$, trong đó sai số trong PDT được thể hiện trên Fig.18a được suy ra bằng cách so sánh các kết quả được thể hiện trên Fig.12a với các kết quả được thể hiện trên Fig.18b. Một lần nữa, gradien màu biểu thị các giá trị pha từ đỏ = π đến xanh da trời = $-\pi$. PDT được tính toán cho phổ pha được hiệu chỉnh $Z_{\text{ch}}^{\text{pha}}(k, n)$ (xem Fig.18b). Có thể thấy rằng PDT của phổ pha được hiệu chỉnh nhắc nhở PDT của tín hiệu ban đầu tốt (xem Fig.12), và sai số là nhỏ đối với các ô thời gian - tần số chứa năng lượng đáng kể (xem Fig.18a). Nên lưu ý rằng sự phi điều hòa của dữ liệu SBR không được hiệu chỉnh đã mất hết phần lớn. Ngoài ra, thuật toán đường như không gây ra các thành phần lạ đáng kể.

Sử dụng $X^{\text{pdt}}(k, n)$ như PDT đích, có thể truyền các giá trị sai số PDT $D_{\text{sm}}^{\text{pdt}}(k, n)$ cho mỗi ô thời gian - tần số. Phương pháp khác tính toán PDT đích sao cho băng thông cho việc truyền được giảm được thể hiện trong phần 9.

Trong các phương án khác, bộ xử lý âm thanh 50 có thể là phần của bộ giải mã 110. Do đó, bộ giải mã 110 để giải mã tín hiệu âm thanh 55 có thể bao gồm bộ xử lý âm thanh 50, bộ giải mã lõi 115, và bộ ráp nối 120. Bộ giải mã lõi 115 được tạo cấu hình để giải mã lõi tín hiệu âm thanh 25 trong khung thời gian 75 với số lượng giảm của các băng phụ đối với tín hiệu âm thanh 55. Bộ ráp nối ráp nối tập hợp các băng phụ 95 của tín hiệu âm thanh được giải mã lõi 25 với số lượng các băng phụ giảm, trong đó tập hợp các băng phụ tạo ra phần ráp nối thứ nhất 30a, với các băng phụ khác trong khung thời gian 75, liên kế với số lượng các băng phụ giảm, để thu được tín hiệu âm thanh 55 với số lượng các băng phụ đều. Ngoài ra, bộ xử lý âm thanh 50 được tạo cấu hình để hiệu chỉnh các pha 45 nằm trong các băng phụ của phần ráp nối thứ nhất 30a theo hàm đích 85. Bộ xử lý âm thanh 50 và tín hiệu âm thanh 55 đã được mô tả đối với Fig.15 và Fig.16, trong đó các dấu hiệu tham chiếu không được mô tả trên

Fig.19 được giải thích. Bộ xử lý âm thanh theo các phương án thực hiện sự hiệu chỉnh pha. Phụ thuộc vào các phương án, bộ xử lý âm thanh có thể còn bao gồm sự hiệu chỉnh biên độ của tín hiệu âm thanh bởi bộ áp dụng tham số mở rộng băng thông 125 áp dụng các tham số BWE và SBR đến các phần ráp nối. Hơn nữa, bộ xử lý âm thanh có thể bao gồm bộ tổng hợp 100, ví dụ, giàn bộ lọc tổng hợp, để tổ hợp, tức là, tổng hợp, các băng phụ của tín hiệu âm thanh để thu được tệp âm thanh thường.

Theo các phương án khác, phần ráp nối 120 được tạo cấu hình để ráp nối tập hợp các băng phụ 95 của tín hiệu âm thanh 25, trong đó tập hợp các băng phụ tạo ra phần ráp nối thứ hai, với các băng phụ khác của khung thời gian, liền kề với phần ráp nối thứ nhất trong đó bộ xử lý âm thanh 50 được tạo cấu hình để hiệu chỉnh pha 45 nằm trong các băng phụ của phần ráp nối thứ hai. Ngoài ra, bộ ráp nối 120 được tạo cấu hình để ráp nối phần ráp nối thứ nhất được hiệu chỉnh với các băng phụ khác của khung thời gian, liền kề với phần ráp nối thứ nhất.

Nói cách khác, trong lựa chọn thứ nhất, phần ráp nối xây dựng tín hiệu âm thanh với số lượng các băng phụ đều từ phần được truyền của tín hiệu âm thanh và sau đó các pha của mỗi phần ráp nối của tín hiệu âm thanh được hiệu chỉnh. Tùy chọn thứ hai đầu tiên hiệu chỉnh các pha của phần ráp nối thứ nhất so với phần được truyền của tín hiệu âm thanh và sau đó tích lũy tín hiệu âm thanh với số lượng các băng phụ đều với phần ráp nối thứ nhất đã được hiệu chỉnh.

Các phương án khác thể hiện bộ giải mã 110 bao gồm bộ trích xuất dòng dữ liệu 130 được tạo cấu hình để trích xuất tần số cơ bản 114 của khung thời gian hiện thời 75 của tín hiệu âm thanh 55 từ dòng dữ liệu 135, trong đó dòng dữ liệu còn bao gồm tín hiệu âm thanh được mã hóa 145 với số lượng băng giảm. Ngoài ra, bộ giải mã có thể bao gồm bộ phân tích tần số cơ bản 150 được tạo cấu hình để phân tích tín hiệu âm thanh được giải mã lỗi 25 để tính toán tần số cơ bản 140. Nói cách khác, các tùy chọn để suy ra tần số cơ bản 140 là, ví dụ, phép phân tích của tín hiệu âm thanh trong bộ giải mã hoặc trong bộ mã hóa, trong đó ở trường hợp sau, tần số cơ bản có thể chính xác hơn với cái giá là tốc độ dữ liệu cao hơn, vì giá trị phải được truyền từ bộ giải mã đến bộ mã hóa.

Fig.20 thể hiện bộ mã hóa 155 để mã hóa tín hiệu âm thanh 55. Bộ mã hóa bao

gồm bộ mã hóa lõi 160 để mã hóa lõi tín hiệu âm thanh 55 để thu được tín hiệu âm thanh được mã hóa lõi 145 có số lượng các băng phụ giảm đối với tín hiệu âm thanh và bộ mã hóa bao gồm bộ phân tích tần số cơ bản 175 để phân tích tín hiệu âm thanh 55 hoặc phiên bản được lọc thông thấp 55 để thu được ước tính tần số cơ bản của tín hiệu âm thanh. Hơn nữa, bộ mã hóa bao gồm bộ trích xuất tham số 165 để trích xuất các tham số của các băng phụ của tín hiệu âm thanh 55 không được nằm trong tín hiệu âm thanh được mã hóa lõi 145 và bộ mã hóa bao gồm bộ tạo tín hiệu đầu ra 170 để tạo ra tín hiệu đầu ra 135 bao gồm tín hiệu âm thanh được mã hóa lõi 145, các tham số và ước tính tần số cơ bản. Trong phương án này, bộ mã hóa 155 có thể bao gồm bộ lọc thông thấp ở trước bộ giải mã lõi 160 và bộ lọc thông cao 185 ở trước bộ trích xuất tham số 165. Theo các phương án khác, bộ tạo tín hiệu đầu ra 170 được tạo cấu hình để tạo tín hiệu đầu ra 135 thành chuỗi các khung, trong đó mỗi khung bao gồm tín hiệu được mã hóa lõi 145, các tham số 190, và trong đó chỉ mỗi khung thứ n bao gồm ước tính tần số cơ bản 140, trong đó $n \geq 2$. Trong các phương án, bộ mã hóa lõi 160 có thể, ví dụ, là bộ mã hóa lập mã âm thanh tiên tiến (Advanced Audio Coding - AAC).

Trong phương án thay thế, bộ mã hóa điền đầy khoảng trống thông minh có thể được sử dụng để mã hóa tín hiệu âm thanh 55. Do đó, bộ mã hóa lõi mã hóa tín hiệu âm thanh bằng thông đầy đủ, trong đó ít nhất một băng phụ của tín hiệu âm thanh bị bỏ qua. Do đó, bộ trích xuất tham số 165 trích xuất các tham số để khôi phục các băng phụ bị bỏ qua từ quy trình mã hóa của bộ mã hóa lõi 160.

Fig.21 thể hiện sự minh họa dưới dạng giản đồ của tín hiệu đầu ra 135. Tín hiệu đầu ra là tín hiệu âm thanh bao gồm tín hiệu âm thanh được mã hóa lõi 145 có số lượng các băng phụ giảm đối với tín hiệu âm thanh ban đầu 55, tham số 190 biểu diễn các băng phụ của tín hiệu âm thanh không được chứa trong tín hiệu âm thanh được mã hóa lõi 145, và ước tính tần số cơ bản 140 của tín hiệu âm thanh 135 hoặc tín hiệu âm thanh ban đầu 55.

Fig.22 thể hiện phương án của tín hiệu âm thanh 135, trong đó tín hiệu âm thanh được tạo thành chuỗi các khung 195, trong đó mỗi khung 195 bao gồm tín hiệu âm thanh được mã hóa lõi 145, các tham số 190, và trong đó chỉ mỗi khung thứ n 195 bao gồm ước tính tần số cơ bản 140, trong đó $n \geq 2$. Điều này có thể mô tả phép truyền

ước tính tần số cơ bản cách đều nhau đối với, ví dụ, mỗi khung thứ 20, hoặc trong đó ước tính tần số cơ bản được truyền không đều, ví dụ, theo nhu cầu hoặc theo mục đích.

Fig.23 thể hiện phương pháp 2300 để xử lý tín hiệu âm thanh với bước 2305 "tính toán số đo pha của tín hiệu âm thanh cho khung thời gian với bộ tính toán đạo hàm pha tín hiệu âm thanh", bước 2310 "xác định số đo pha đích cho khung thời gian đã nêu với bộ xác định đạo hàm pha đích", và bước 2315 "hiệu chỉnh các pha của tín hiệu âm thanh cho khung thời gian với bộ hiệu chỉnh pha sử dụng số đo pha tính toán và số đo pha đích để thu được tín hiệu âm thanh được xử lý".

Fig.24 thể hiện phương pháp 2400 để giải mã tín hiệu âm thanh với bước 2405 "giải mã tín hiệu âm thanh trong khung thời gian với số lượng các băng phụ giảm so với tín hiệu âm thanh", bước 2410 "ráp nối tập hợp các băng phụ của tín hiệu âm thanh được giải mã với số lượng các băng phụ giảm, trong đó tập hợp các băng phụ tạo ra phần ráp nối thứ nhất, với các băng phụ khác trong khung thời gian, liền kề với số lượng các băng phụ giảm, để thu được tín hiệu âm thanh với số lượng các băng phụ đều", và bước 2415 "hiệu chỉnh các pha nằm trong các băng phụ của phần ráp nối thứ nhất theo hàm đích với bộ xử lý âm thanh".

Fig.25 thể hiện phương pháp 2500 để mã hóa tín hiệu âm thanh với bước 2505 "mã hóa lõi tín hiệu âm thanh với bộ mã hóa lõi để thu được tín hiệu âm thanh được mã hóa lõi có số lượng các băng phụ giảm đối với tín hiệu âm thanh", bước 2510 "phân tích tín hiệu âm thanh hoặc phiên bản được lọc thông thấp của tín hiệu âm thanh với bộ phân tích tần số cơ bản để thu được ước tính tần số cơ bản cho tín hiệu âm thanh", bước 2515 "trích xuất các tham số của các băng phụ của tín hiệu âm thanh không nằm trong tín hiệu âm thanh được mã hóa lõi với bộ trích xuất tham số", và bước 2520 "tạo ra tín hiệu đầu ra bao gồm tín hiệu âm thanh được mã hóa lõi, các tham số, và ước tính tần số cơ bản với bộ tạo tín hiệu đầu ra".

Các phương pháp 2300, 2400 và 2500 có thể được thực hiện trong mã chương trình của chương trình máy tính để thực hiện các phương pháp khi chương trình máy tính chạy trên máy tính.

8.2 Hiệu chỉnh các sai số theo thời gian - Sự hiệu chỉnh đạo hàm pha dọc

Như được thảo luận trước đó, con người có thể nhận biết sai số trong vị trí theo thời gian của sóng hài nếu các sóng hài được đồng bộ theo tần số và nếu tần số cơ bản là thấp. Trong phần 5, sáng chế đã thể hiện rằng các sóng hài được đồng bộ nếu đạo hàm pha theo tần số là bất biến trong miền QMF. Do đó, sẽ có lợi để có ít nhất một sóng hài trong mỗi băng. Nếu không, các băng "trống" sẽ có các pha ngẫu nhiên và sẽ làm nhiễu loạn số đo này. May mắn là, con người nhạy với vị trí theo thời gian của các sóng hài chỉ khi tần số cơ bản là thấp (xem phần 7). Do đó, đạo hàm pha theo tần số có thể được sử dụng làm số đo để xác định các ảnh hưởng đáng kể về cảm quan do các chuyển động theo thời gian của các sóng hài.

Fig.26 thể hiện sơ đồ khối dạng giản lược của bộ xử lý âm thanh 50' để xử lý tín hiệu âm thanh 55, trong đó bộ xử lý âm thanh 50' bao gồm bộ xác định số đo pha đích 65', bộ tính toán sai số pha 200, và bộ hiệu chỉnh pha 70'. Bộ xác định số đo pha đích 65' xác định số đo pha đích 85' cho tín hiệu âm thanh 55 trong khung thời gian 75. Bộ tính toán sai số pha 200 tính toán sai số pha 105' sử dụng pha của tín hiệu âm thanh 55 trong khung thời gian 75 và số đo pha đích 85'. Bộ hiệu chỉnh pha 70' hiệu chỉnh pha của tín hiệu âm thanh 55 trong khung thời gian sử dụng sai số pha 105' tạo ra tín hiệu âm thanh được xử lý 90'.

Fig.27 thể hiện sơ đồ khối dạng giản lược của bộ xử lý âm thanh 50' theo phương án khác. Do đó, tín hiệu âm thanh 55 bao gồm nhiều băng phụ 95 cho khung thời gian 75. Theo đó, bộ xác định số đo pha đích 65' được tạo cấu hình để xác định số đo pha đích thứ nhất 85a' cho tín hiệu băng phụ thứ nhất 95a và số đo pha đích thứ hai 85b' cho tín hiệu băng phụ thứ hai 95b. Bộ tính toán sai số pha 200 tạo ra vectơ của các sai số pha 105', trong đó phần tử thứ nhất của vectơ đề cập đến độ lệch thứ nhất 105a' của pha của tín hiệu băng phụ thứ nhất 95 và số đo pha đích thứ nhất 85a' và trong đó phần tử thứ hai của vectơ đề cập đến độ lệch thứ hai 105b' của pha của tín hiệu băng phụ thứ hai 95b và số đo pha đích thứ hai 85b'. Do đó, bộ xử lý âm thanh 50 bao gồm bộ tổng hợp tín hiệu âm thanh 100 để tổng hợp tín hiệu âm thanh được hiệu chỉnh 90' sử dụng tín hiệu băng phụ thứ nhất được hiệu chỉnh 90a' và tín hiệu băng phụ thứ hai được hiệu chỉnh 90b'.

Về các phương án khác, nhiều băng phụ 95 được nhóm thành băng cơ sở 30 và

tập hợp các phần ráp nối tần số 40, băng cơ sở 30 bao gồm một băng phụ 95 của tín hiệu âm thanh 55 và tập hợp các phần ráp nối tần số 40 bao gồm ít nhất một băng phụ 95 của băng cơ sở 30 tại tần số cao hơn tần số của ít nhất một băng phụ trong băng cơ sở. Phải lưu ý rằng phần ráp nối của tín hiệu âm thanh đã được mô tả đối với Fig.3 và do đó sẽ không được mô tả cụ thể trong phần này của phần mô tả. Phải đề cập đến việc các phần ráp nối tần số 40 có thể là tín hiệu băng cơ sở thô được sao chép cho các tần số cao hơn được nhận với thừa số khuếch đại trong đó sự hiệu chỉnh pha có thể được áp dụng. Hơn nữa, theo phương án được ưu tiên, phép nhân độ khuếch đại và sự hiệu chỉnh pha có thể được chuyển sao cho các pha của tín hiệu băng cơ sở thô được sao chép đến các tần số cao hơn trước khi được nhân với thừa số khuếch đại. Phương án khác thể hiện bộ tính toán sai số pha 200 tính toán trung bình các phần tử của vectơ của các sai số pha 105' đề cập đến phần ráp nối thứ nhất 40a trong số tập hợp các phần ráp nối tần số 40 để thu được sai số pha trung bình 105". Do đó, bộ tính toán đạo hàm pha tín hiệu âm thanh 210 được thể hiện để tính toán trung bình của các đạo hàm pha theo tần số 215 cho băng cơ sở 30.

Fig.28a thể hiện sự mô tả chi tiết hơn của bộ hiệu chỉnh pha 70' theo sơ đồ khối. Bộ hiệu chỉnh pha 70' ở phía trên cùng trên Fig.28a được tạo cấu hình để hiệu chỉnh pha của các tín hiệu băng phụ 95 trong các phần ráp nối tần số thứ nhất và tiếp theo trong số tập hợp các phần ráp nối tần số. Trong phương án trên Fig.28a, đã được minh họa rằng các băng phụ 95c và 95d thuộc về phần ráp nối 40a và các băng phụ 95e và 95f thuộc về phần ráp nối 40b. Các pha được hiệu chỉnh sử dụng sai số pha trung bình được gán trọng số, trong đó sai số pha trung bình 105 được gán trọng số theo chỉ số của phần ráp nối tần số 40 để thu được tín hiệu phần ráp nối cải biên 40'.

Phương án khác được mô tả tại phần dưới cùng trên Fig.28a. Ở góc bên trái trên cùng của bộ hiệu chỉnh pha 70', phương án đã được mô tả được thể hiện để thu được tín hiệu phần ráp nối cải biên 40' từ các phần ráp nối 40 và sai số pha trung bình 105". Hơn nữa, bộ hiệu chỉnh pha 70' tính toán theo bước khởi tạo tín hiệu phần ráp nối cải biên 40" khác với phần ráp nối tần số thứ nhất được tối ưu hóa bằng cách cộng trung bình của các đạo hàm pha theo tần số 215, được gán trọng số bởi chỉ số băng phụ hiện thời vào pha của tín hiệu băng phụ với chỉ số băng phụ cao nhất trong băng cơ sở 30 của tín hiệu âm thanh 55. Đối với bước khởi tạo, bộ chuyển 220a là trong vị trí bên trái

của nó. Đối với bước xử lý khác bất kỳ, bộ chuyển sẽ nằm trong vị trí khác tạo ra sự kết nối định hướng dọc.

Trong phương án khác, bộ tính toán đạo hàm pha tín hiệu âm thanh 210 được tạo cấu hình để tính toán trung bình của các đạo hàm pha theo tần số 215 cho nhiều tín hiệu băng phụ bao gồm các tần số cao hơn là tín hiệu băng cơ sở 30 để phát hiện các chuyển tiếp trong tín hiệu băng phụ 95. Phải lưu ý rằng sự hiệu chỉnh chuyển tiếp tương tự với sự hiệu chỉnh pha dọc của bộ xử lý âm thanh 50' với sự khác biệt là các tần số trong băng cơ sở 30 không phản xạ các tần số cao hơn của sự chuyển tiếp. Do đó, các tần số này phải được xem xét cho sự hiệu chỉnh pha của sự chuyển tiếp.

Sau bước khởi tạo, bộ hiệu chỉnh pha 70' được tạo cấu hình để cập nhật đệ quy, dựa trên các phần rập nối tần số 40, tín hiệu phần rập nối cải biên 40'' khác bằng cách cộng trung bình của các đạo hàm pha theo tần số 215, được gán trọng số bởi chỉ số băng phụ của băng phụ hiện thời 95, vào pha của tín hiệu băng phụ với chỉ số băng phụ cao nhất trong phần rập nối tần số trước đó. Phương án được ưu tiên là sự tổ hợp của các phương án được mô tả trước đó, trong đó bộ hiệu chỉnh pha 70' tính toán trung bình được gán trọng số của tín hiệu phần rập nối cải biên 40' và tín hiệu phần rập nối cải biên 40'' khác để thu được tín hiệu phần rập nối cải biên được tổ hợp 40'''. Do đó, bộ hiệu chỉnh pha 70' cập nhật đệ quy, dựa trên các phần rập nối tần số 40, tín hiệu phần rập nối cải biên được tổ hợp 40''' bằng cách cộng trung bình của các đạo hàm pha theo tần số 215, được gán trọng số bởi chỉ số băng phụ của băng phụ hiện thời 95, vào pha của tín hiệu băng phụ với chỉ số băng phụ cao nhất trong phần rập nối tần số trước đó của tín hiệu phần rập nối cải biên được tổ hợp 40'''. Để thu được các phần rập nối cải biên được tổ hợp 40a'', 40b'', v.v., bộ chuyển 220b được dịch chuyển đến vị trí tiếp theo sau phép đệ quy, bắt đầu từ phần rập nối cải biên được tổ hợp 40''' cho bước khởi tạo, chuyển thành phần rập nối cải biên được tổ hợp 40b'' sau phép đệ quy thứ nhất và cứ như vậy.

Hơn nữa, bộ hiệu chỉnh pha 70' có thể tính toán trung bình được gán trọng số của tín hiệu phần rập nối 40' và tín hiệu phần rập nối cải biên 40'' sử dụng trung bình lặp của tín hiệu phần rập nối 40' trong phần rập nối tần số hiện thời được gán trọng số với hàm gán trọng số cụ thể thứ nhất và tín hiệu phần rập nối cải biên 40'' trong phần

ráp nối tần số hiện thời được gán trọng số với hàm gán trọng số cụ thể thứ hai.

Để cung cấp khả năng tương tác giữa bộ xử lý âm thanh 50 và bộ xử lý âm thanh 50', bộ hiệu chỉnh pha 70' có thể tạo ra vectơ của các độ lệch pha, trong đó các độ lệch pha được tính toán sử dụng tín hiệu phần ráp nối cải biên được tổ hợp 40''' và tín hiệu âm thanh 55.

Fig.28b minh họa các bước hiệu chỉnh pha từ quan điểm khác. Đối với khung thời gian thứ nhất, tín hiệu phần ráp nối 40' được suy ra bằng cách sử dụng chế độ hiệu chỉnh pha thứ nhất trên các phần ráp nối của tín hiệu âm thanh 55. Tín hiệu phần ráp nối 40' được sử dụng trong bước khởi tạo của chế độ hiệu chỉnh thứ hai để thu được tín hiệu phần ráp nối cải biên 40''. Sự tổ hợp tín hiệu phần ráp nối 40' và tín hiệu phần ráp nối cải biên 40'' dẫn đến tín hiệu phần ráp nối cải biên được tổ hợp 40'''.

Chế độ hiệu chỉnh thứ hai do đó được áp dụng trên tín hiệu phần ráp nối cải biên được tổ hợp 40''' để thu được tín hiệu phần ráp nối cải biên 40'' cho khung thời gian thứ hai 75b. Ngoài ra, chế độ hiệu chỉnh thứ nhất được sử dụng cho các phần ráp nối của tín hiệu âm thanh 55 trên khung thời gian thứ hai 75b để thu được tín hiệu phần ráp nối 40'. Một lần nữa, sự tổ hợp tín hiệu phần ráp nối 40' và tín hiệu phần ráp nối cải biên 40'' dẫn đến tín hiệu phần ráp nối cải biên được tổ hợp 40''' . Sơ đồ xử lý được mô tả cho khung thời gian thứ hai được áp dụng cho khung thời gian thứ ba 75c và bất kỳ khung thời gian nào của tín hiệu âm thanh 55 theo đó.

Fig.29 thể hiện sơ đồ khối chi tiết của bộ xác định số đo pha đích 65'. Theo phương án, bộ xác định số đo pha đích 65' bao gồm bộ trích xuất dòng dữ liệu 130' để trích xuất vị trí đỉnh 230 và tần số cơ bản của các vị trí đỉnh 235 trong khung thời gian hiện thời của tín hiệu âm thanh 55 từ dòng dữ liệu 135. Ngoài ra, bộ xác định số đo pha đích 65' bao gồm bộ phân tích tín hiệu âm thanh 225 để phân tích tín hiệu âm thanh 55 trong khung thời gian hiện thời để tính toán vị trí đỉnh 230 và tần số cơ bản của các vị trí đỉnh 235 trong khung thời gian hiện thời. Ngoài ra, bộ xác định số đo pha đích bao gồm bộ tạo phổ đích 240 để ước tính các vị trí đỉnh khác trong khung thời gian hiện thời sử dụng vị trí đỉnh 230 và tần số cơ bản của các vị trí đỉnh 235.

Fig.30 minh họa sơ đồ khối chi tiết của bộ tạo phổ đích 240 được mô tả trên Fig.29. Bộ tạo phổ đích 240 bao gồm bộ tạo đỉnh 245 để tạo ra chuỗi xung 265 theo

thời gian. Bộ tạo tín hiệu 250 điều chỉnh tần số của chuỗi xung theo tần số cơ bản của các vị trí đỉnh 235. Hơn nữa, bộ định vị xung 255 điều chỉnh pha của chuỗi xung 265 theo vị trí đỉnh 230. Nói cách khác, bộ tạo tín hiệu 250 thay đổi dạng của tần số ngẫu nhiên của chuỗi xung 265 sao cho tần số của chuỗi xung bằng với tần số cơ bản của các vị trí đỉnh của tín hiệu âm thanh 55. Hơn nữa, bộ định vị xung 255 dịch chuyển pha của chuỗi xung sao cho một trong số các đỉnh của chuỗi xung bằng với vị trí đỉnh 230. Sau đó, bộ phân tích phổ 260 tạo ra phổ pha của chuỗi xung được điều chỉnh, trong đó phổ pha của tín hiệu miền thời gian là số đo pha đích 85'.

Fig.31 thể hiện sơ đồ khối dạng giản lược của bộ giải mã 110' để giải mã tín hiệu âm thanh 55. Bộ giải mã 110 bao gồm bộ giải mã lõi 115 được tạo cấu hình để giải mã tín hiệu âm thanh 25 trong khung thời gian của băng cơ sở, và bộ ráp nối 120 để ráp nối tập hợp các băng phụ 95 của băng cơ sở được giải mã, trong đó tập hợp các băng phụ tạo ra phần ráp nối, với các băng phụ khác trong khung thời gian, liền kề với băng cơ sở, để thu được tín hiệu âm thanh 32 bao gồm các tần số cao hơn các tần số trong băng cơ sở. Hơn nữa, bộ giải mã 110' bao gồm bộ xử lý âm thanh 50' để hiệu chỉnh các pha của các băng phụ của phần ráp nối theo số đo pha đích.

Theo các phương án khác, phần ráp nối 120 được tạo cấu hình để ráp nối tập hợp các băng phụ 95 của tín hiệu âm thanh 25, trong đó tập hợp các băng phụ tạo ra phần ráp nối khác, với các băng phụ khác của khung thời gian, liền kề với phần ráp nối trong đó bộ xử lý âm thanh 50' được tạo cấu hình để hiệu chỉnh các pha nằm trong các băng phụ của phần ráp nối khác. Ngoài ra, bộ ráp nối 120 được tạo cấu hình để ráp nối phần ráp nối được hiệu chỉnh với các băng phụ khác của khung thời gian, liền kề với phần ráp nối.

Phương án khác đề cập đến bộ giải mã để giải mã tín hiệu âm thanh bao gồm sự chuyển tiếp, trong đó bộ xử lý âm thanh 50' được tạo cấu hình để hiệu chỉnh pha của sự chuyển tiếp. Việc xử lý sự chuyển tiếp được mô tả cách khác trong phần 8.4. Do đó, bộ giải mã 110 bao gồm bộ xử lý âm thanh 50' khác để nhận đạo hàm pha theo tần số khác và để hiệu chỉnh các phần chuyển tiếp trong tín hiệu âm thanh 32 sử dụng đạo hàm pha theo tần số được nhận. Hơn nữa, phải lưu ý rằng bộ giải mã 110' trên Fig.31 tương tự với bộ giải mã 110 trên Fig.19, sao cho phần mô tả liên quan đến các phần tử

chính là có thể trao đổi lẫn nhau trong các trường hợp không liên quan đến sự khác biệt của các bộ xử lý âm thanh 50 và 50'.

Fig.32 thể hiện bộ mã hóa 155' để mã hóa tín hiệu âm thanh 55. Bộ mã hóa 155' bao gồm bộ mã hóa lõi 160, bộ phân tích tần số cơ bản 175', bộ trích xuất tham số 165, và bộ tạo tín hiệu đầu ra 170. Bộ mã hóa lõi 160 được tạo cấu hình để mã hóa lõi tín hiệu âm thanh 55 để thu được tín hiệu âm thanh được mã hóa lõi 145 có số lượng các băng phụ giảm so với tín hiệu âm thanh 55. Bộ phân tích tần số cơ bản 175' phân tích các vị trí đỉnh 230 trong tín hiệu âm thanh 55 hoặc phiên bản được lọc thông thấp của tín hiệu âm thanh để thu được ước tính tần số cơ bản của các vị trí đỉnh 235 trong tín hiệu âm thanh. Hơn nữa, bộ trích xuất tham số 165 trích xuất các tham số 190 của các băng phụ của tín hiệu âm thanh 55 không được nằm trong tín hiệu âm thanh được mã hóa lõi 145 bộ tạo tín hiệu đầu ra 170 tạo ra tín hiệu đầu ra 135 bao gồm tín hiệu âm thanh được mã hóa lõi 145, các tham số 190, tần số cơ bản của các vị trí đỉnh 235, và một trong số các vị trí đỉnh 230. Theo các phương án, bộ tạo tín hiệu đầu ra 170 được tạo cấu hình để tạo tín hiệu đầu ra 135 thành chuỗi các khung, trong đó mỗi khung bao gồm tín hiệu âm thanh được mã hóa lõi 145, các tham số 190, và trong đó chỉ mỗi khung thứ n bao gồm ước tính tần số cơ bản của các vị trí đỉnh 235 và vị trí đỉnh 230, trong đó $n \geq 2$.

Fig.33 thể hiện phương án của tín hiệu âm thanh 135 bao gồm tín hiệu âm thanh được mã hóa lõi 145 bao gồm số lượng các băng phụ giảm so với tín hiệu âm thanh ban đầu 55, tham số 190 biểu diễn các băng phụ của tín hiệu âm thanh không nằm trong tín hiệu âm thanh được mã hóa lõi, ước tính tần số cơ bản của các vị trí đỉnh 235, và ước tính vị trí đỉnh 230 của tín hiệu âm thanh 55. Ngoài ra, tín hiệu âm thanh 135 được tạo ra thành chuỗi các khung, trong đó mỗi khung bao gồm tín hiệu âm thanh được mã hóa lõi 145, các tham số 190, và trong đó chỉ mỗi khung thứ n bao gồm ước tính tần số cơ bản của các vị trí đỉnh 235 và vị trí đỉnh 230, trong đó $n \geq 2$. Ý tưởng đã được mô tả đối với Fig.22.

Fig.34 thể hiện phương pháp 3400 để xử lý tín hiệu âm thanh với bộ xử lý âm thanh. Phương pháp 3400 bao gồm bước 3405 "xác định số đo pha đích cho tín hiệu âm thanh trong khung thời gian với số đo pha đích", bước 3410 "tính toán sai số pha

với bộ tính toán sai số pha sử dụng pha của tín hiệu âm thanh trong khung thời gian và số đo pha đích", và bước 3415 "hiệu chỉnh pha của tín hiệu âm thanh trong khung thời gian với pha được hiệu chỉnh sử dụng sai số pha".

Fig.35 thể hiện phương pháp 3500 để giải mã tín hiệu âm thanh với bộ giải mã. Phương pháp 3500 bao gồm bước 3505 "giải mã tín hiệu âm thanh trong khung thời gian của băng cơ sở với bộ giải mã lõi", bước 3510 "ráp nối tập hợp các băng phụ của băng cơ sở được giải mã với bộ ráp nối, trong đó tập hợp các băng phụ tạo ra phần ráp nối, với các băng phụ khác trong khung thời gian, liền kề với băng cơ sở, để thu được tín hiệu âm thanh bao gồm các tần số cao hơn các tần số trong băng cơ sở", và bước 3515 "hiệu chỉnh các pha với các băng phụ của phần ráp nối thứ nhất với bộ xử lý âm thanh theo số đo pha đích".

Fig.36 thể hiện phương pháp 3600 để mã hóa tín hiệu âm thanh với bộ mã hóa. Phương pháp 3600 bao gồm bước 3605 "mã hóa lõi tín hiệu âm thanh với bộ mã hóa lõi để thu được tín hiệu âm thanh được mã hóa lõi có số lượng các băng phụ giảm đối với tín hiệu âm thanh", bước 3610 "phân tích tín hiệu âm thanh hoặc phiên bản được lọc thông thấp của tín hiệu âm thanh với bộ phân tích tần số cơ bản để thu được ước tính tần số cơ bản của các vị trí đỉnh trong tín hiệu âm thanh", bước 3615 "trích xuất các tham số của các băng phụ của tín hiệu âm thanh không nằm trong tín hiệu âm thanh được mã hóa lõi với bộ trích xuất tham số", và bước 3620 "tạo ra tín hiệu đầu ra với bộ tạo tín hiệu đầu ra bao gồm tín hiệu âm thanh được mã hóa lõi, các tham số, tần số cơ bản của các vị trí đỉnh, và vị trí đỉnh".

Nói cách khác, thuật toán được gợi ý để hiệu chỉnh các sai số trong các vị trí theo thời gian của các hàm sóng hài hoạt động như sau. Đầu tiên, sự chênh lệch giữa phổ pha của tín hiệu đích và tín hiệu được xử lý SBR ($Z_{tv}^{pha}(k, n)$ và Z^{pha}) được tính toán:

$$D^{pha}(k, n) = Z^{pha}(k, n) - Z_{tv}^{pha}(k, n), \quad (20a)$$

mà được mô tả trên Fig.37. Fig.37 thể hiện sai số trong phổ pha $D^{pha}(k, n)$ của tín hiệu kèn trombon trong miền QMF sử dụng SBR sao chép trực tiếp. Tại mức này, phổ pha đích có thể được giả thiết là bằng phổ pha đích của tín hiệu đầu vào:

$$Z_{tv}^{pha}(k, n) = X^{pha}(k, n) \quad (20b)$$

Sau đó, nó sẽ được biểu diễn làm thế nào phổ pha đích có thể thu được với tốc độ bit thấp.

Sự hiệu chỉnh đạo hàm pha dọc được thực hiện sử dụng hai phương pháp, và phổ pha được hiệu chỉnh cuối cùng thu được như là hỗn hợp của chúng.

Đầu tiên, có thể thấy rằng sai số mà bất biến tương đối bên trong phần rập nối tần số, và sai số nhảy đến giá trị mới khi đi vào phần rập nối tần số mới. Điều này hợp lý, vì pha thay đổi với giá trị bất biến theo thời gian tại tất cả các tần số trong tín hiệu ban đầu. Sai số được tạo ra tại điểm giao chéo và sai số duy trì bất biến bên trong phần rập nối. Do đó, giá trị đơn là đủ để hiệu chỉnh sai số pha cho toàn bộ phần rập nối tần số. Hơn nữa, sai số pha của các phần rập nối tần số cao hơn có thể được hiệu chỉnh sử dụng giá trị sai số giống nhau sau phép nhân với chỉ số của phần rập nối tần số.

Do đó, trung bình lặp của sai số pha được tính toán cho phần rập nối tần số thứ nhất:

$$D_{avg}^{pha}(n) = \text{circmean}\{D^{pha}(k, n)\}, 8 \leq k \leq 13. \quad (21)$$

Phổ pha có thể được hiệu chỉnh sử dụng nó:

$$Y_{cv1}^{pha}(k, n, i) = Y^{pha}(k, n, i) - i \cdot D_{avg}^{pha}(n). \quad (22)$$

Sự hiệu chỉnh thô này tạo ra kết quả chính xác nếu PDF đích, tức là, đạo hàm pha theo tần số $X^{pdf}(k, n)$, bất biến một cách chính xác tại tất cả các tần số. Tuy nhiên, như có thể thấy trên Fig.12, thường có sự biến thiên nhẹ theo tần số trong giá trị. Do đó, các kết quả tốt hơn có thể thu được bằng cách sử dụng việc xử lý nâng cao tại các điểm giao chéo để tránh bất kì sự gián đoạn nào trong PDF được tạo ra. Nói cách khác, sự hiệu chỉnh này tạo ra các giá trị đúng cho PDF về trung bình, nhưng có thể có các sự gián đoạn nhẹ tại các tần số giao chéo của các phần rập nối tần số. Để tránh khỏi chúng, phương pháp hiệu chỉnh được áp dụng. Phổ pha được hiệu chỉnh cuối cùng $Y_{cv}^{pha}(k, n, i)$ thu được như hỗn hợp của hai phương pháp hiệu chỉnh.

Phương pháp hiệu chỉnh khác bắt đầu bằng việc tính toán trung bình của PDF trong băng cơ sở:

$$X_{\text{avg}}^{\text{pdf}}(n) = \text{circmean}\{X_{\text{base}}^{\text{pdf}}(k, n)\}. \quad (23)$$

Phổ pha có thể được hiệu chỉnh sử dụng số đo này bằng cách giả định rằng pha thay đổi với giá trị trung bình này, tức là:

$$\begin{aligned} Y_{\text{cv2}}^{\text{pha}}(k, n, 1) &= X_{\text{base}}^{\text{pha}}(6, n) + k \cdot X_{\text{avg}}^{\text{pdf}}(n), \\ Y_{\text{cv2}}^{\text{pha}}(k, n, i) &= Y_{\text{cv}}^{\text{pha}}(6, n, i - 1) + k \cdot X_{\text{avg}}^{\text{pdf}}(n), \end{aligned} \quad (24)$$

trong đó $Y_{\text{cv}}^{\text{pha}}$ là tín hiệu phần ráp nối được tổ hợp của hai phương pháp hiệu chỉnh.

Sự hiệu chỉnh cung cấp chất lượng tốt tại các điểm giao chéo, nhưng có thể gây ra sự lệch hướng trong PDF về phía các tần số cao hơn. Để tránh điều này, hai phương pháp hiệu chỉnh được tổ hợp bằng cách tính toán trung bình lặp được gán trọng số của chúng:

$$Y_{\text{cv}}^{\text{pha}}(k, n, i) = \text{circmean}\{Y_{\text{cv12}}^{\text{pha}}(k, n, i, c), W_{\text{fc}}(k, c)\}, \quad (25)$$

trong đó c biểu thị phương pháp hiệu chỉnh ($Y_{\text{cv1}}^{\text{pha}}$ hoặc $Y_{\text{cv2}}^{\text{pha}}$ và $W_{\text{fc}}(k, c)$ là hàm gán trọng số:

$$\begin{aligned} W_{\text{fc}}(k, 1) &= [0,2, \ 0,45, \ 0,7, \ 1,1,1], \\ W_{\text{fc}}(k, 2) &= [0,8, \ 0,55, \ 0,3, \ 0,0,0]. \end{aligned} \quad (26a)$$

Phổ pha thu được $Y_{\text{cv}}^{\text{pha}}(k, n, i)$ không phải chịu sự gián đoạn hoặc lệch hướng. Sai số so với phổ ban đầu và PDF của phổ pha được hiệu chỉnh được mô tả trên Fig.38. Fig.38a thể hiện sai số trong phổ pha $D_{\text{cv}}^{\text{pha}}(k, n)$ của tín hiệu kèn trombon trong miền QMF sử dụng tín hiệu SBR được hiệu chỉnh pha, trong đó Fig.38b thể hiện đạo hàm pha theo tần số tương ứng $Z_{\text{cv}}^{\text{pdf}}(k, n)$. Có thể thấy rằng sai số là nhỏ hơn đáng kể so với không có sự hiệu chỉnh, và PDF không phải chịu sự gián đoạn lớn. Có các sai số đáng kể tại các khung thời gian nhất định, nhưng các khung này có năng lượng thấp (xem Fig.4), để chúng có hiệu quả cảm quan không đáng kể. Các khung theo thời gian với năng lượng đáng kể được hiệu chỉnh tương đối tốt. Có thể lưu ý rằng các thành phần lạ của SBR được giảm nhẹ đáng kể.

Phổ pha được hiệu chỉnh $Z_{cv}^{pha}(k, n)$ thu được bằng cách kết nối các phần ráp nối tần số được hiệu chỉnh $Y_{cv}^{pha}(k, n, i)$. Để tương thích với chế độ hiệu chỉnh ngang, sự hiệu chỉnh pha dọc có thể được thể hiện cũng sử dụng ma trận điều biến (xem phương trình 18):

$$Q^{pha}(k, n) = Z_{cv}^{pha}(k, n) - Z^{pha}(k, n). \quad (26b)$$

8.3 Chuyển đổi giữa các phương pháp hiệu chỉnh pha khác nhau

Các phần 8.1 và 8.2 đã thể hiện rằng các sai số pha được dẫn SBR có thể được hiệu chỉnh bằng cách sử dụng hiệu chỉnh PDT cho đàn viôlông và hiệu chỉnh PDF cho kèn trombon. Tuy nhiên, nó không xem xét làm thế nào để biết hiệu chỉnh nào trong số các hiệu chỉnh nên được áp dụng cho tín hiệu không biết, hoặc nếu bất kỳ trong số chúng nên được áp dụng. Phần này đề xuất phương pháp để tự động lựa chọn hướng hiệu chỉnh. Hướng hiệu chỉnh (ngang/dọc) được quyết định dựa trên sự biến thiên của các đạo hàm pha của tín hiệu đầu vào.

Do đó, trên Fig.39, bộ tính toán để xác định dữ liệu hiệu chỉnh pha cho tín hiệu âm thanh 55 được thể hiện. Bộ xác định sự biến thiên 275 xác định độ biến thiên của pha 45 của tín hiệu âm thanh 55 trong chế độ biến thiên thứ nhất và thứ hai. Bộ so sánh biến thiên 280 so sánh độ biến thiên thứ nhất 290a được xác định sử dụng chế độ biến thiên thứ nhất và chế độ biến thiên thứ hai 290b được xác định sử dụng chế độ biến thiên thứ hai và bộ tính toán dữ liệu hiệu chỉnh tính toán dữ liệu hiệu chỉnh pha 295 phù hợp với chế độ biến thiên thứ nhất hoặc chế độ biến thiên thứ hai dựa trên kết quả của bộ so sánh.

Hơn nữa, bộ xác định độ biến thiên 275 có thể được tạo cấu hình để xác định số đo độ lệch tiêu chuẩn của đạo hàm pha theo thời gian (phase derivative over time - PDT) cho nhiều khung thời gian của tín hiệu âm thanh 55 như độ biến thiên 290a của pha trong chế độ biến thiên thứ nhất và để xác định số đo độ lệch tiêu chuẩn của pha đạo hàm pha theo tần số (phase derivative over frequency - PDF) cho nhiều băng phụ của tín hiệu âm thanh 55 như độ biến thiên 290b của pha trong chế độ biến thiên thứ hai. Do đó, bộ so sánh độ biến thiên 280 so sánh số đo của đạo hàm pha theo thời gian như độ biến thiên thứ nhất 290a và số đo của đạo hàm pha theo tần số như độ biến

thiên thứ hai 290b cho các khung thời gian của tín hiệu âm thanh.

Các phương án thể hiện bộ xác định độ biến thiên 275 để xác định độ lệch tiêu chuẩn lặp của đạo hàm pha theo thời gian của khung hiện thời và nhiều khung trước đó của tín hiệu âm thanh 55 như số đo độ lệch tiêu chuẩn và để xác định độ lệch tiêu chuẩn lặp của đạo hàm pha theo thời gian của khung hiện thời và nhiều khung tương lai của tín hiệu âm thanh 55 cho khung thời gian hiện thời như số đo độ lệch tiêu chuẩn. Hơn nữa, bộ xác định độ biến thiên 275 tính toán, khi xác định độ biến thiên thứ nhất 290a, giá trị tối thiểu của cả hai độ lệch tiêu chuẩn lặp. Trong phương án khác, bộ xác định độ biến thiên 275 tính toán độ biến thiên 290a trong chế độ biến thiên thứ nhất như tổ hợp của số đo độ lệch tiêu chuẩn cho nhiều băng phụ 95 trong khung thời gian 75 để tạo ra số đo độ lệch tiêu chuẩn trung bình theo tần số. Bộ so sánh độ lệch 280 được tạo cấu hình để thực hiện sự tổ hợp của các số đo độ lệch tiêu chuẩn bằng cách tính toán trung bình được gán trọng số năng lượng của các số đo độ lệch tiêu chuẩn của nhiều băng phụ sử dụng các giá trị cường độ của tín hiệu băng phụ 95 trong khung thời gian hiện thời như số đo năng lượng.

Trong phương án được ưu tiên, bộ xác định độ lệch 275 làm nhãn số đo độ lệch tiêu chuẩn trung bình, khi xác định độ lệch thứ nhất 290a, theo khung thời gian hiện thời, nhiều khung thời gian trước đó và nhiều khung thời gian tương lai. Sự làm nhãn như được gán trọng số theo năng lượng được tính toán sử dụng các khung thời gian tương ứng và hàm tạo cửa sổ. Hơn nữa, bộ xác định độ biến thiên 275 được tạo cấu hình để làm nhãn số đo độ lệch tiêu chuẩn, khi xác định độ biến thiên thứ hai 290b trên khung thời gian hiện thời, nhiều khung thời gian trước, và nhiều khung thời gian tương lai 75, trong đó việc làm nhãn được gán trọng số theo năng lượng được tính toán sử dụng các khung thời gian tương ứng 75 và hàm tạo cửa sổ. Do đó, bộ so sánh độ biến thiên 280 so sánh số đo độ lệch tiêu chuẩn trung bình như độ biến thiên thứ nhất 290a được xác định sử dụng chế độ biến thiên thứ nhất và so sánh số đo độ lệch tiêu chuẩn như độ biến thiên thứ hai 290b được xác định sử dụng chế độ biến thiên thứ hai.

Phương án được ưu tiên được mô tả trên Fig.40. Theo phương án này, bộ xác định độ biến thiên 275 bao gồm hai đường xử lý để tính toán độ biến thiên thứ nhất và thứ hai. Phần ráp nối xử lý thứ nhất bao gồm bộ tính toán PDT 300a, để tính toán số

đo độ lệch tiêu chuẩn của đạo hàm pha theo thời gian 305a từ tín hiệu âm thanh 55 hoặc pha của tín hiệu âm thanh. Bộ tính toán độ lệch tiêu chuẩn lặp 310a xác định độ lệch tiêu chuẩn lặp thứ nhất 315a và độ lệch tiêu chuẩn lặp thứ hai 315b từ số đo độ lệch tiêu chuẩn của đạo hàm pha theo thời gian 305a. Các độ lệch tiêu chuẩn lặp thứ nhất và thứ hai 315a và 315b được so sánh bởi bộ so sánh 320. Bộ so sánh 320 tính toán giá trị tối thiểu 325 của hai số đo độ lệch tiêu chuẩn lặp 315a và 315b. Bộ tổ hợp tổ hợp giá trị tối thiểu 325 theo tần số để tạo ra số đo độ lệch tiêu chuẩn trung bình 335a. Bộ làm nhẵn 340a làm nhẵn các số đo độ lệch tiêu chuẩn trung bình để tạo ra số đo độ lệch tiêu chuẩn trung bình nhẵn 345a.

Đường xử lý thứ hai bao gồm bộ tính toán PDF 300b để tính toán đạo hàm pha theo tần số 305b từ tín hiệu âm thanh 55 hoặc pha của tín hiệu âm thanh. Bộ tính toán đạo hàm tiêu chuẩn lặp 310b tạo ra các số đo độ lệch tiêu chuẩn 335b của đạo hàm pha theo tần số 305. Số đo độ lệch tiêu chuẩn 305 được làm nhẵn bởi bộ làm nhẵn 340b để tạo ra số đo độ lệch tiêu chuẩn nhẵn 345b. Các số đo độ lệch tiêu chuẩn trung bình được làm nhẵn 345a và số đo độ lệch tiêu chuẩn được làm nhẵn 345b lần lượt là độ biến thiên thứ nhất và thứ hai. Bộ so sánh độ biến thiên 280 so sánh độ biến thiên thứ nhất và thứ hai và bộ tính toán dữ liệu hiệu chỉnh 285 tính toán dữ liệu hiệu chỉnh pha 295 dựa trên việc so sánh độ biến thiên thứ nhất và thứ hai.

Các phương án khác thể hiện bộ tính toán 270 xử lý ba chế độ hiệu chỉnh pha khác nhau. Sơ đồ khối tương trưng được thể hiện trên Fig.41. Fig.41 thể hiện bộ xác định độ biến thiên 275 còn xác định độ biến thiên thứ ba 290c của pha của tín hiệu âm thanh 55 trong chế độ biến thiên thứ ba, trong đó chế độ biến thiên thứ ba là chế độ phát hiện chuyển tiếp. Bộ so sánh độ biến thiên 280 so sánh độ biến thiên thứ nhất 290a, được xác định sử dụng chế độ biến thiên thứ nhất, độ biến thiên thứ hai 290b, được xác định sử dụng chế độ biến thiên thứ hai, và độ biến thiên thứ ba 290c, được xác định sử dụng chế độ biến thiên thứ ba. Do đó, bộ tính toán dữ liệu hiệu chỉnh 285 tính toán dữ liệu hiệu chỉnh pha 295 phù hợp với chế độ hiệu chỉnh thứ nhất, chế độ hiệu chỉnh thứ hai, hoặc chế độ hiệu chỉnh thứ ba, dựa trên kết quả của sự so sánh. Để tính toán độ biến thiên thứ ba 290c trong chế độ biến thiên thứ ba, bộ so sánh độ biến thiên 280 có thể được tạo cấu hình để tính toán ước tính năng lượng tức thời của khung thời gian hiện thời và ước tính năng lượng trung bình theo thời gian của nhiều khung

thời gian 75. Do đó, bộ so sánh độ biến thiên 280 được tạo cấu hình để tính toán tỉ lệ của ước tính năng lượng tức thời và ước tính năng lượng trung bình theo thời gian và được tạo cấu hình để so sánh tỉ lệ với ngưỡng được xác định để phát hiện sự chuyển tiếp trong khung thời gian 75.

Bộ so sánh độ biến thiên 280 phải xác định chế độ hiệu chỉnh thích hợp dựa trên ba độ biến thiên. Dựa trên quyết định này, bộ tính toán dữ liệu hiệu chỉnh 285 tính toán dữ liệu hiệu chỉnh pha 295 phù hợp với chế độ biến thiên thứ ba nếu sự chuyển tiếp được phát hiện. Hơn nữa, bộ tính toán dữ liệu hiệu chỉnh 85 tính toán dữ liệu hiệu chỉnh pha 295 phù hợp với chế độ biến thiên thứ nhất, nếu sự vắng mặt của sự chuyển tiếp được phát hiện và nếu độ biến thiên thứ nhất 290a được xác định trong chế độ biến thiên thứ nhất, nhỏ hơn hoặc bằng độ biến thiên thứ hai 290b, được xác định trong chế độ biến thiên thứ hai. Theo đó, dữ liệu hiệu chỉnh pha 295 được tính toán phù hợp với chế độ biến thiên thứ hai, nếu sự không có mặt của sự chuyển tiếp được phát hiện và nếu độ biến thiên thứ hai 290b, được xác định trong chế độ biến thiên thứ hai, là nhỏ hơn độ biến thiên thứ nhất 290a, được xác định bởi chế độ biến thiên thứ nhất.

Bộ tính toán dữ liệu hiệu chỉnh còn được tạo cấu hình để tính toán dữ liệu hiệu chỉnh pha 295 cho độ biến thiên thứ ba 290c cho khung thời gian hiện thời, một hoặc nhiều khung thời gian trước đó và một hoặc nhiều khung thời gian tương lai. Theo đó, bộ tính toán dữ liệu hiệu chỉnh 285 được tạo cấu hình để tính toán dữ liệu hiệu chỉnh pha 295 cho độ biến thiên thứ hai 290b cho khung thời gian hiện thời, một hoặc nhiều khung thời gian trước đó và một hoặc nhiều khung thời gian tương lai. Mặt khác, bộ tính toán dữ liệu hiệu chỉnh 285 được tạo cấu hình để tính toán dữ liệu hiệu chỉnh 295 cho sự hiệu chỉnh pha ngang và chế độ hiệu chỉnh thứ nhất, tính toán dữ liệu hiệu chỉnh 295 cho sự hiệu chỉnh pha dọc trong chế độ biến thiên thứ hai, và tính toán dữ liệu hiệu chỉnh 295 cho sự hiệu chỉnh chuyển tiếp trong chế độ biến thiên thứ ba.

Fig.42 thể hiện phương pháp 4200 để xác định dữ liệu hiệu chỉnh pha từ tín hiệu âm thanh. Phương pháp 4200 bao gồm bước 4205 "xác định độ biến thiên của pha của tín hiệu âm thanh với bộ xác định độ biến thiên trong chế độ thứ nhất và thứ hai", bước 4210 "so sánh độ biến thiên được xác định sử dụng chế độ biến thiên thứ nhất và

thứ hai với bộ so sánh độ biến thiên", và bước 4215 "tính toán sự hiệu chỉnh pha với bộ tính toán dữ liệu hiệu chỉnh phù hợp với chế độ biến thiên thứ nhất hoặc chế độ biến thiên thứ hai trên cơ sở kết quả của sự so sánh".

Nói cách khác, PDT của đàn viôlông là nhẵn theo thời gian trong khi PDF của kèn trombon là nhẵn theo tần số. Do đó, độ lệch tiêu chuẩn (standard deviation - STD) của các số đo này như số đo của độ lệch có thể được sử dụng để lựa chọn phương pháp hiệu chỉnh thích hợp. STD của đạo hàm pha theo thời gian có thể được tính toán như:

$$\begin{aligned} X^{\text{stdt1}}(k, n) &= \text{circstd}\{X^{\text{pdt}}(k, n + l)\}, -23 \leq l \leq 0, \\ X^{\text{stdt2}}(k, n) &= \text{circstd}\{X^{\text{pdt}}(k, n + l)\}, 0 \leq l \leq 23, \\ X^{\text{stdt}}(k, n) &= \min\{X^{\text{stdt1}}(k, n), X^{\text{stdt2}}(k, n)\}, \end{aligned} \quad (27)$$

và STD của đạo hàm pha theo tần số như:

$$X^{\text{stdf}}(n) = \text{circstd}\{X^{\text{pdf}}(k, n)\}, 2 \leq k \leq 13, \quad (28)$$

trong đó $\text{circstd}\{\}$ biểu thị việc tính toán STD lặp (các giá trị góc có thể có khả năng được gán trọng số bởi năng lượng để tránh STD cao do các ngăn năng lượng thấp nhiều, hoặc sự tính toán STD có thể được giới hạn đến các ô với năng lượng đủ). Các STD cho đàn viôlông và kèn trombon được thể hiện lần lượt trên các hình vẽ Fig.43a, Fig.43b và Fig.43c, Fig.43d. Các hình vẽ Fig.43a và Fig.43c thể hiện đạo hàm pha theo thời gian $X^{\text{stdt}}(k, n)$ trong miền QMF, trong đó các hình vẽ Fig.43b và Fig.43d thể hiện độ lệch tiêu chuẩn theo tần số $X^{\text{stdf}}(n)$ mà không có sự hiệu chỉnh pha. Gradient màu biểu thị các giá trị từ đỏ = 1 đến xanh da trời = 0. Có thể nhận ra rằng STD và PDT thấp hơn cho đàn viôlông trong đó STD và PDF thấp hơn đối với kèn trombon (đặc biệt là đối với các ô mà có năng lượng cao).

Phương pháp hiệu chỉnh được sử dụng đối với mỗi khung theo thời gian được lựa chọn trên cơ sở của các STD là thấp. Đối với điều đó, các giá trị $X^{\text{stdt}}(k, n)$ phải được tổ hợp theo tần số. Trình kết hợp được thực hiện bằng cách tính toán trung bình được gán trọng số năng lượng đối với phạm vi tần số được định trước:

$$X^{\text{stdt}}(k, n) = \frac{\sum_{k=2}^{19} X^{\text{stdt}}(k, n) X^{\text{mag}}(k, n)}{\sum_{k=2}^{19} X^{\text{mag}}(k, n)}. \quad (29)$$

Ước tính độ lệch được làm nhẵn theo thời gian để có sự chuyển nhẵn, và do đó tránh được các thành phần lạ tiềm tàng. Việc làm nhẵn được thực hiện sử dụng cửa sổ Hann và được gán trọng số bằng năng lượng của khung theo thời gian:

$$X_{\text{sm}}^{\text{stdt}}(n) = \frac{\sum_{l=-10}^{10} X^{\text{stdt}}(n+l) X^{\text{mag}}(n+l) W(l)}{\sum_{l=-10}^{10} X^{\text{mag}}(n+l) W(l)}, \quad (30)$$

trong đó $W(l)$ là hàm cửa sổ và $X^{\text{mag}}(n) = \sum_{k=1}^{64} X^{\text{mag}}(k, n)$ là tổng của $X^{\text{mag}}(k, n)$ theo tần số. Phương trình tương ứng được sử dụng để làm nhẵn $X^{\text{stdf}}(n)$.

Phương pháp hiệu chỉnh pha được xác định bằng cách so sánh $X_{\text{sm}}^{\text{stdt}}(n)$ và $X_{\text{sm}}^{\text{stdf}}(n)$. Phương pháp mặc định là hiệu chỉnh (ngang) PDT, và nếu $X_{\text{sm}}^{\text{stdf}}(n) < X_{\text{sm}}^{\text{stdt}}(n)$, hiệu chỉnh (dọc) PDF được áp dụng cho khoảng $[n-5, n+5]$. Nếu cả hai độ lệch đều lớn, ví dụ, lớn hơn giá trị ngưỡng được định trước, hoặc của các phương pháp hiệu chỉnh được áp dụng, và có thể thực hiện việc tiết kiệm tốc độ bit.

8.4 Việc xử lý chuyển tiếp - Sự hiệu chỉnh đạo hàm pha cho sự chuyển tiếp

Tín hiệu đàn viôlông với tiếng vĩ tay được thêm vào giữa được thể hiện trên Fig.44. Cường độ $X^{\text{mag}}(k, n)$ của tín hiệu đàn viôlông + tiếng vĩ tay trong miền QMF được thể hiện trên Fig.44a, và phổ pha tương ứng $X^{\text{pha}}(k, n)$ trên Fig.44b. Xét Fig.44a, gradien màu biểu thị các giá trị từ đỏ = 0 dB đến xanh da trời = -80 dB. Theo đó, đối với Fig.44b, gradien pha biểu thị các giá trị pha từ đỏ = π đến xanh da trời = $-\pi$. Các đạo hàm pha theo thời gian và theo tần số được biểu diễn trên Fig.45. Đạo hàm pha theo thời gian $X^{\text{pdt}}(k, n)$ của tín hiệu đàn viôlông + tiếng vĩ tay trong miền QMF được thể hiện trên Fig.45a, và đạo hàm pha theo tần số $X^{\text{pdf}}(k, n)$ tương ứng trên Fig.45b. Gradien màu biểu thị các giá trị pha từ đỏ = π đến xanh da trời = $-\pi$. Có thể thấy rằng PDT là nhiều đối với tiếng vĩ tay, nhưng PDF nhẵn hơn một chút, ít nhất tại các tần số cao. Do đó, sự hiệu chỉnh PDF nên được áp dụng cho tiếng vĩ tay để duy trì độ nét của nó. Tuy nhiên, phương pháp hiệu chỉnh được gợi ý trong phần 8.2 có thể

không hoạt động chính xác với tín hiệu này, vì âm đàn viôlông nhiễu loạn các đạo hàm tại các tần số thấp. Kết quả là, phổ pha của băng cơ sở không phản xạ các tần số cao, và do đó, sự hiệu chỉnh pha của các phần ráp nối tần số sử dụng giá trị đơn có thể không hoạt động. Hơn nữa, việc phát hiện các chuyển tiếp trên cơ sở độ biến thiên của giá trị PDF (xem Phần 8.3) sẽ khó do các giá trị PDF nhiễu tại các tần số thấp.

Giải pháp cho vấn đề rất dễ dàng. Thứ nhất, các chuyển tiếp được phát hiện sử dụng phương pháp trên cơ sở năng lượng đơn giản. Năng lượng tức thời của các tần số trung/cao được so sánh với ước tính năng lượng được làm nhẵn. Năng lượng tức thời của các tần số trung/cao được tính toán như:

$$X^{\text{magmh}}(n) = \sum_{k=6}^{64} X^{\text{mag}}(k, n). \quad (31)$$

Sự làm nhẵn được thực hiện sử dụng bộ lọc IIR thứ tự thứ nhất:

$$X_{\text{sm}}^{\text{magmh}}(n) = 0,1 \cdot X^{\text{magmh}}(n) + 0,9 \cdot X_{\text{sm}}^{\text{magmh}}(n-1). \quad (32)$$

Nếu $X^{\text{magmh}}(n)/X_{\text{sm}}^{\text{magmh}}(n) > \theta$, sự chuyển tiếp đã được phát hiện. Ngưỡng θ có thể được tinh chỉnh để phát hiện lượng chuyển tiếp mong muốn. Ví dụ, $\theta = 2$ có thể được sử dụng. Khung được phát hiện không được lựa chọn trực tiếp để là khung chuyển tiếp. Thay vào đó, giá trị tối đa năng lượng cục bộ được tìm kiếm từ âm thanh nổi của nó. Trong việc thực hiện hiện thời, khoảng được lựa chọn là $[n-2, n+7]$. Khung theo thời gian với năng lượng tối đa bên trong khoảng này được lựa chọn để là sự chuyển tiếp.

Về lý thuyết, chế độ hiệu chỉnh dọc có thể cũng được áp dụng cho sự chuyển tiếp. Tuy nhiên, trong trường hợp chuyển tiếp, phổ pha của băng cơ sở thường không phản xạ các tần số cao. Điều này có thể dẫn đến các tiếng vang trước hoặc sau trong tín hiệu được xử lý. Do đó, việc xử lý cải biên nhẹ được gợi ý cho các chuyển tiếp.

PDF trung bình của sự chuyển tiếp tại các tần số cao được tính toán:

$$X_{\text{avghi}}^{\text{pdf}}(n) = \text{circmean}\{X^{\text{pdf}}(k, n)\}, -11 \leq k \leq 36. \quad (33)$$

Phổ pha cho khung chuyển tiếp được tổng hợp sử dụng sự thay đổi pha bất biến

như phương trình 24, nhưng $X_{\text{avg}}^{\text{pdf}}(n)$ được thay thế bởi $X_{\text{avghi}}^{\text{pdf}}(n)$. Sự hiệu chỉnh giống nhau được áp dụng cho các khung theo thời gian nằm trong khoảng $[n - 2, n + 2]$ (π được cộng vào PDF của các khung $n - 1$ và $n + 1$ do các đặc tính của QMF, xem phần 6). Sự hiệu chỉnh này đã đưa ra sự chuyển tiếp đến vị trí thích hợp, nhưng hình dạng của sự chuyển tiếp không cần nhiệt như mong muốn, và các thùy bên đáng kể (tức là các chuyển tiếp bổ sung) có thể được có mặt do sự chồng lấp theo thời gian đáng kể của các khung QMF. Do đó, góc pha tuyệt đối cũng phải chính xác. Góc tuyệt đối được hiệu chỉnh bằng cách tính toán sai số trung bình giữa phổ pha được tổng hợp và phổ pha ban đầu. Sự hiệu chỉnh được thực hiện riêng lẻ cho mỗi khung theo thời gian của sự chuyển tiếp.

Kết quả của sự hiệu chỉnh chuyển tiếp được thể hiện trên Fig.46. Đạo hàm pha theo thời gian $X^{\text{pdt}}(k, n)$ của tín hiệu đàn viôlông + tiếng vỗ tay trong miền QMF sử dụng SBR được hiệu chỉnh pha được thể hiện. Fig.47b thể hiện đạo hàm pha theo tần số $X^{\text{pdf}}(k, n)$. Một lần nữa, gradien màu biểu thị các giá trị pha từ đỏ = π đến xanh da trời = $-\pi$. Có thể nhận biết rằng tiếng vỗ tay được hiệu chỉnh pha có chung độ nét như tín hiệu ban đầu, mặc dù sự chênh lệch được so sánh với sự sao chép trực tiếp không lớn. Do đó, sự hiệu chỉnh chuyển tiếp không nhất thiết được yêu cầu trong tất cả các trường hợp chỉ khi sự sao chép trực tiếp được kích hoạt. Ngược lại, nếu sự hiệu chỉnh PDT được kích hoạt, sẽ là quan trọng để xử lý sự chuyển tiếp, như sự hiệu chỉnh PDT nếu không sẽ làm nhòe các chuyển tiếp một cách rất xấu.

9. Sự nén của dữ liệu hiệu chỉnh

Phần 8 thể hiện các sai số pha có thể được hiệu chỉnh, như tốc độ bit thích hợp cho việc hiệu chỉnh không được xem xét. Phần này đề xuất các phương pháp làm thế nào để biểu diễn dữ liệu hiệu chỉnh với tốc độ bit thấp.

9.1 Nén dữ liệu hiệu chỉnh PTD - Tạo phổ đích cho sự hiệu chỉnh ngang

Có nhiều tham số khả thi mà có thể được truyền để kích hoạt sự hiệu chỉnh PDT. Tuy nhiên, do $D_{\text{sm}}^{\text{pdt}}(k, n)$ được làm phẳng theo thời gian, nó là ứng cử tiềm năng cho sự truyền tốc độ bit thấp.

Thứ nhất, tốc độ cập nhật thích hợp cho các tham số được thảo luận. Giá trị chỉ

được cập nhật cho mỗi N khung và được nội suy tuyến tính ở giữa. Khoảng cập nhật cho chất lượng tốt là khoảng 40ms. Đối với các tín hiệu nhất định ít bit hơn là có lợi và đối với các tín hiệu khác nhiều bit hơn là có lợi. Các bài kiểm tra nghe chính quy là hữu ích để đánh giá tốc độ cập nhật tối ưu. Tuy nhiên, khoảng cập nhật dài tương đối có vẻ chấp nhận được.

Độ chính xác góc thích hợp cho $D_{sm}^{pdt}(k, n)$ cũng được nghiên cứu. 6 bit (64 giá trị góc khả thi) là đủ cho chất lượng tốt về cảm quan. Ngoài ra, việc truyền chỉ sự thay đổi trong giá trị được kiểm tra. Các giá trị thường chỉ thay đổi một ít, do vậy lượng tử hóa không đều có thể được áp dụng để có độ chính xác hơn cho các thay đổi nhỏ. Sử dụng cách tiếp cận này, 4 bit (16 giá trị góc khả thi) được phát hiện ra là cung cấp chất lượng tốt.

Điều cuối cùng cần xem xét là độ chính xác phổ thích hợp. Như có thể thấy trên Fig.17, nhiều băng dường như đại khái có cùng giá trị giống nhau. Do đó, một giá trị có thể được sử dụng để biểu diễn một vài băng. Ngoài ra, tại các tần số cao có nhiều sóng hài bên trong một băng, do đó có lẽ cần độ chính xác ít hơn. Tuy nhiên, cách tiếp cận khác có thể tốt hơn được tìm ra, do đó các tùy chọn này không được nghiên cứu một cách kỹ lưỡng. Cách tiếp cận được đề xuất hiệu quả hơn sẽ được thảo luận sau đây.

9.1.1 Sử dụng ước tính tần số để nén dữ liệu hiệu chỉnh PDT

Như được thảo luận trong phần 5, đạo hàm pha theo thời gian về cơ bản là tần số của đường hình sin được tạo ra. Các PDT của QMF phức 64 băng được áp dụng có thể được biến đổi thành các tần số sử dụng phương trình sau:

$$X^{freq}(k, n) = \frac{f_s}{64} \left\lfloor \frac{(k-1.5)}{2} \right\rfloor + \left(\left[\left(\frac{X^{pdt}(k, n)}{2\pi} \bmod 1 \right) + \frac{(-1)^k}{4} + \frac{1}{2} \right] \bmod 1 \right). \quad (34)$$

Các tần số được tạo ra bên trong khoảng $f_{inter}(k) = [f_c(k) - f_{BW}, f_c(k) + f_{BW}]$, trong đó $f_c(k)$ là tần số trung tâm của băng k và f_{BW} là 375 Hz. Kết quả được thể hiện trên Fig.47 trong phép biểu diễn thời gian - tần số của các tần số của các băng QMF $X^{freq}(k, n)$ cho tín hiệu đàn viôlông. Có thể thấy rằng các tần số dường như theo sau các phép nhân tần số cơ bản của âm và các sóng hài do đó được đặt cách trong tần

số bằng tần số cơ bản. Ngoài ra, tiếng rung đường như gây ra sự điều biến tần số.

Các đồ thị giống nhau có thể được áp dụng cho $Z^{\text{freq}}(k, n)$ sao chép trực tiếp và SBR $Z_{\text{ch}}^{\text{freq}}(k, n)$ được hiệu chỉnh (lần lượt xem Fig.48a và Fig.48b). Fig.48a thể hiện phép biểu diễn thời gian - tần số của các tần số của các băng QMF của tín hiệu SBR sao chép trực tiếp $Z^{\text{freq}}(k, n)$ so với tín hiệu ban đầu $X^{\text{freq}}(k, n)$, được thể hiện trên Fig.47. Fig.48 thể hiện đồ thị tương ứng cho tín hiệu SBR được hiệu chỉnh $Z_{\text{ch}}^{\text{freq}}(k, n)$. Trong các đồ thị trên Fig.48a và Fig.48b, tín hiệu ban đầu được vẽ bằng màu xanh da trời, trong đó các tín hiệu SBR sao chép trực tiếp và tín hiệu SBR được hiệu chỉnh được vẽ bằng màu đỏ. Sự phi điều hòa của SBR sao chép trực tiếp có thể được thấy trong hình vẽ, đặc biệt khi bắt đầu và kết thúc mẫu. Ngoài ra có thể thấy rằng độ sâu điều biến-tần số nhỏ hơn rõ ràng so với độ sâu của tín hiệu ban đầu. Ngược lại, trong trường hợp SBR được hiệu chỉnh, các tần số của các sóng hài dường như theo sau các tần số của tín hiệu ban đầu. Ngoài ra, độ sâu điều biến có vẻ chính xác. Do đó, đồ thị này dường như xác nhận sự hợp lệ của phương pháp hiệu chỉnh được đề xuất. Do đó, nó tập trung vào việc nén thực tế của dữ liệu hiệu chỉnh tiếp theo.

Do các tần số của $X^{\text{freq}}(k, n)$ được đặt cách với lượng giống nhau, các tần số này của tất cả các băng có thể được làm gần đúng nếu khoảng cách giữa các tần số được ước tính và được truyền. Trong trường hợp tín hiệu sóng hài, khoảng cách nên bằng với tần số cơ bản của âm. Do đó, chỉ giá trị đơn được truyền để biểu diễn tất cả các băng. Trong trường hợp nhiều tín hiệu không đều, cần nhiều giá trị để mô tả trạng thái sóng hài. Ví dụ, khoảng cách của sóng hài tăng nhẹ trong trường hợp âm đàn piano [14]. Đơn giản, nó giả định sau đây rằng các sóng hài được đặt cách với lượng giống nhau. Tuy nhiên, điều này không hạn chế tính tổng quát của việc xử lý âm thanh được mô tả.

Do đó, tần số cơ bản của âm được ước tính để ước tính các tần số của sóng hài. Ước tính tần số cơ bản là đề tài được nghiên cứu rộng rãi (ví dụ, xem [14]). Do đó, phương pháp ước tính đơn giản được thực hiện để tạo ra dữ liệu được sử dụng cho các bước xử lý tiếp theo. Phương pháp tính toán một cách cơ bản các khoảng cách của các sóng hài, và tổ hợp kết quả theo một số phương pháp giải quyết vấn đề bằng cách đánh giá kinh nghiệm (bao nhiêu năng lượng, độ ổn định của giá trị theo tần số và thời gian

như thế nào, v.v). Trong bất kỳ trường hợp nào, kết quả là ước tính tần số cơ bản cho mỗi khung thời gian $X^{f_0}(n)$. Nói cách khác, đạo hàm pha theo thời gian liên quan đến tần số của ngân QMF tương ứng. Ngoài ra, các thành phần lạ liên quan đến các sai số trong PDT có thể nhận biết hầu hết với các tín hiệu sóng hài. Do đó, đề xuất rằng PDT đích (xem phương trình 16a) có thể được ước tính sử dụng ước tính của tần số cơ bản f_0 . Ước tính của tần số cơ bản là đề tài được nghiên cứu rộng rãi, và có nhiều phương pháp mạnh có sẵn để thu được các ước tính đáng tin cậy của tần số cơ bản.

Tại đây, tần số cơ bản $X^{f_0}(n)$, như đã biết bộ giải mã trước khi thực hiện BWE và sử dụng sự hiệu chỉnh pha theo sáng chế trong BWE, được giả định. Do đó, sẽ là thuận lợi rằng giai đoạn mã hóa truyền tần số cơ bản được ước tính $X^{f_0}(n)$. Ngoài ra, đối với hiệu quả lập mã được cải thiện, giá trị có thể chỉ được cập nhật cho, ví dụ, mỗi khung thời gian thứ 20 (tương ứng với khoảng là -27ms) và được nội suy ở giữa.

Ngoài ra, tần số cơ bản có thể được ước tính trong giai đoạn giải mã, và không có thông tin nào được truyền. Tuy nhiên, các ước tính tốt hơn có thể được dự kiến nếu ước tính được thực hiện với tín hiệu ban đầu trong giai đoạn mã hóa.

Việc xử lý bộ giải mã bắt đầu bằng việc thu được ước tính tần số cơ bản $X^{f_0}(n)$ cho mỗi khung theo thời gian.

Các tần số của sóng hài có thể được thu được bằng việc nhân nó với véc tơ chỉ số:

$$\forall \kappa \in \mathbb{N} : X^{\text{harm}}(\kappa, n) = \kappa \cdot X^{f_0}(n) \quad (35)$$

Kết quả được mô tả trên Fig.49. Fig.49 thể hiện phép biểu diễn thời gian tần số của các tần số được ước tính của các sóng hài $X^{\text{harm}}(\kappa, n)$ so với các tần số của các băng QMF của tín hiệu ban đầu $X^{\text{freq}}(k, n)$. Một lần nữa, màu xanh da trời chỉ tín hiệu ban đầu và màu đỏ chỉ tín hiệu được ước tính. Các tần số của các sóng hài được ước tính khá khớp với tín hiệu ban đầu. Các tần số này có thể được cho là các tần số "được cho phép". Nếu thuật toán tạo ra các tần số này, sự phi điều hòa liên quan đến các thành phần lạ nên được tránh.

Các tham số được truyền của thuật toán là tần số cơ bản $X^{f_0}(n)$. Để hiệu quả mã hóa được cải thiện, giá trị chỉ được cập nhật cho mỗi khung thời gian thứ 20 (tức

là, mỗi 27 ms). Giá trị này có vẻ cung cấp chất lượng cảm quan tốt trên cơ sở việc nghe không chính thức. Tuy nhiên, các bài kiểm tra nghe chính thức là hữu ích để đánh giá giá trị tối ưu hơn tốc độ cập nhật.

Bước tiếp theo của thuật toán là tìm giá trị thích hợp cho mỗi băng số. Điều này được thực hiện bằng việc lựa chọn giá trị $X^{\text{harm}}(\kappa, n)$ mà gần nhất với tần số trung tâm của mỗi băng $f_c(k)$ để phản xạ băng đó. Nếu giá trị gần nhất nằm ngoài giá trị khả thi của băng số ($f_{\text{inter}}(k)$), giá trị biên của băng được sử dụng. Ma trận kết quả $X_{\text{eh}}^{\text{freq}}(k, n)$ chứa tần số cho mỗi ô thời gian - tần số.

Bước cuối của thuật toán nén dữ liệu hiệu chỉnh là chuyển đổi dữ liệu tần số trở về dữ liệu PDT:

$$X_{\text{eh}}^{\text{pdt}}(k, n) = 2\pi \cdot \left(\frac{64 \cdot X_{\text{estim}}^{\text{freq}}(k, n)}{f_s} \mod 1 \right), \quad (36)$$

trong đó $\text{mod}()$ biểu thị bộ vận hành môđun. Thuật toán hiệu chỉnh thực tế làm việc như được thể hiện trong phần 8.1. $Z_{\text{th}}^{\text{pdt}}(k, n)$ trong phương trình 16a được thay thế bằng $X_{\text{eh}}^{\text{pdt}}(k, n)$ như PDT đích, và các phương trình từ 17 đến 19 được sử dụng như trong phần 8.1. Kết quả của thuật toán hiệu chỉnh với dữ liệu hiệu chỉnh được nén được thể hiện trên Fig.50. Fig.50 thể hiện sai số trong PDT $D_{\text{sm}}^{\text{pdt}}(k, n)$ của tín hiệu đàn viôlông trong miền QMF của SBR được hiệu chỉnh với dữ liệu hiệu chỉnh được nén. Fig.50b thể hiện đạo hàm pha theo tần số tương ứng $Z_{\text{ch}}^{\text{pdt}}(k, n)$. Các gradien màu biểu thị các giá trị pha từ đỏ = π đến xanh da trời = $-\pi$. Các giá trị PDT theo sau các giá trị PDT của tín hiệu ban đầu với độ chính xác tương tự như phương pháp hiệu chỉnh mà không có sự nén dữ liệu (xem Fig.18). Do đó, thuật toán nén là hợp lệ. Chất lượng được nhận biết có và không có sự nén của dữ liệu hiệu chỉnh là tương tự.

Các phương án sử dụng độ chính xác cao hơn cho các tần số thấp và thấp hơn cho các tần số cao, sử dụng tổng cộng 12 bit cho mỗi giá trị. Tốc độ bit thu được là khoảng 0,5 kbps (không có bất kỳ sự nén, như mã hóa entropy). Độ chính xác này tạo ra chất lượng cảm quan như không có sự lượng tử hóa. Tuy nhiên, tốc độ bit thấp hơn một cách đáng kể có thể được sử dụng trong nhiều trường hợp tạo ra chất lượng cảm quan đủ tốt.

Một tùy chọn cho các sơ đồ tốc độ bit thấp là để ước tính tần số cơ bản trong pha giải mã sử dụng tín hiệu được truyền. Trong trường hợp này không giá trị nào được truyền. Tùy chọn khác là để ước tính tần số cơ bản sử dụng tín hiệu được truyền, so sánh nó với ước tính được thu sử dụng tín hiệu tín hiệu băng thông rộng, và chỉ để truyền sự chênh lệch. Có thể giả định rằng độ chênh lệch có thể được biểu diễn sử dụng tốc độ bit rất thấp.

9.2 Nén dữ liệu hiệu chỉnh PDF

Như được thảo luận trong phần 8.2, dữ liệu thích hợp cho việc hiệu chỉnh PDF là sai số pha trung bình của phần ráp nối tần số thứ nhất $D_{avg}^{pha}(n)$. Việc hiệu chỉnh có thể được thực hiện cho tất cả các phần ráp nối tần số với sự hiểu biết về giá trị này, do đó việc truyền chỉ một giá trị cho mỗi khung thời gian được yêu cầu. Tuy nhiên, việc truyền ngay cả giá trị đơn cho mỗi khung thời gian có thể sinh ra tốc độ bit rất cao.

Xem xét Fig.12 cho kèn trombon, có thể thấy rằng PDF có giá trị bất biến tương đối theo tần số, và giá trị tương tự được biểu diễn cho một vài khung theo thời gian. Giá trị là bất biến theo thời gian miễn là sự chuyển tiếp tương tự đang chiếm năng lượng của cửa sổ phân tích QMF. Khi sự chuyển tiếp mới bắt đầu chiếm ưu thế, giá trị mới được biểu diễn. Sự thay đổi góc giữa các giá trị PDF có vẻ như tương tự từ một sự chuyển tiếp tới sự chuyển tiếp khác. Điều này là có nghĩa, do PDF đang điều khiển vị trí theo thời gian của sự chuyển tiếp, và nếu tín hiệu có tần số cơ bản không đổi, khoảng cách giữa các sự chuyển tiếp nên không đổi.

Do đó, PDF (hoặc vị trí của sự chuyển tiếp) có thể chỉ được truyền rải rác theo thời gian, và trạng thái PDF giữa các các thời điểm có thể được ước tính sử dụng sự hiểu biết về tần số cơ bản. Sự hiệu chỉnh PDF có thể được thực hiện sử dụng thông tin này. Ý tưởng này thực tế là đối ngẫu với sự hiệu chỉnh PDF, trong đó các tần số của các sóng hài được giả định đặt cách đều nhau. Ở đây, ý tưởng tương tự được sử dụng, nhưng thay vào đó, các vị trí theo thời gian của sự chuyển tiếp được giả định được đặt cách đều nhau. Phương pháp được đề xuất dưới đây là dựa trên cơ sở phát hiện các vị trí của các đỉnh trong dạng sóng, và sử dụng thông tin này, phổ tham chiếu được tạo ra cho sự hiệu chỉnh pha.

9.2.1 Sử dụng sự phát hiện đỉnh cho việc nén dữ liệu hiệu chỉnh PDF - Tạo phổ đích cho sự hiệu chỉnh dọc

Các vị trí của các đỉnh phải được ước tính để thực hiện sự hiệu chỉnh PDF thành công. Một giải pháp để tính toán các vị trí của các đỉnh sử dụng giá trị PDF, tương tự như trong phương trình 34, và để ước tính các vị trí của các đỉnh ở giữa sử dụng tần số cơ bản được ước tính. Tuy nhiên, phương thức này đòi hỏi ước tính tần số cơ bản tương đối ổn định. Các phương án thể hiện sự đơn giản, nhanh chóng để thực hiện, phương pháp thay thế, mà thể hiện rằng cách tiếp cận việc nén được đề xuất này là khả thi.

Phép biểu diễn thời gian-tần số của tín hiệu kèn trombon được thể hiện trên Fig.51. Fig.51a thể hiện dạng sóng của tín hiệu kèn trombon trong phép biểu diễn miền thời gian. Fig.51b biểu diễn tín hiệu miền thời gian tương ứng mà chỉ chứa các đỉnh được ước tính, trong đó các vị trí của các đỉnh được sử dụng siêu dữ liệu được truyền. Tín hiệu trên Fig.51b là chuỗi xung 265 được mô tả, ví dụ đối với Fig.30. Thuật toán bắt đầu bằng việc phân tích các vị trí của các đỉnh ở dạng sóng. Điều này được thực hiện bằng việc tìm kiếm các cực đại cục bộ. Đối với mỗi 27ms (tức là, đối với mỗi 20 khung QMF), vị trí của đỉnh gần nhất với điểm trung tâm của khung được truyền. Giữa các vị trí đỉnh được truyền, các đỉnh được giả định là cách đều nhau theo thời gian. Do đó, bằng việc biết tần số cơ bản, các vị trí của các đỉnh có thể được ước tính. Trong phương án này, số lượng các đỉnh được phát hiện được truyền (cần lưu ý rằng điều này đòi hỏi việc phát hiện thành công tất cả các đỉnh; tần số cơ bản dựa trên việc ước tính có thể sinh ra các kết quả mạnh hơn). Tốc độ bit thu được là khoảng 0,5kbps (không có bất kỳ sự nén, như mã hóa entropy), mà bao gồm việc truyền vị trí của đỉnh cho mỗi 27ms sử dụng 9 bit và truyền số lượng các chuyển tiếp ở giữa sử dụng 4 bit. Các tác giả sáng chế tìm ra rằng độ chính xác này tạo ra chất lượng cảm quan như không có sự lượng tử hóa. Tuy nhiên, tốc độ bit thấp hơn một cách đáng kể có thể được sử dụng trong nhiều trường hợp tạo ra chất lượng cảm quan đủ tốt.

Việc sử dụng siêu dữ liệu được truyền, tín hiệu miền thời gian được tạo ra, mà bao gồm các xung ở các vị trí của các đỉnh được ước tính (xem Fig.51b). Việc phân tích QMF được thực hiện đối với tín hiệu này, và phổ pha $X_{ev}^{pha}(k, n)$ được tính toán.

Sự hiệu chỉnh PDF thực tế được thực hiện bằng cách khác như được đề xuất trong Phần 8.2, nhưng $Z_{th}^{pha}(k, n)$ trong phương trình 20a được thay thế bằng $X_{ev}^{pha}(k, n)$.

Dạng sóng của các tín hiệu có sự nhất quán pha dọc thường yếu và gợi nhớ về chuỗi xung. Do đó, đề xuất rằng phổ pha đích cho sự hiệu chỉnh dọc có thể được ước tính bằng việc tạo mẫu nó như phổ pha của chuỗi xung mà có các đỉnh tại các vị trí tương ứng và tần số cơ bản tương ứng.

Các vị trí gần trung tâm của khung theo thời gian nhất được truyền cho, ví dụ, mỗi khung thời gian thứ 20 (tương ứng với khoảng là -27ms). Tần số cơ bản được ước tính, mà được truyền với tốc độ bằng nhau, được sử dụng để nội suy các vị trí đỉnh ở giữa các vị trí được truyền.

Ngoài ra, tần số cơ bản và các vị trí đỉnh có thể được ước tính trong giai đoạn giải mã, và không có thông tin nào được truyền. Tuy nhiên, các ước tính tốt hơn có thể được dự kiến nếu ước tính được thực hiện với tín hiệu ban đầu trong giai đoạn mã hóa.

Việc xử lý bộ giải mã bắt đầu bằng việc thu được ước tính tần số cơ bản $X^{f_0}(n)$ cho mỗi khung theo thời gian và ngoài ra, các vị trí đỉnh ở dạng sóng cũng được ước tính. Các vị trí đỉnh được sử dụng để tạo ra tín hiệu miền thời gian mà bao gồm các xung tại các vị trí này. Phân tích QMF được sử dụng để tạo ra phổ pha tương ứng $X_{ev}^{pha}(k, n)$. Phổ pha được ước lượng này có thể được sử dụng trong phương trình 20a như phổ pha đích:

$$Z_{tv}^{pha}(k, n) = X_{ev}^{pha}(k, n). \quad (37)$$

Phương pháp được đề xuất sử dụng giai đoạn mã hóa chỉ để truyền các vị trí đỉnh được ước tính và các tần số cơ bản với tốc độ cập nhật, ví dụ, 27ms. Ngoài ra, nên lưu ý rằng các sai số trong đạo hàm pha dọc có thể nhận biết được chỉ khi tần số cơ bản là tương đối thấp. Do đó, tần số cơ bản có thể được truyền với tốc độ bit tương đối thấp.

Kết quả của thuật toán hiệu chỉnh với dữ liệu hiệu chỉnh được nén được thể hiện trên Fig.52. Fig.52a thể hiện sai số trong phổ pha $D_{cv}^{pha}(k, n)$ của tín hiệu kèn trombon trong miền QMF với SBR được hiệu chỉnh và dữ liệu hiệu chỉnh được nén. Theo đó,

Fig.52b thể hiện đạo hàm pha tương ứng theo tần số $Z_{cv}^{pdf}(k, n)$. Gradient màu biểu thị các giá trị từ đỏ = π tới xanh da trời = $-\pi$. Các giá trị PDF theo sau các giá trị PDF của tín hiệu ban đầu với độ chính xác tương tự như phương pháp hiệu chỉnh mà không có sự nén dữ liệu (xem Fig.13). Do đó, thuật toán nén là hợp lệ. Chất lượng được nhận biết có và không có sự nén của dữ liệu hiệu chỉnh là tương tự.

9.3 Nén dữ liệu xử lý chuyển tiếp

Do các sự chuyển tiếp có thể được giả định là rải rác tương đối, có thể giả định rằng dữ liệu này có thể được truyền trực tiếp. Các phương án thể hiện việc truyền sáu giá trị cho mỗi sự chuyển tiếp: một giá trị cho PDF trung bình, và năm giá trị cho các sai số trong góc pha tuyệt đối (một giá trị cho mỗi khung thời gian bên trong khoảng $[n - 2, n + 2]$). Cách khác là để truyền vị trí của sự chuyển tiếp (tức là một giá trị) và để ước tính phổ pha đích $X_{et}^{pha}(k, n)$ như trong trường hợp hiệu chỉnh dọc.

Nếu tốc độ bit cần thiết để được nén cho sự chuyển tiếp, cách tiếp cận tương tự có thể được sử dụng như cho hiệu chỉnh PDF (xem Phần 9.2). Vị trí của sự chuyển tiếp có thể được truyền, tức là giá trị đơn. Phổ pha đích và PDF đích có thể thu được sử dụng giá trị vị trí này như trong phần 9.2.

Ngoài ra, vị trí chuyển tiếp có thể được ước tính trong giai đoạn giải mã, và không có thông tin nào được truyền. Tuy nhiên, các ước tính tốt hơn có thể được dự kiến nếu ước tính được thực hiện với tín hiệu ban đầu trong giai đoạn mã hóa.

Tất cả các phương án được mô tả trước đây có thể được thấy một cách riêng biệt từ các phương án khác hoặc trong tổ hợp các phương án khác. Do đó, Fig.53 đến Fig.57 thể hiện bộ mã hóa và bộ giải mã tổ hợp một số phương án được mô tả trước đó.

Fig.53 thể hiện bộ giải mã 110'' để giải mã tín hiệu âm thanh. Bộ giải mã 110'' bao gồm bộ tạo phổ đích thứ nhất 65a, bộ hiệu chỉnh pha thứ nhất 70a và bộ tính toán tín hiệu băng phụ âm thanh 350. Bộ tạo phổ đích thứ nhất 65a, cũng được đề cập đến như bộ xác định số đo pha đích, tạo ra phổ đích 85a'' cho khung thời gian thứ nhất của tín hiệu băng phụ của tín hiệu âm thanh 32 sử dụng dữ liệu hiệu chỉnh thứ nhất 295a. Bộ hiệu chỉnh pha thứ nhất 70a hiệu chỉnh pha 45 của tín hiệu băng phụ trong khung

thời gian thứ nhất của tín hiệu âm thanh 32 được xác định với thuật toán hiệu chỉnh pha, trong đó sự hiệu chỉnh được hiện bằng việc giảm sự chênh lệch giữa số đo của tín hiệu băng phụ trong khung thời gian thứ nhất của tín hiệu âm thanh 32 và phổ đích 85a". Bộ tính toán tín hiệu băng phụ âm thanh 350 tính toán tín hiệu băng phụ âm thanh 355 cho khung thời gian thứ nhất sử dụng pha được hiệu chỉnh 91a cho khung thời gian. Ngoài ra, bộ tính toán tín hiệu băng phụ âm thanh 350 tính toán tín hiệu băng phụ âm thanh cho khung thời gian thứ hai khác với khung thời gian thứ nhất sử dụng số đo của tín hiệu băng phụ 85a" trong khung thời gian thứ hai hoặc sử dụng sự tính toán pha được hiệu chỉnh phù hợp với thuật toán hiệu chỉnh pha khác khác với thuật toán hiệu chỉnh pha này. Fig.53 còn thể hiện bộ phân tích 360 mà phân tích tùy chọn tín hiệu âm thanh 32 với cường độ 47 và pha 45. Thuật toán hiệu chỉnh pha khác có thể được thực hiện trong bộ hiệu chỉnh pha thứ hai 70b hoặc bộ hiệu chỉnh pha thứ ba 70c. Các bộ hiệu chỉnh pha này sẽ được minh họa đối với Fig.54. Bộ tính toán tín hiệu băng phụ âm thanh 250 tính toán tín hiệu băng phụ âm thanh cho khung thời gian thứ nhất sử dụng pha được hiệu chỉnh 91 cho khung thời gian thứ nhất và giá trị cường độ 47 của tín hiệu băng phụ âm thanh của khung thời gian thứ nhất, trong đó giá trị cường độ 47 là cường độ của tín hiệu âm thanh 32, trong khung thời gian thứ nhất hoặc cường độ được xử lý của tín hiệu âm thanh 35 trong khung thời gian thứ nhất.

Fig.54 thể hiện phương án nữa của bộ giải mã 110". Do đó, bộ giải mã 110" bao gồm bộ tạo phổ đích thứ hai 65b, trong đó bộ tạo phổ đích thứ hai 65b tạo ra phổ đích 85b" cho khung thời gian thứ hai của băng phụ của tín hiệu âm thanh 32 sử dụng dữ liệu hiệu chỉnh thứ hai 295b. Ngoài ra bộ phát hiện 110" bao gồm bộ hiệu chỉnh pha thứ hai 70b để hiệu chỉnh pha 45 của tín hiệu băng phụ trong khung thời gian của tín hiệu âm thanh 32 được xác định với thuật toán hiệu chỉnh pha thứ hai, trong đó sự hiệu chỉnh được thực hiện bằng cách giảm chênh lệch giữa số đo của khung thời gian của băng phụ của tín hiệu âm thanh và phổ đích 85b".

Theo đó, bộ giải mã 110" bao gồm bộ tạo phổ đích thứ ba 65c, trong đó bộ tạo phổ đích thứ ba 65c tạo ra phổ đích cho khung thời gian thứ ba của băng phụ của tín hiệu âm thanh 32 sử dụng dữ liệu hiệu chỉnh thứ ba 295c. Ngoài ra, bộ giải mã 110" bao gồm bộ hiệu chỉnh pha thứ ba 70c để hiệu chỉnh pha 45 của tín hiệu băng phụ và khung thời gian của tín hiệu âm thanh 32 được xác định với thuật toán hiệu chỉnh pha

thứ ba, trong đó sự hiệu chỉnh được thực hiện bằng cách giảm chênh lệch giữa số đo của khung thời gian của băng phụ của tín hiệu âm thanh và phổ đích 85c. Bộ tính toán tín hiệu băng phụ âm thanh 350 có thể tính toán tín hiệu băng phụ âm thanh cho khung thời gian thứ ba khác với các khung thời gian thứ nhất và thứ hai sử dụng sự hiệu chỉnh pha của bộ hiệu chỉnh pha thứ ba.

Theo phương án, bộ hiệu chỉnh pha thứ nhất 70a được tạo cấu hình để lưu trữ tín hiệu băng phụ được hiệu chỉnh pha 91a của khung thời gian trước đó của tín hiệu âm thanh hoặc để nhận tín hiệu băng phụ được hiệu chỉnh pha của khung thời gian trước đó 375 của tín hiệu âm thanh từ bộ hiệu chỉnh pha thứ hai 70b của bộ hiệu chỉnh pha thứ ba 70c. Hơn nữa, bộ hiệu chỉnh pha thứ nhất 70a hiệu chỉnh pha 45 của tín hiệu âm thanh 32 trong khung thời gian hiện thời của tín hiệu băng phụ âm thanh trên cơ sở tín hiệu băng phụ được hiệu chỉnh pha được lưu trữ hoặc được nhận của khung thời gian trước đó 91a, 375.

Các phương án khác thể hiện bộ hiệu chỉnh pha thứ nhất 70a thực hiện sự hiệu chỉnh pha ngang, bộ hiệu chỉnh pha thứ hai 70b thực hiện sự hiệu chỉnh pha dọc, và bộ hiệu chỉnh pha thứ ba 70c thực hiện sự hiệu chỉnh pha cho các sự chuyển tiếp.

Theo quan điểm khác, Fig.54 thể hiện sơ đồ khối của giai đoạn giải mã trong thuật toán hiệu chỉnh pha. Đầu vào của việc xử lý là tín hiệu BWE trong miền thời gian - tần số và siêu dữ liệu. Một lần nữa, trong các ứng dụng thực tiễn sự hiệu chỉnh đạo hàm pha theo sáng chế được ưu tiên để đồng sử dụng giàn bộ lọc hoặc biến đổi sơ đồ BWE hiện có. Trong ví dụ hiện thời, đây là miền QMF như được sử dụng trong SBR. Bộ tách kênh thứ nhất (không được thể hiện) trích xuất dữ liệu hiệu chỉnh đạo hàm pha từ dòng bit của BWE được trang bị bộ mã hóa-giải mã cảm quan mà được nâng cao bằng sự hiệu chỉnh theo sáng chế.

Bộ tách kênh thứ hai 130 (DEMUX) đầu tiên chia siêu dữ liệu được nhận 135 thành dữ liệu kích hoạt 365 và dữ liệu hiệu chỉnh 295a-c cho các chế độ hiệu chỉnh khác nhau. Dựa trên dữ liệu kích hoạt, sự tính toán của phổ đích được kích hoạt cho chế độ hiệu chỉnh đúng (sự tính toán của phổ đích khác có thể không hoạt động). Sử dụng phổ đích, sự hiệu chỉnh pha được thực hiện cho tín hiệu BWE được nhận sử dụng chế độ hiệu chỉnh được mong muốn. Nên lưu ý rằng do sự hiệu chỉnh ngang 70 được

thực hiện một cách đệ quy (nói cách khác: phụ thuộc vào các khung tín hiệu trước đó), nó cũng nhận các ma trận hiệu chỉnh trước đó từ các chế độ hiệu chỉnh khác 70b, c. Cuối cùng, tín hiệu được hiệu chỉnh, hoặc tín hiệu chưa được xử lý được thiết lập cho đầu ra dựa trên dữ liệu kích hoạt.

Sau khi hiệu chỉnh dữ liệu pha, sự tổng hợp BWE cơ bản xuôi dòng khác được tiếp tục, trong trường hợp ví dụ hiện thời là sự tổng hợp SBR. Các biến thiên có thể tồn tại trong đó sự hiệu chỉnh pha được chèn vào dòng tín hiệu tổng hợp BWE. Ưu tiên hơn, sự hiệu chỉnh đạo hàm pha được hoàn thành như sự điều chỉnh ban đầu trên các phần ráp nối phổ thô có các pha $Z^{pha}(k, n)$ và tất cả việc xử lý BWE bổ sung hoặc các bước điều chỉnh (trong SBR điều này có thể sự bổ sung nhiễu âm, lọc nghịch đảo, các đường hình sin bị thiếu, v.v) được thực hiện xuôi chiều theo các pha được hiệu chỉnh $Z_c^{pha}(k, n)$.

Fig.55 thể hiện phương án nữa của bộ giải mã 110". Theo phương án này, bộ giải mã 110" bao gồm bộ giải mã lõi 115, bộ ráp nối 120, bộ tổng hợp 100 và khối A, mà bộ giải mã 110" theo các phương án trước đó được thể hiện trên Fig.54. Bộ giải mã lõi 115 được tạo cấu hình để giải mã tín hiệu âm thanh 25 trong khung thời gian với số lượng các băng phụ giảm so với tín hiệu âm thanh 55. Bộ ráp nối 120 ráp nối tập hợp các băng phụ của tín hiệu âm thanh được giải mã lõi 25 với số lượng các băng phụ giảm, trong đó tập hợp các băng phụ tạo ra phần ráp nối thứ nhất, với các băng phụ khác trong khung thời gian, liên kế với số lượng các băng phụ giảm, để thu được tín hiệu âm thanh 32 với số lượng các băng phụ đều. Bộ xử lý cường độ 125' xử lý các giá trị cường độ của tín hiệu băng phụ âm thanh 355 trong khung thời gian. Theo các bộ giải mã trước đó 110 và 110', bộ xử lý cường độ có thể là bộ áp dụng tham số mở rộng băng thông 125.

Nhiều phương án có thể suy xét trong đó các khối xử lý tín hiệu được chuyển. Ví dụ, bộ xử lý cường độ 125' và khối A có thể hoán đổi. Do đó, khối A làm việc trên tín hiệu âm thanh được khôi phục 35, trong đó các giá trị cường độ của các phần ráp nối đã được hiệu chỉnh. Ngoài ra, bộ tính toán tín hiệu băng phụ âm thanh 350 có thể được định vị sau bộ xử lý cường độ 125' để tạo ra tín hiệu âm thanh được hiệu chỉnh 355 từ phần được hiệu chỉnh pha và hiệu chỉnh cường độ của tín hiệu âm thanh.

Hơn nữa, bộ giải mã 110'' bao gồm bộ tổng hợp 100 để tổng hợp tín hiệu âm thanh được hiệu chỉnh pha và cường độ để thu được tín hiệu âm thanh được xử lý được tổ hợp tần số 90. Một cách tùy chọn, do cả sự hiệu chỉnh cường độ và pha được áp dụng trên tín hiệu âm thanh không được giải mã lỗi 25, tín hiệu âm thanh đã nêu có thể được truyền trực tiếp đến bộ tổng hợp 100. Bất kỳ khối xử lý tùy chọn nào trong một trong số các bộ giải mã được mô tả trước đó 110 hoặc 110' có thể cũng được áp dụng trong bộ giải mã 110''.

Fig.56 thể hiện bộ mã hóa 155' để mã hóa tín hiệu âm thanh 55. Bộ mã hóa 155'' bao gồm bộ xác định pha 380 được kết nối với bộ tính toán 270, bộ mã hóa lỗi 160, bộ trích xuất tham số 165, và bộ tạo tín hiệu đầu ra 170. Bộ xác định pha 380 xác định pha 45 của tín hiệu âm thanh 55 trong đó bộ tính toán 270 xác định dữ liệu hiệu chỉnh pha 295 cho tín hiệu âm thanh 55 trên cơ sở pha được xác định 45 của tín hiệu âm thanh 55. Bộ mã hóa lỗi 160 mã hóa lỗi tín hiệu âm thanh 55 để thu được tín hiệu âm thanh được mã hóa lỗi 145 có số lượng các băng phụ giảm so với tín hiệu âm thanh 55. Bộ trích xuất tham số 165 trích xuất các tham số 190 từ tín hiệu âm thanh 55 để thu được các phép biểu diễn tham số độ phân giải thấp cho tập hợp thứ hai gồm các băng phụ không gồm trong tín hiệu âm thanh được mã hóa lỗi. Bộ tạo tín hiệu đầu ra 170 tạo ra tín hiệu đầu ra 135 bao gồm các tham số 190, tín hiệu âm thanh được mã hóa lỗi 145 và dữ liệu hiệu chỉnh pha 295'. Một cách tùy chọn, bộ mã hóa 155'' bao gồm bộ lọc thông thấp 180 trước khi mã hóa lỗi tín hiệu âm thanh 55 và bộ lọc thông cao 185 trước khi trích xuất các tham số 190 từ tín hiệu âm thanh 55. Ngoài ra, thay vì lọc thông thấp hoặc thông cao tín hiệu âm thanh 55, thuật toán điền đầy khoảng trống có thể được sử dụng, trong đó bộ mã hóa lỗi 160 mã hóa lỗi số lượng giảm của các băng phụ, trong đó ít nhất một băng phụ trong số tập hợp các băng phụ không được mã hóa lỗi. Hơn nữa, bộ trích xuất tham số trích xuất các tham số 190 từ ít nhất một băng phụ không được mã hóa với bộ mã hóa lỗi 160.

Theo các phương án, bộ tính toán 270 bao gồm tập hợp các bộ tính toán dữ liệu hiệu chỉnh 285a-c để hiệu chỉnh sự hiệu chỉnh pha phù hợp với chế độ biến thiên thứ nhất, chế độ biến thiên thứ hai, hoặc chế độ biến thiên thứ ba. Ngoài ra, bộ tính toán 270 xác định dữ liệu kích hoạt 365 để kích hoạt một bộ tính toán dữ liệu hiệu chỉnh của tập hợp các bộ tính toán dữ liệu hiệu chỉnh 285a-c. Bộ tạo tín hiệu đầu ra 170 tạo

ra tín hiệu đầu ra bao gồm dữ liệu kích hoạt, các tham số, tín hiệu âm thanh được mã hóa lỗi và dữ liệu hiệu chỉnh pha.

Fig.57 thể hiện sự thực hiện thay thế của bộ tính toán 270 mà có thể được sử dụng trong bộ mã hóa 155” được thể hiện trên Fig.56. Bộ tính toán chế độ hiệu chỉnh 385 bao gồm bộ xác định độ biến thiên 275 và bộ so sánh độ biến thiên 280. Dữ liệu kích hoạt 365 là kết quả của việc so sánh các độ biến thiên khác nhau. Ngoài ra, dữ liệu kích hoạt 365 kích hoạt một trong số các bộ tính toán dữ liệu hiệu chỉnh 185a-c theo độ biến thiên đã xác định. Dữ liệu hiệu chỉnh đã được tính toán 295a, 295b, hoặc 295c có thể là đầu vào của bộ tạo tín hiệu đầu ra 170 của bộ mã hóa 155” và do đó là phần của tín hiệu đầu ra 135.

Các phương án thể hiện bộ tính toán 270 bao gồm bộ tạo siêu dữ liệu 390, mà tạo ra dòng siêu dữ liệu 295’ bao gồm dữ liệu hiệu chỉnh được tính toán 295a, 295b, hoặc 295c và dữ liệu kích hoạt 365. Dữ liệu kích hoạt 365 có thể được truyền tới bộ giải mã nếu bản thân dữ liệu hiệu chỉnh không bao gồm thông tin đủ của chế độ hiệu chỉnh hiện thời. Thông tin đủ có thể, ví dụ, là số lượng các bit được sử dụng để biểu diễn dữ liệu hiệu chỉnh, mà khác với dữ liệu hiệu chỉnh 295a, dữ liệu hiệu chỉnh 295b, và dữ liệu hiệu chỉnh 295c. Hơn nữa, bộ tạo tín hiệu đầu ra 170 có thể sử dụng bổ sung dữ liệu kích hoạt 365, sao cho bộ tạo siêu dữ liệu 390 có thể được bỏ qua.

Theo quan điểm khác, sơ đồ khối trên Fig.57 thể hiện giai đoạn mã hóa trong thuật toán hiệu chỉnh pha. Đầu vào của bộ xử lý là tín hiệu âm thanh ban đầu 55 và miền thời gian - tần số. Trong các ứng dụng thực tiễn, sự hiệu chỉnh đạo hàm pha theo sáng chế được ưu tiên để đồng sử dụng giàn bộ lọc hoặc phép biến đổi sơ đồ BWE hiện có. Trong ví dụ hiện thời, đây là miền QMF được sử dụng trong SBR.

Khối tính toán chế độ hiệu chỉnh đầu tiên tính toán chế độ hiệu chỉnh mà được áp dụng cho mỗi khung theo thời gian. Dựa trên dữ liệu kích hoạt 365, sự tính toán dữ liệu hiệu chỉnh 295a-c được kích hoạt trong chế độ hiệu chỉnh đúng (sự tính toán dữ liệu hiệu chỉnh khác có thể không hoạt động). Cuối cùng, bộ dồn kênh (multiplexer - MUX) tổ hợp dữ liệu kích hoạt và dữ liệu hiệu chỉnh từ các chế độ hiệu chỉnh khác nhau.

Bộ dồn kênh khác (không được thể hiện) hợp nhất dữ liệu hiệu chỉnh đạo hàm

pha vào dòng bit của BWE và bộ mã hóa cảm quan mà được nâng cao bằng sự hiệu chỉnh theo sáng chế.

Fig.58 thể hiện bộ giải mã 5800 để giải mã tín hiệu âm thanh. Phương pháp 5800 bao gồm bước 5805 "tạo ra phổ đích cho khung thời gian thứ nhất của tín hiệu băng phụ của tín hiệu âm thanh với bộ tạo phổ đích thứ nhất sử dụng dữ liệu hiệu chỉnh thứ nhất", bước 5810 "hiệu chỉnh phổ của tín hiệu băng phụ trong khung thời gian thứ nhất của tín hiệu âm thanh với bộ hiệu chỉnh pha thứ nhất được xác định với thuật toán hiệu chỉnh pha, trong đó sự hiệu chỉnh được thực hiện bằng việc giảm sự chênh lệch giữa số đo của tín hiệu băng phụ trong khung thời gian thứ nhất của tín hiệu âm thanh và phổ đích, và bước 5815 "tính toán tín hiệu băng phụ âm thanh cho khung thời gian thứ nhất với bộ tính toán tín hiệu băng phụ âm thanh sử dụng pha được hiệu chỉnh của khung thời gian và để tính toán các tín hiệu băng phụ âm thanh cho khung thời gian thứ hai khác khung thời gian thứ nhất sử dụng số đo của tín hiệu băng phụ trong khung thời gian thứ hai hoặc sử dụng sự tính toán pha được hiệu chỉnh phù hợp với thuật toán hiệu chỉnh pha khác khác với thuật toán hiệu chỉnh pha."

Fig.59 thể hiện phương pháp 5900 để mã hóa tín hiệu âm thanh. Phương pháp 5900 bao gồm bước 5905 "xác định pha của tín hiệu âm thanh với bộ xác định pha", bước 5910 "xác định dữ liệu hiệu chỉnh pha cho tín hiệu âm thanh với bộ tính toán trên cơ sở pha được xác định của tín hiệu âm thanh", bước 5915 "mã hóa lỗi tín hiệu âm thanh với bộ mã hóa lỗi để thu được tín hiệu âm thanh được mã hóa lỗi có số lượng giảm của các băng phụ so với tín hiệu âm thanh", bước 5920 "trích xuất các tham số từ tín hiệu âm thanh với bộ trích xuất tham số để thu được phép biểu diễn tham số độ phân giải thấp cho tập hợp thứ hai gồm các băng phụ không nằm trong tín hiệu âm thanh được mã hóa lỗi", và bước 5925 "tạo ra tín hiệu đầu ra với bộ tạo tín hiệu đầu ra bao gồm các tham số, tín hiệu âm thanh được mã hóa lỗi, và dữ liệu hiệu chỉnh pha".

Các phương pháp 5800 và 5900 cũng như các phương pháp 2300, 2400, 2500, 3400, 3500, 3600 và 4200 được mô tả trước đó, có thể được thực hiện trong chương trình máy tính chạy trên máy tính.

Phải lưu ý rằng tín hiệu âm thanh 55 được sử dụng như thuật ngữ chung cho tín hiệu âm thanh, đặc biệt là cho tín hiệu âm thanh ban đầu, tức là, tín hiệu âm thanh

chưa được xử lý, phần được truyền của tín hiệu âm thanh $X_{trans}(k, n)$ 25, tín hiệu băng cơ sở $X_{base}(k, n)$ 30, tín hiệu âm thanh được xử lý bao gồm các tần số cao hơn 32 khi được so sánh với tín hiệu âm thanh ban đầu, tín hiệu âm thanh được khôi phục 35, phần ráp nối tần số được hiệu chỉnh cường độ $Y(k, n, i)$ 40, pha 45 của tín hiệu âm thanh, hoặc cường độ 47 của tín hiệu âm thanh. Do đó, các tín hiệu âm thanh khác nhau có thể được trao lẫn nhau do ngữ cảnh của phương án.

Các phương án thay thế liên quan đến giàn bộ lọc khác hoặc các miền biến đổi được sử dụng cho việc xử lý thời gian - tần số theo sáng chế, ví dụ biến đổi Fourier thời gian ngắn (short time Fourier transform - STFT), biến đổi cosin rời rạc cải biên phức (Complex Modified Discrete Cosine Transform - CMDCT), hoặc miền biến đổi Fourier rời rạc (Discrete Fourier Transform - DFT). Do đó, các đặc tính pha cụ thể liên quan tới phép biến đổi có thể được xem xét. Cụ thể, nếu ví dụ các hệ số sao chép được sao chép từ số chẵn tới số lẻ hoặc ngược lại, tức là băng phụ thứ hai của tín hiệu âm thanh ban đầu được sao chép tới băng phụ thứ chín thay cho băng phụ thứ tám như được mô tả trong các phương án, liên hợp phức của phần ráp nối có thể được sử dụng để xử lý. Điều tương tự áp dụng cho phần phản chiếu của các phần ráp nối thay vì sử dụng, ví dụ, thuật toán sao chép để khắc phục thứ tự bị đảo ngược của các góc pha trong phần ráp nối.

Các phương án khác có thể bỏ thông tin phụ từ bộ mã hóa và ước tính một số hoặc tất cả các tham số hiệu chỉnh cần thiết tại khu vực bộ giải mã. Các phương án khác có thể có các sơ đồ ráp nối BWE cơ bản mà ví dụ sử dụng các phần băng cơ sở khác nhau, số lượng hoặc kích thước khác nhau của các phần ráp nối hoặc các kỹ thuật chuyển vị khác nhau, ví dụ phản chiếu phổ hoặc điều biến băng biên đơn (single side band modulation - SSB). Các biến thiên có thể tồn tại trong đó sự hiệu chỉnh pha được phối hợp một cách chính xác vào dòng tín hiệu tổng hợp BWE. Hơn nữa, việc làm nhẵn được thực hiện sử dụng cửa sổ Hann trượt, mà có thể được thay thế cho hiệu suất tính toán tốt hơn bằng, ví dụ, IIR thứ tự thứ nhất.

Việc sử dụng các bộ mã hóa-giải mã âm thanh cảm quan theo tình trạng kỹ thuật thường làm suy yếu độ nhất quán pha của các thành phần phổ của tín hiệu âm thanh, đặc biệt tại tốc độ bit thấp, trong đó các kỹ thuật mã hóa theo tham số như mở

rộng băng thông được áp dụng. Điều này dẫn đến sự thay đổi của đạo hàm pha của tín hiệu âm thanh. Tuy nhiên, trong các loại tín hiệu nhất định sự bảo toàn của đạo hàm pha là quan trọng. Kết quả là, chất lượng cảm quan của các âm thanh này bị suy yếu. Sáng chế điều chỉnh lại đạo hàm pha theo tần số ("dọc") hoặc theo thời gian ("ngang") của các tín hiệu này nếu sự khôi phục của đạo hàm pha có lợi về cảm quan. Ngoài ra, quyết định được đưa ra liệu việc điều chỉnh đạo hàm pha dọc hoặc ngang được ưu tiên hơn một cách cảm quan. Sự truyền chỉ thông tin phụ nhỏ gọn là cần thiết để điều khiển việc xử lý hiệu chỉnh đạo hàm pha. Do đó, sáng chế cải thiện chất lượng âm thanh của các bộ mã hóa âm thanh cảm quan với cái giá là thông tin phụ vừa phải.

Nói cách khác, sự sao chép băng phổ (spectral band replication - SBR) có thể dẫn đến các sai số trong phổ pha. Cảm giác của con người về các sai số này được nghiên cứu tiết lộ hai ảnh hưởng đáng kể một cách cảm quan: các chênh lệch trong các tần số và các vị trí theo thời gian của sóng hài. Các sai số tần số có vẻ nhận biết được chỉ khi tần số cơ bản là đủ cao mà chỉ có một sóng hài bên trong băng ERB. Do đó, các sai số thời gian-vị trí có vẻ cảm thấy được chỉ khi tần số cơ bản thấp và nếu các pha của các sóng hài được căn chỉnh theo tần số.

Các sai số thời gian-vị trí có thể được phát hiện bằng việc tính toán đạo hàm pha trên tần số (PDF). Nếu các giá trị PDF ổn định theo thời gian, các chênh lệch giữa chúng giữa các tín hiệu được xử lý SBR và các tín hiệu ban đầu nên được hiệu chỉnh. Điều này hiệu chỉnh một cách hiệu quả các tần số của các sóng hài, và do đó, cảm quan của sự phi điều hòa được tránh.

Các sai số thời gian-vị trí có thể được phát hiện bằng việc tính toán đạo hàm pha theo tần số (PDF). Nếu các giá trị PDF ổn định theo tần số, các chênh lệch giữa chúng giữa các tín hiệu được xử lý SBR và các tín hiệu ban đầu nên được hiệu chỉnh. Điều này hiệu chỉnh một cách hiệu quả vị trí theo thời gian của các sóng hài, và do đó, cảm quan của việc điều biến các nhiễu âm tại các tần số giao chéo được tránh.

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

các chức năng được thực hiện bởi các khối phần cứng logic hoặc vật lý.

Mặc dù một số khía cạnh đã được mô tả trong phạm vi của thiết bị, rõ ràng rằng các khía cạnh này cũng thể hiện sự mô tả của phương pháp tương ứng, trong đó khối hoặc thiết bị tương ứng với bước phương pháp hoặc dấu hiệu của bước phương pháp. Tương tự, các khía cạnh được mô tả trong 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 vài hoặc tất cả các bước phương pháp có thể được thực hiện bởi (hoặc sử dụng) thiết bị phần cứng, tương tự 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 truyền hoặc được mã hóa đổi mới 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 hiện nhất định, các phương án của sáng chế có thể được thực hiện 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 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 đọc được bằng máy tính.

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

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

Các phương án khác bao gồm chương trình máy tính để thực hiện một trong các

phương pháp được mô tả ở đây, được lưu trữ trên vật mang có thể đọc được bằng máy.

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

Do đó, phương án khác của các phương pháp theo sáng chế là, vật mang dữ liệu (hoặc vật ghi lưu trữ không tạm thời hoặc vật ghi lưu trữ số, hoặc vật ghi có thể đọc được bằng máy tính) gồm có, đã được ghi lại trên đó, chương trình máy tính để thực hiện một trong số các phương pháp được mô tả ở đây. Vật mang dữ liệu, vật ghi lưu trữ số hoặc vật ghi được ghi lại thường là hữu hình và/hoặc không tạm thời.

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

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

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

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

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

hiện bởi thiết bị phần cứng bất kỳ.

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

YÊU CẦU BẢO HỘ

1. Bộ xử lý âm thanh (50) để xử lý tín hiệu âm thanh (55) bao gồm:

bộ tính toán số đo pha tín hiệu âm thanh (60) được tạo cấu hình để tính toán số đo pha (80) của tín hiệu âm thanh cho khung thời gian (75a);

bộ xác định số đo pha đích (65) để xác định số đo pha đích (85) cho khung thời gian (75a) đã nêu;

bộ hiệu chỉnh pha (70) được tạo cấu hình để hiệu chỉnh các pha (45) của tín hiệu âm thanh (55) cho khung thời gian (75a) sử dụng số đo pha được tính toán (80) và số đo pha đích (85) để thu được tín hiệu âm thanh được xử lý (90).

2. Bộ xử lý âm thanh (50) theo điểm 1,

trong đó tín hiệu âm thanh (55) bao gồm nhiều tín hiệu băng phụ (95a,b) cho khung thời gian (75a);

trong đó bộ xác định số đo pha đích được tạo cấu hình để xác định số đo pha đích thứ nhất (85a) cho tín hiệu băng phụ thứ nhất (95a) và số đo pha đích thứ hai (85b) cho tín hiệu băng phụ thứ hai (95b);

trong đó bộ tính toán số đo pha tín hiệu âm thanh (60) được tạo cấu hình để xác định số đo pha thứ nhất (80a) cho tín hiệu băng phụ thứ nhất (95a) và số đo pha thứ hai (80b) cho tín hiệu băng phụ thứ hai (95b);

trong đó bộ hiệu chỉnh pha (70) được tạo cấu hình để hiệu chỉnh pha thứ nhất (45a) của tín hiệu băng phụ thứ nhất (95a) sử dụng số đo pha thứ nhất (80a) của tín hiệu âm thanh (55) và số đo pha đích thứ nhất (85) để thu được tín hiệu băng phụ được xử lý thứ nhất (90a) và để hiệu chỉnh pha thứ hai (45b) của tín hiệu băng phụ thứ hai (95b) sử dụng số đo pha thứ hai (80b) của tín hiệu âm thanh (55) và số đo pha đích thứ hai (85b) để thu được tín hiệu băng phụ được xử lý thứ hai (90b); và

bộ tổng hợp tín hiệu âm thanh (100) để tổng hợp tín hiệu âm thanh được xử lý (90) sử dụng tín hiệu băng phụ thứ nhất được xử lý (90a) và tín hiệu băng phụ thứ hai được xử lý (90b).

3. Bộ xử lý âm thanh (50) theo một trong số các điểm 1 hoặc 2,

trong đó số đo pha (80) là đạo hàm pha theo thời gian;

trong đó bộ tính toán số đo pha tín hiệu âm thanh (60) được tạo cấu hình để tính toán, cho mỗi băng phụ (95) của nhiều băng phụ, đạo hàm pha của giá trị pha của khung thời gian hiện thời (75b) và giá trị pha của khung thời gian tương lai (75c);

trong đó bộ hiệu chỉnh pha (70) được tạo cấu hình để tính toán, cho mỗi băng phụ (95) của nhiều băng phụ của khung thời gian hiện thời (75b), độ lệch (105) giữa đạo hàm pha đích (85) và đạo hàm pha theo thời gian (80);

trong đó sự hiệu chỉnh được thực hiện bằng bộ hiệu chỉnh pha (70) được thực hiện sử dụng độ lệch.

4. Bộ xử lý âm thanh (50) theo một trong số các điểm từ 1 đến 3,

trong đó bộ hiệu chỉnh pha (70) được tạo cấu hình để hiệu chỉnh các tín hiệu băng phụ (95) của các băng phụ khác nhau của tín hiệu âm thanh (55) trong khung thời gian (75), do đó các tần số của các tín hiệu băng phụ được hiệu chỉnh (90a,b) có các giá trị tần số được phân bố một cách hài hòa đến tần số cơ bản của tín hiệu âm thanh (55).

5. Bộ xử lý âm thanh (50) theo một trong số các điểm từ 1 đến 4,

trong đó bộ hiệu chỉnh pha (70) được tạo cấu hình để làm nhẵn độ lệch (105) cho mỗi băng phụ (95) của nhiều băng phụ trên khung thời gian trước (75a), hiện thời (75b), và tương lai (75c) và được tạo cấu hình để giảm các thay đổi nhanh chóng của độ lệch (105) trong băng phụ (95).

6. Bộ xử lý âm thanh (50) theo điểm 5,

trong đó việc làm nhẵn là trung bình được gán trọng số;

trong đó bộ hiệu chỉnh pha (70) được tạo cấu hình để tính toán trung bình được gán trọng số trên khung trước (75a), hiện thời (75b), và tương lai (75c), được gán trọng số bởi cường độ (47) của tín hiệu âm thanh (55) trong khung trước (75a), hiện thời (75b), và tương lai (75c).

7. Bộ xử lý âm thanh (50) theo một trong số các điểm từ 1 đến 6,

trong đó bộ hiệu chỉnh pha (70) được tạo cấu hình để tạo ra vectơ của các độ lệch (105), trong đó phần tử thứ nhất của vectơ đề cập đến độ lệch thứ nhất (105a) cho băng phụ thứ nhất (95a) của nhiều băng phụ và phần tử thứ hai của vectơ đề cập đến độ lệch thứ hai (105b) cho băng phụ thứ hai (95b) của nhiều băng phụ từ khung thời gian trước (75a) đến khung thời gian hiện thời (75b);

trong đó bộ hiệu chỉnh pha (70) được tạo cấu hình để áp dụng vectơ của các độ lệch (105) vào các pha (45) của tín hiệu âm thanh, trong đó phần tử thứ nhất của vectơ được áp dụng cho pha (45a) của tín hiệu âm thanh (55) trong băng phụ thứ nhất (95a) của nhiều băng phụ của tín hiệu âm thanh (55) và phần tử thứ hai của vectơ được áp dụng cho pha (45b) của tín hiệu âm thanh (55) trong băng phụ thứ hai (95b) của nhiều băng phụ của tín hiệu âm thanh (55).

8. Bộ xử lý âm thanh (50) theo một trong số các điểm từ 1 đến 7,

trong đó bộ xác định số đo pha đích (65) được tạo cấu hình để thu được ước tính tần số cơ bản (85) cho khung thời gian (75);

trong đó bộ xác định số đo pha đích (65) được tạo cấu hình để tính toán ước tính tần số (85) cho mỗi băng phụ (95) của nhiều băng phụ của khung thời gian (75) sử dụng tần số cơ bản của khung thời gian (75).

9. Bộ xử lý âm thanh (50) theo điểm 8,

trong đó bộ xác định số đo pha đích (65) được tạo cấu hình để chuyển đổi các ước tính tần số (85) cho mỗi băng phụ (95) của nhiều băng phụ thành đạo hàm pha theo thời gian (85) sử dụng tổng số lượng các băng phụ (95) và tần số lấy mẫu của tín hiệu âm thanh (55).

10. Bộ xử lý âm thanh (50) theo điểm 8 hoặc 9,

trong đó bộ xác định số đo pha đích (65) được tạo cấu hình để tạo vectơ của các ước tính tần số (85) cho mỗi băng phụ (95) của nhiều băng phụ, trong đó phần tử thứ nhất của vectơ đề cập đến ước tính tần số (85a) cho băng phụ thứ nhất (95a) và phần tử thứ hai của vectơ đề cập đến ước tính tần số (85b) cho băng phụ thứ hai (95b);

trong đó bộ xác định số đo pha đích (65) được tạo cấu hình để tính toán ước tính tần số (85) sử dụng các bội số của tần số cơ bản, trong đó ước tính tần số (85) của

băng phụ hiện thời (95) là bội số của tần số cơ bản mà gần nhất với trung tâm của băng phụ (95), hoặc trong đó ước tính tần số (85) của băng phụ hiện thời (95) là tần số biên của băng phụ hiện thời (95) nếu không có bội số nào trong số các bội số của tần số cơ bản trong băng phụ hiện thời (95).

11. Bộ giải mã (110) để giải mã tín hiệu âm thanh (55), bộ giải mã (110) bao gồm:

bộ xử lý âm thanh (50) theo một trong số các điểm từ 1 đến 10;

bộ giải mã lỗi (115) được tạo cấu hình để giải mã lỗi tín hiệu âm thanh (25) trong khung thời gian (75) với số lượng các băng phụ giảm so với tín hiệu âm thanh (55);

bộ ráp nối (120) được tạo cấu hình để ráp nối tập hợp các băng phụ (95) của tín hiệu âm thanh được giải mã lỗi (25) với số lượng các băng phụ giảm, trong đó tập hợp các băng phụ tạo ra phần ráp nối thứ nhất (30a), với các băng phụ khác trong khung thời gian (75), liền kề với số lượng các băng phụ giảm, để thu được tín hiệu âm thanh (55) với số lượng các băng phụ đều;

trong đó bộ xử lý âm thanh (50) được tạo cấu hình để hiệu chỉnh các pha (45) trong các băng phụ của phần ráp nối thứ nhất (30a) theo hàm đích (85).

12. Bộ giải mã (110) theo điểm 11,

trong đó bộ ráp nối (120) được tạo cấu hình để ráp nối tập hợp các băng phụ (95) của tín hiệu âm thanh (25), trong đó tập hợp các băng phụ tạo phần ráp nối thứ hai, với các băng phụ khác của khung thời gian (75), liền kề phần ráp nối thứ nhất; và

trong đó bộ xử lý âm thanh (50) được tạo cấu hình để hiệu chỉnh các pha (45) trong băng phụ (95) của phần ráp nối thứ hai; hoặc

trong đó bộ ráp nối (120) được tạo cấu hình để ráp nối phần ráp nối thứ nhất được hiệu chỉnh với các băng phụ khác của khung thời gian, liền kề với phần ráp nối thứ nhất.

13. Bộ giải mã (110) theo điểm 11 hoặc 12, bộ giải mã bao gồm:

bộ trích xuất dòng dữ liệu (130) được tạo cấu hình để trích xuất tần số cơ bản (140) của khung thời gian hiện thời (75) của tín hiệu âm thanh (55) từ dòng dữ liệu

(135), trong đó dòng dữ liệu còn bao gồm tín hiệu âm thanh được mã hóa (145) với số lượng băng phụ giảm; hoặc

bộ phân tích tần số cơ bản (150) được tạo cấu hình để phân tích tín hiệu âm thanh được giải mã lỗi (25) để tính toán tần số cơ bản (140).

14. Bộ mã hóa (155) để mã hóa tín hiệu âm thanh (55), bộ mã hóa (155) bao gồm:

bộ mã hóa lỗi (160) được tạo cấu hình để mã hóa lỗi tín hiệu âm thanh (55) để thu được tín hiệu âm thanh được mã hóa lỗi (145) có số lượng các băng phụ giảm so với tín hiệu âm thanh (55);

bộ phân tích tần số cơ bản (175) để phân tích tín hiệu âm thanh (55) hoặc phiên bản được lọc thông thấp của tín hiệu âm thanh để thu được ước tính tần số cơ bản (140) của tín hiệu âm thanh (155);

bộ trích xuất tham số (165) được tạo cấu hình để trích xuất các tham số của các băng phụ của tín hiệu âm thanh (55) không nằm trong tín hiệu âm thanh được mã hóa lỗi (145);

bộ tạo tín hiệu đầu ra (170) được tạo cấu hình để tạo ra tín hiệu đầu ra (135) bao gồm tín hiệu âm thanh được mã hóa lỗi (145), các tham số (190), và ước tính tần số cơ bản (140).

15. Bộ mã hóa (155) theo điểm 14, trong đó bộ tạo tín hiệu đầu ra (170) được tạo cấu hình để tạo ra tín hiệu đầu ra (135) thành chuỗi các khung, trong đó mỗi khung bao gồm tín hiệu âm thanh được mã hóa lỗi (145), các tham số (190), và trong đó mỗi khung thứ N bao gồm ước tính tần số cơ bản (140), trong đó N lớn hơn hoặc bằng 2.

16. Phương pháp (2300) xử lý tín hiệu âm thanh (55), phương pháp (2300) bao gồm các bước sau:

tính toán số đo pha của tín hiệu âm thanh (55) cho khung thời gian với bộ tính toán số đo pha tín hiệu âm thanh (60);

xác định số đo pha đích cho khung thời gian đã nêu với bộ xác định số đo pha đích (65);

hiệu chỉnh các pha của tín hiệu âm thanh (55) cho khung thời gian với bộ hiệu chỉnh pha (70) sử dụng số đo pha được tính toán và số đo pha đích để thu tín hiệu âm thanh chưa xử lý (90).

17. Phương pháp (2400) giải mã tín hiệu âm thanh (55), phương pháp (2400) bao gồm các bước sau:

giải mã tín hiệu âm thanh (25) trong khung thời gian với số lượng các băng phụ giảm so với tín hiệu âm thanh (55);

ráp nối tập hợp các băng phụ của tín hiệu âm thanh được giải mã lỗi (25) với số lượng các băng phụ giảm, trong đó tập hợp các băng phụ tạo ra phần ráp nối thứ nhất, với các băng phụ khác trong khung thời gian, liền kề với số lượng các băng phụ giảm, để thu được tín hiệu âm thanh (55) với số lượng các băng phụ đều;

hiệu chỉnh các pha trong các băng phụ của phần ráp nối thứ nhất theo hàm đích với phương pháp theo điểm 16.

18. Phương pháp mã hóa tín hiệu âm thanh (55), phương pháp bao gồm các bước sau:

mã hóa lỗi tín hiệu âm thanh với bộ mã hóa lỗi (160) để thu được tín hiệu âm thanh được mã hóa lỗi có số lượng các băng phụ giảm so với tín hiệu âm thanh (55);

phân tích tín hiệu âm thanh (55) hoặc phiên bản được lọc thông thấp của tín hiệu âm thanh với bộ phân tích tần số cơ bản để thu ước tính tần số cơ bản (140) của tín hiệu âm thanh;

trích xuất các tham số của các băng phụ của tín hiệu âm thanh (55) không nằm trong tín hiệu âm thanh được mã hóa lỗi (145) với bộ trích xuất tham số (165);

tạo ra tín hiệu đầu ra (135) bao gồm tín hiệu âm thanh được mã hóa lỗi (145), các tham số (190), và ước tính tần số cơ bản (140) với bộ tạo tín hiệu đầu ra (170).

19. Vật ghi có thể đọc được bằng máy tính bao gồm chương trình máy tính có mã chương trình được thích ứng để thực hiện phương pháp theo điểm 16 khi chương trình máy tính chạy trên máy tính.

20. Vật ghi có thể đọc được bằng máy tính bao gồm chương trình máy tính có mã chương trình được thích ứng để thực hiện phương pháp theo điểm 17 khi chương trình máy tính chạy trên máy tính.

21. Vật ghi có thể đọc được bằng máy tính bao gồm chương trình máy tính có mã chương trình được thích ứng để thực hiện phương pháp theo điểm 18 khi chương trình máy tính chạy trên máy tính.

22. Vật ghi có thể đọc được bằng máy tính lưu trữ trong đó tín hiệu âm thanh (135), tín hiệu âm thanh (135) bao gồm:

tín hiệu âm thanh được mã hóa lỗi (145) có số lượng các băng phụ giảm so với tín hiệu âm thanh ban đầu (55);

tham số (190) biểu diễn các băng phụ của tín hiệu âm thanh không nằm trong tín hiệu âm thanh được mã hóa lỗi (145);

ước tính tần số cơ bản (140) của tín hiệu âm thanh (135) hoặc tín hiệu âm thanh ban đầu (55).

23. Vật ghi có thể đọc được bằng máy tính lưu trữ trong đó tín hiệu âm thanh (135) theo điểm 22,

trong đó tín hiệu âm thanh (135) được tạo thành chuỗi các khung, trong đó mỗi khung bao gồm tín hiệu âm thanh được mã hóa lỗi (145), các tham số (190), và trong đó chỉ mỗi khung thứ N bao gồm ước tính tần số cơ bản (140), trong đó N lớn hơn hoặc bằng 2.

ban đầu: phổ cường độ (dB)

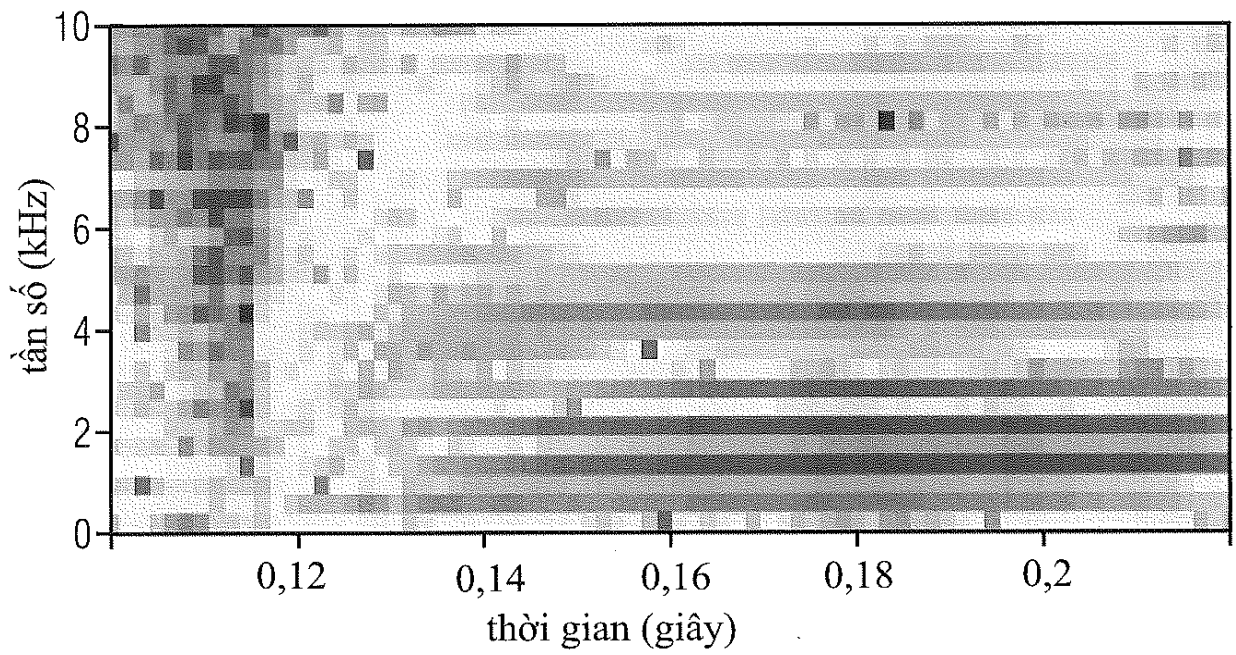


FIG 1A

ban đầu: phổ pha (radian)

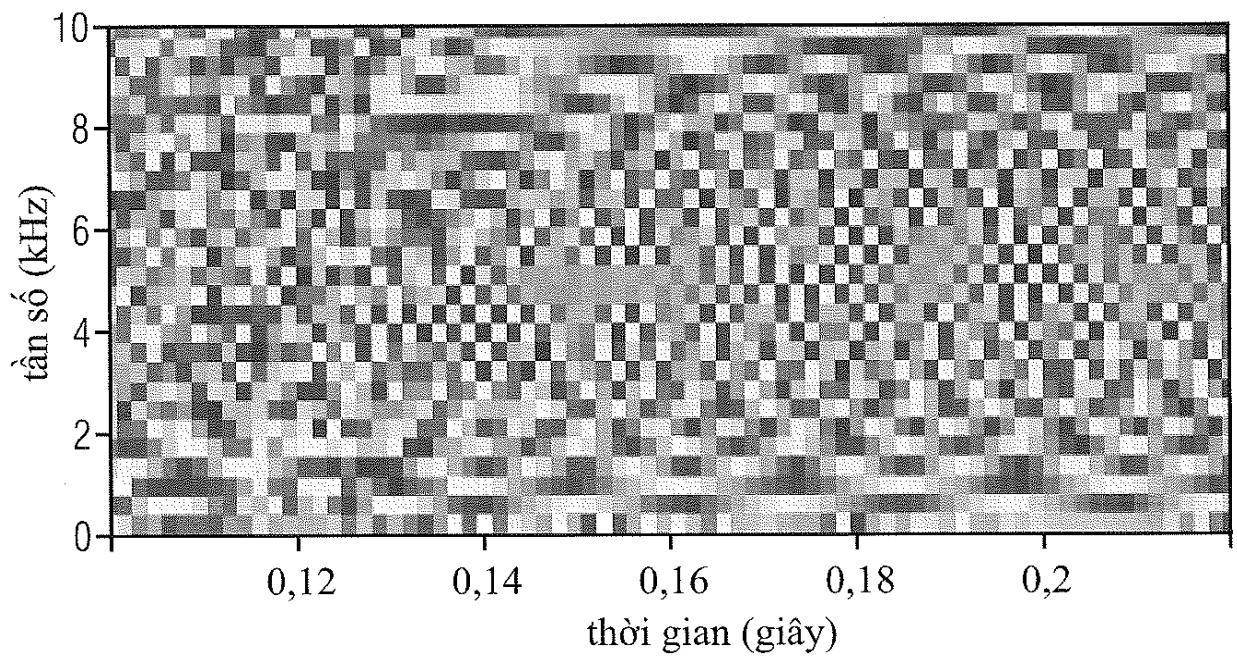


FIG 1B

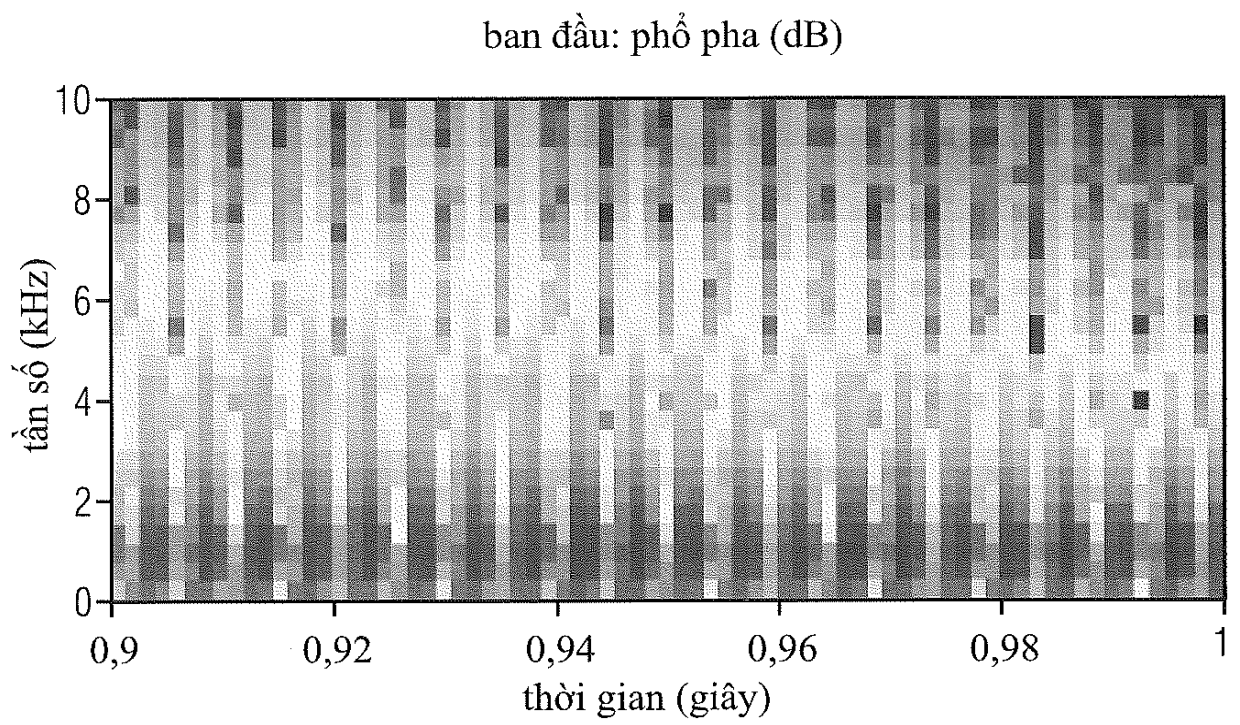


FIG 1C

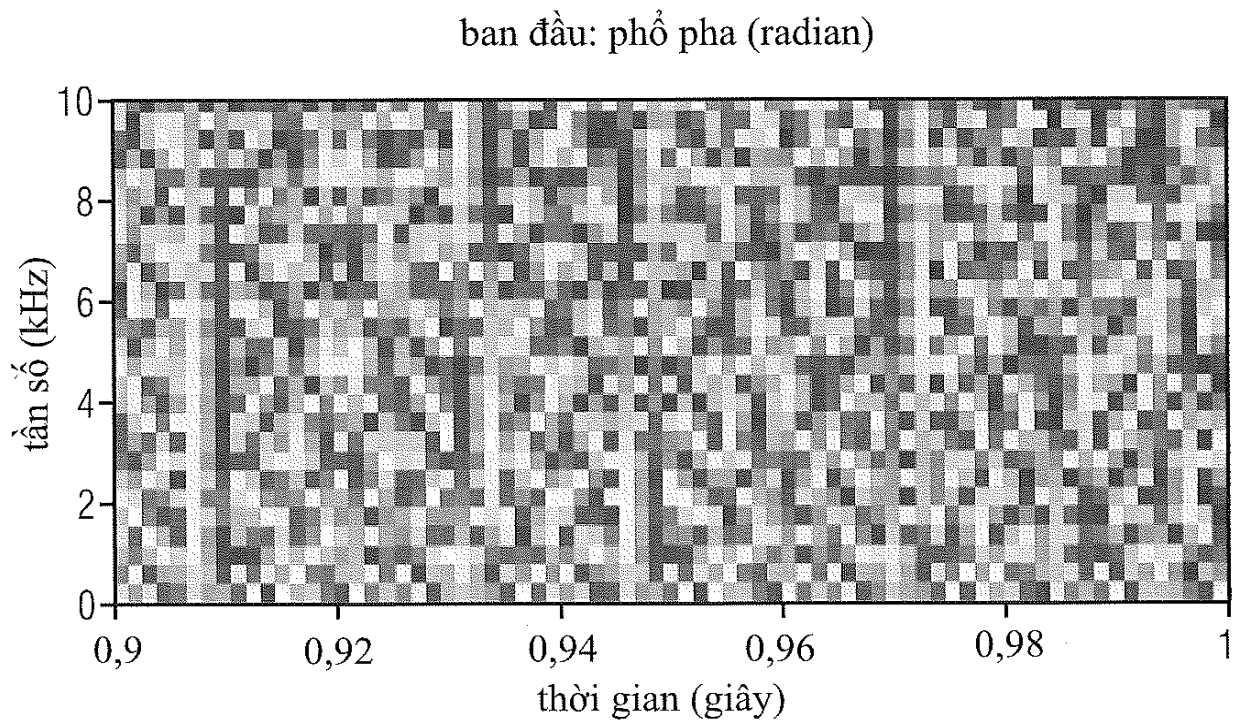


FIG 1D

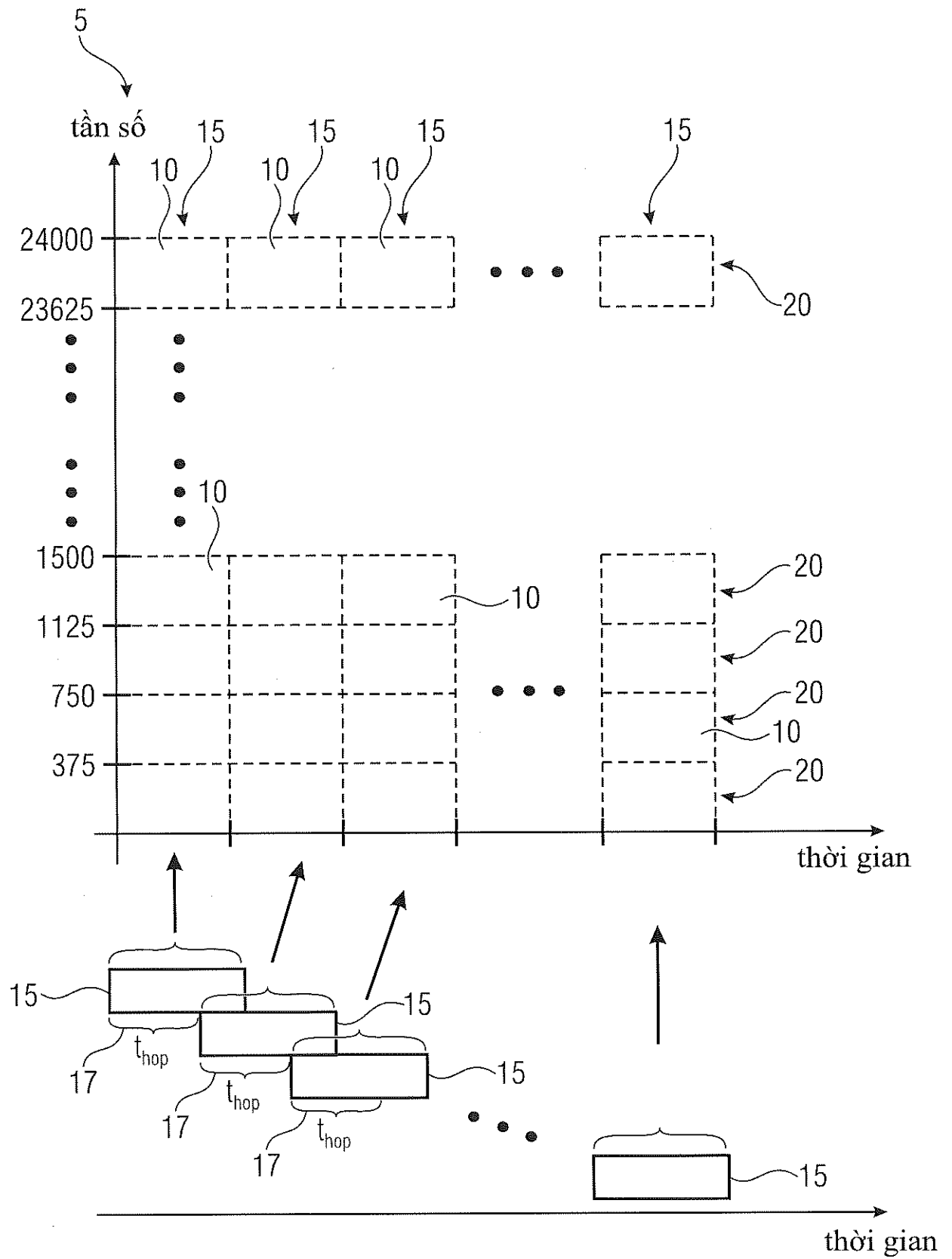


FIG 2

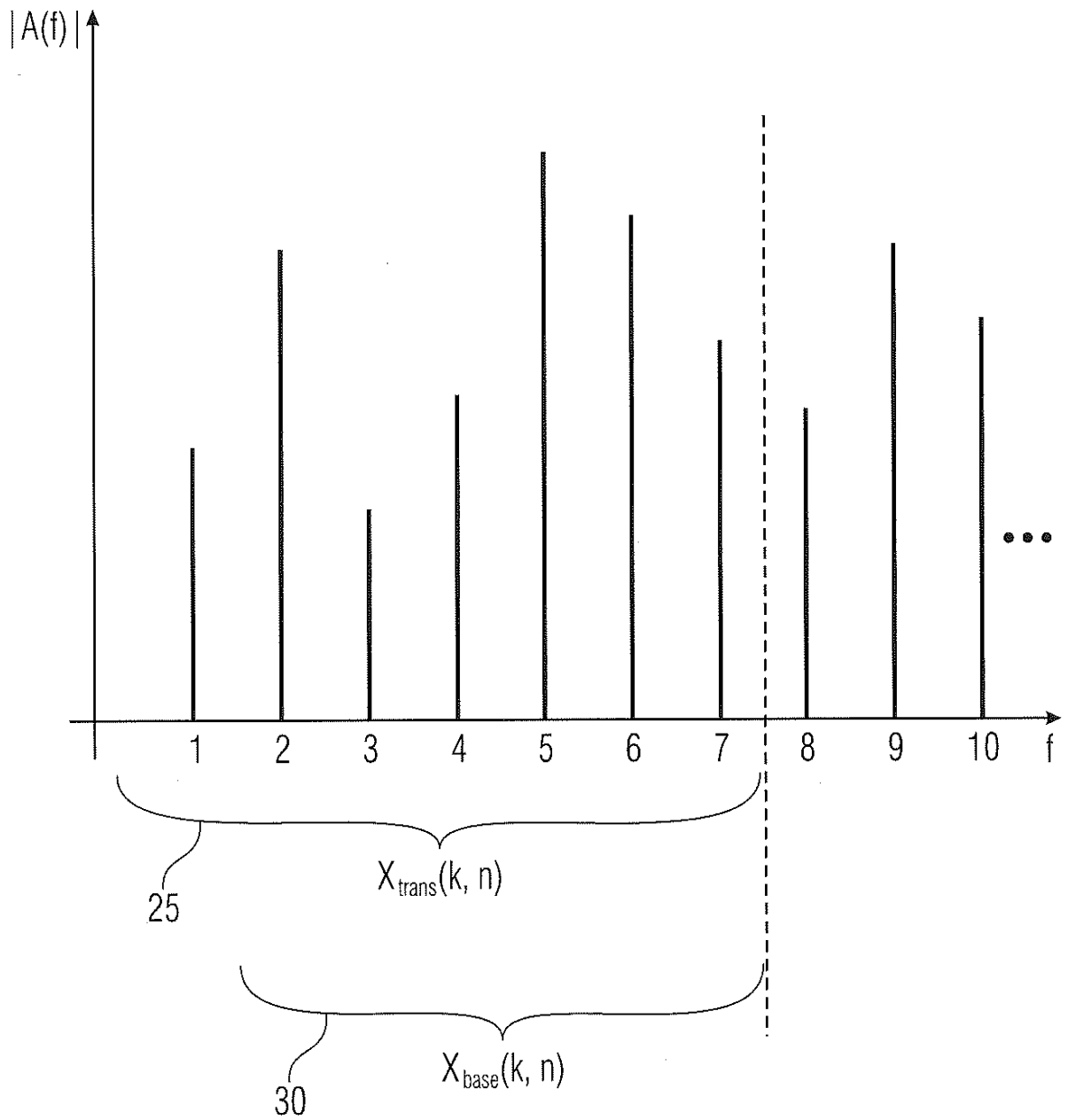


FIG 3A

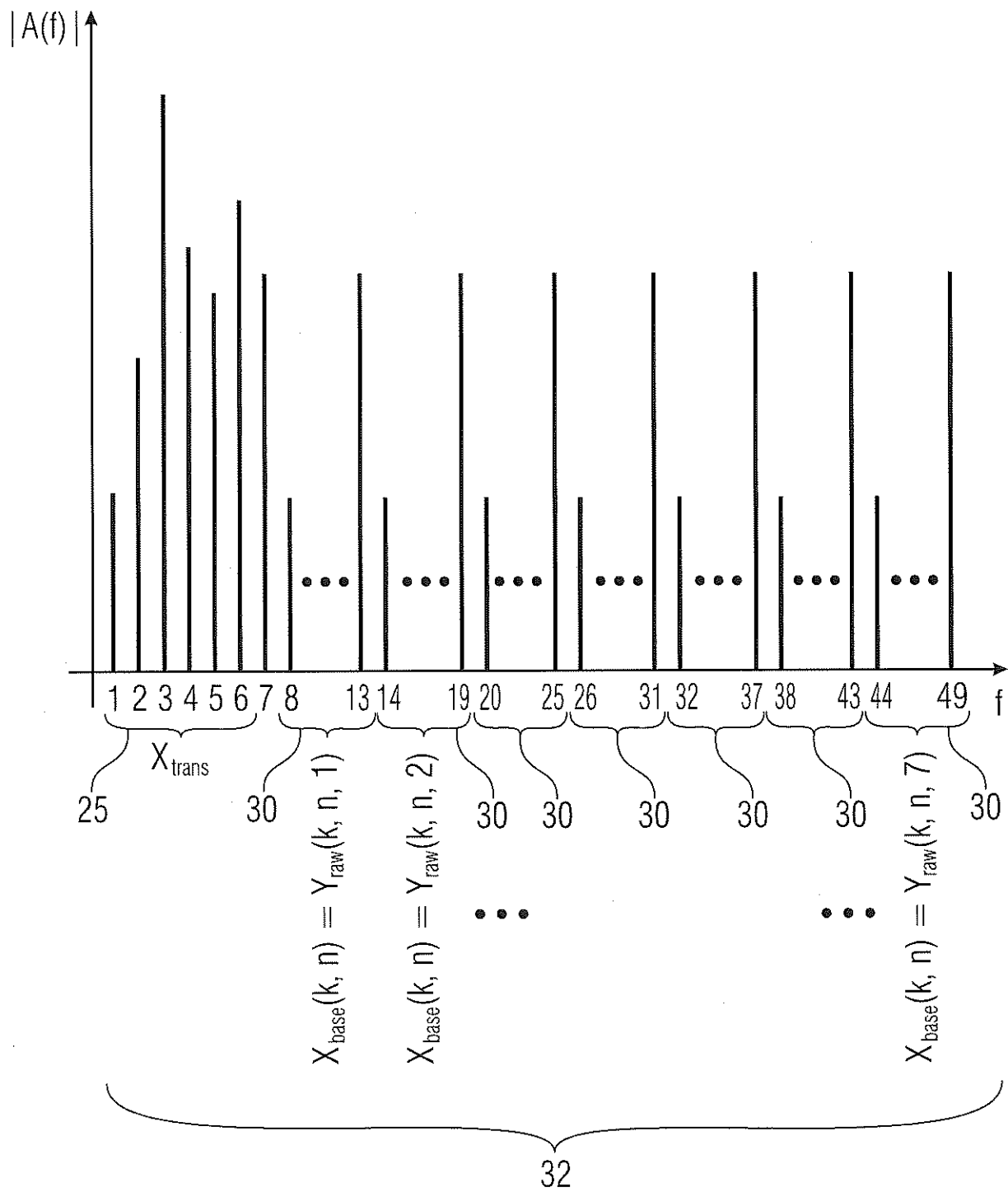


FIG 3B

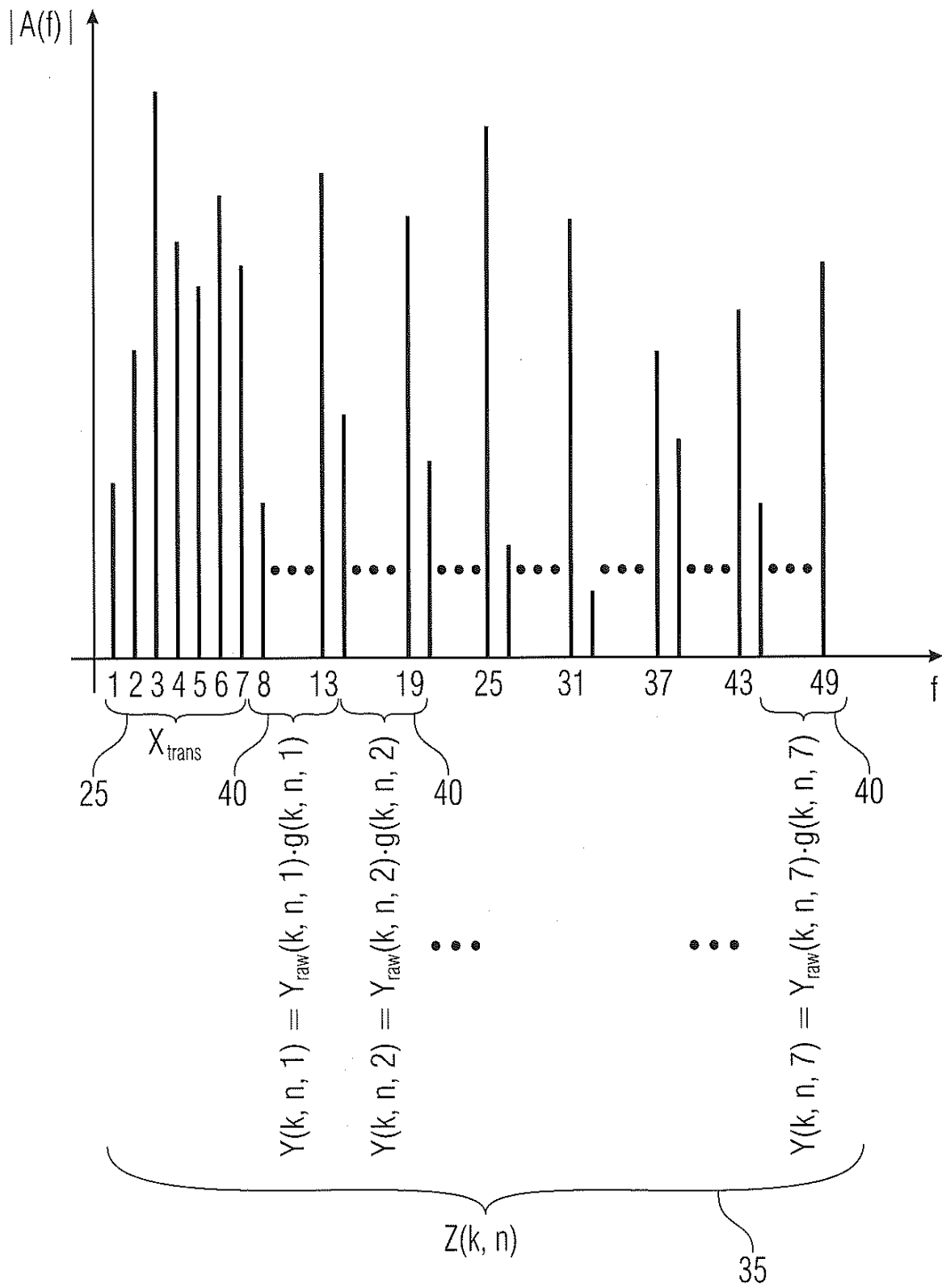


FIG 3C

sao chép: phổ cường độ (dB)

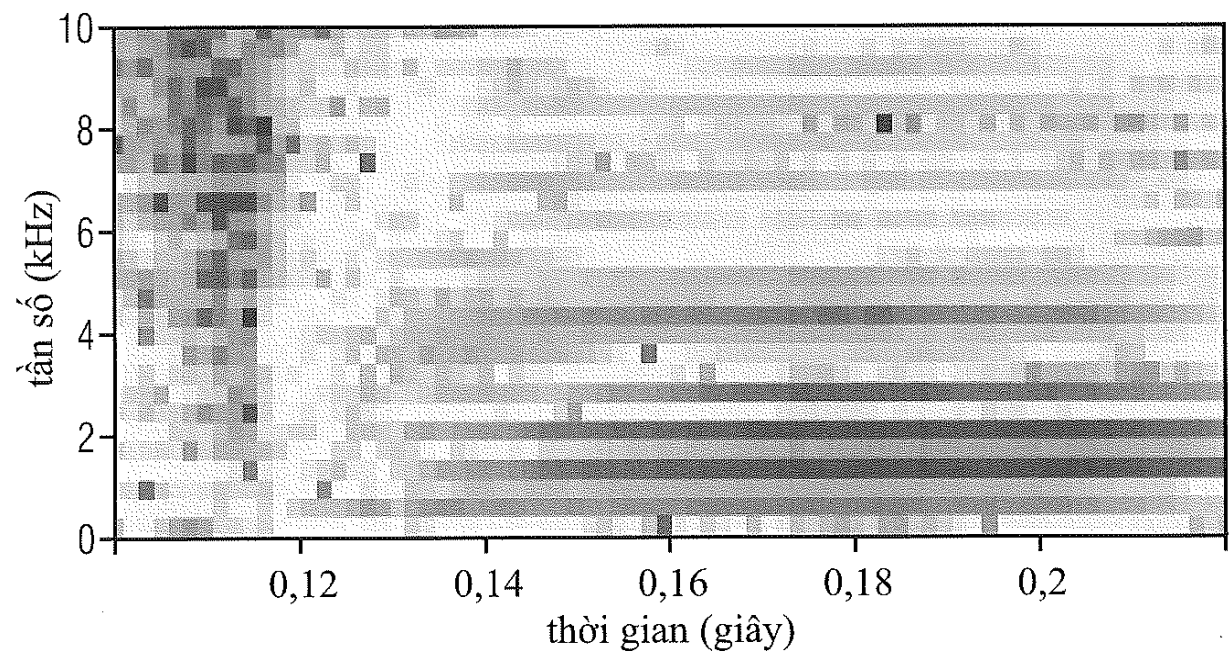


FIG 4A

sao chép: phổ pha (radian)

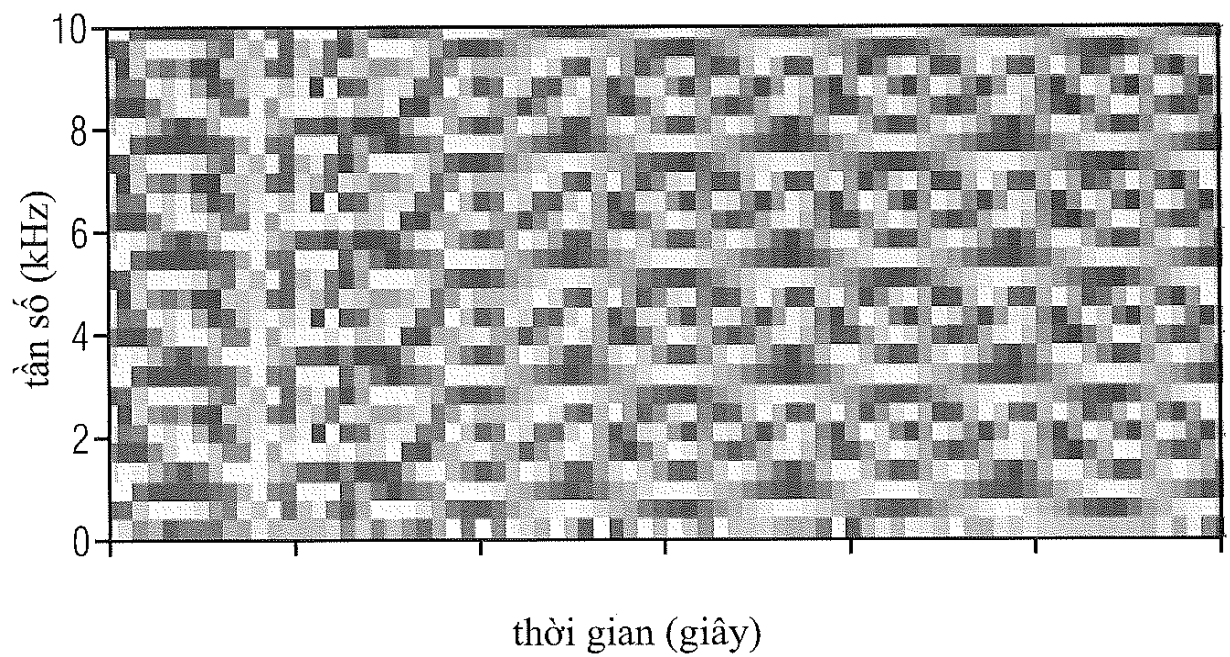


FIG 4B

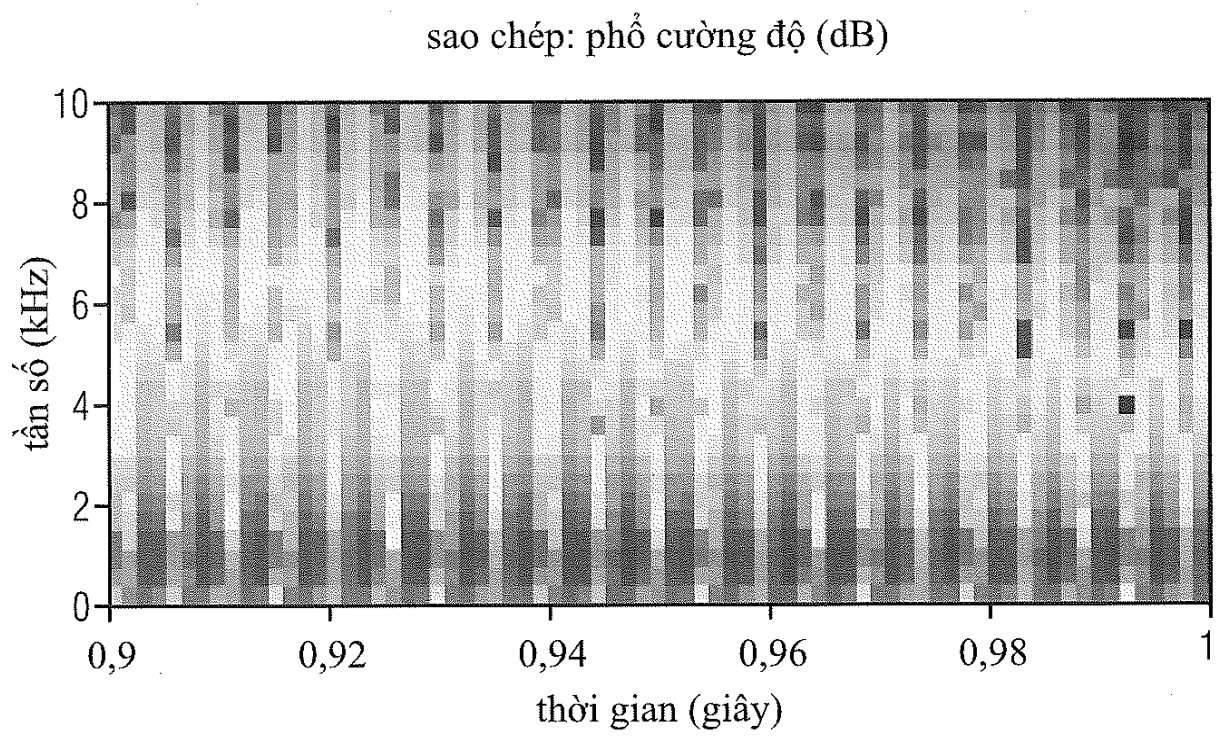


FIG 4C

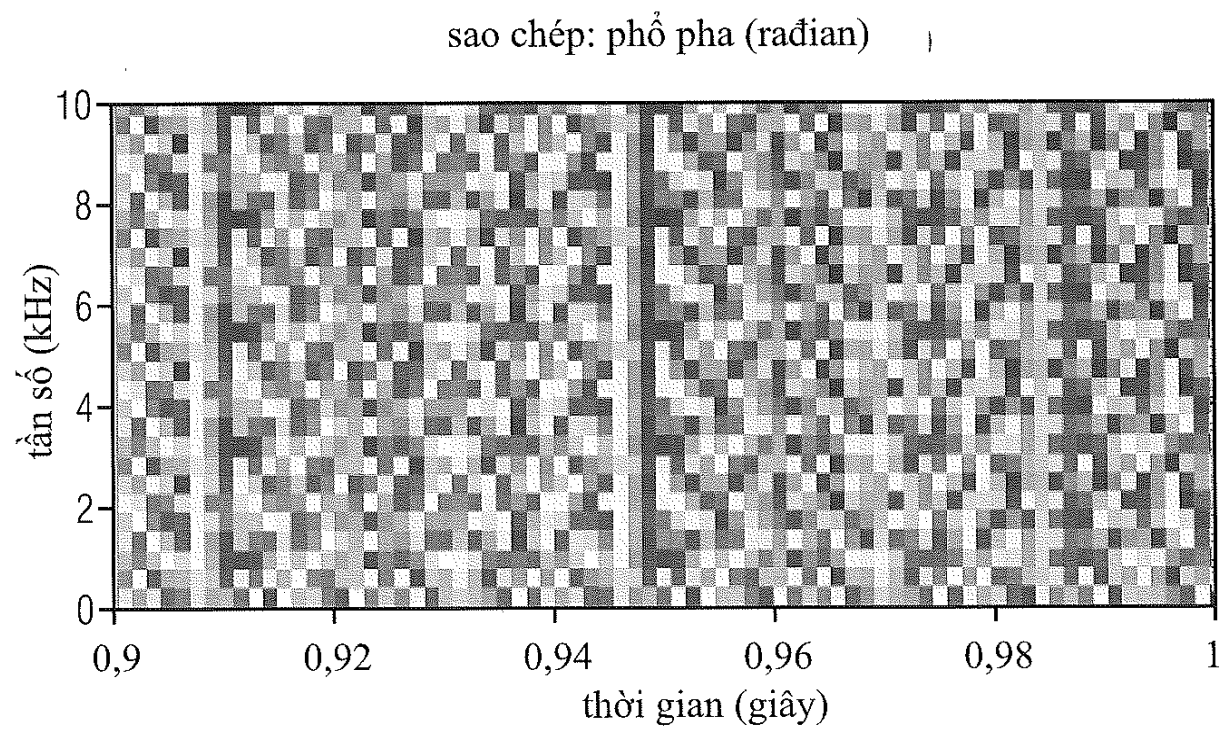


FIG 4D

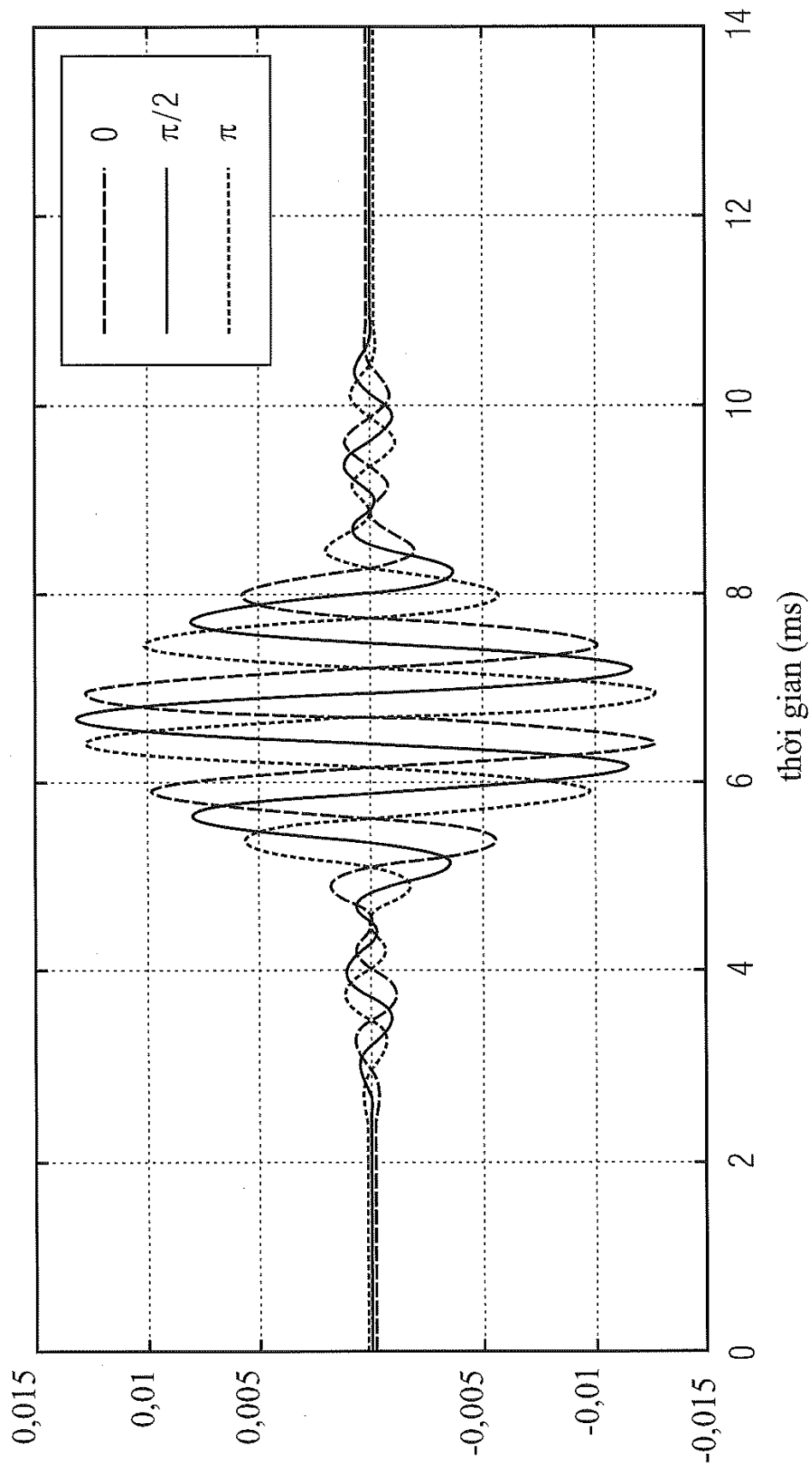


FIG 5

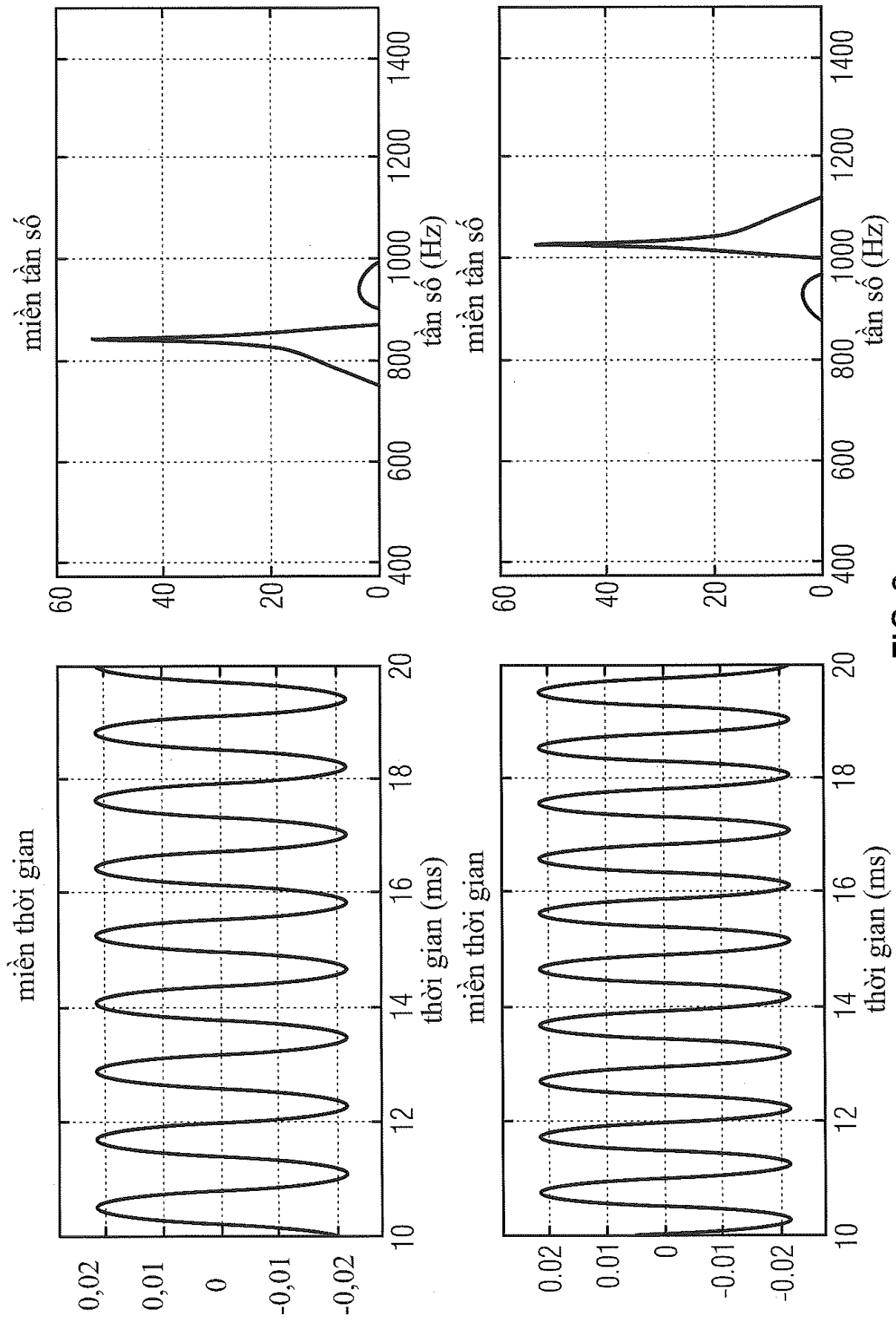


FIG 6

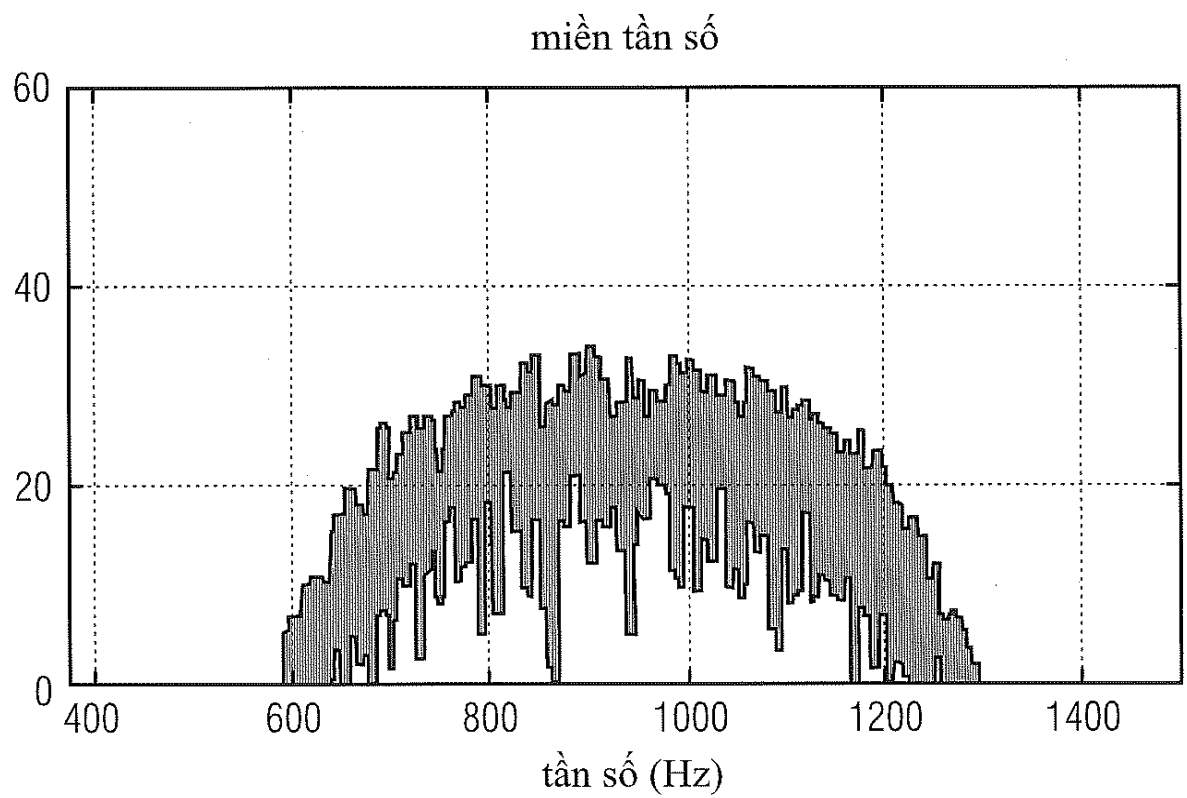
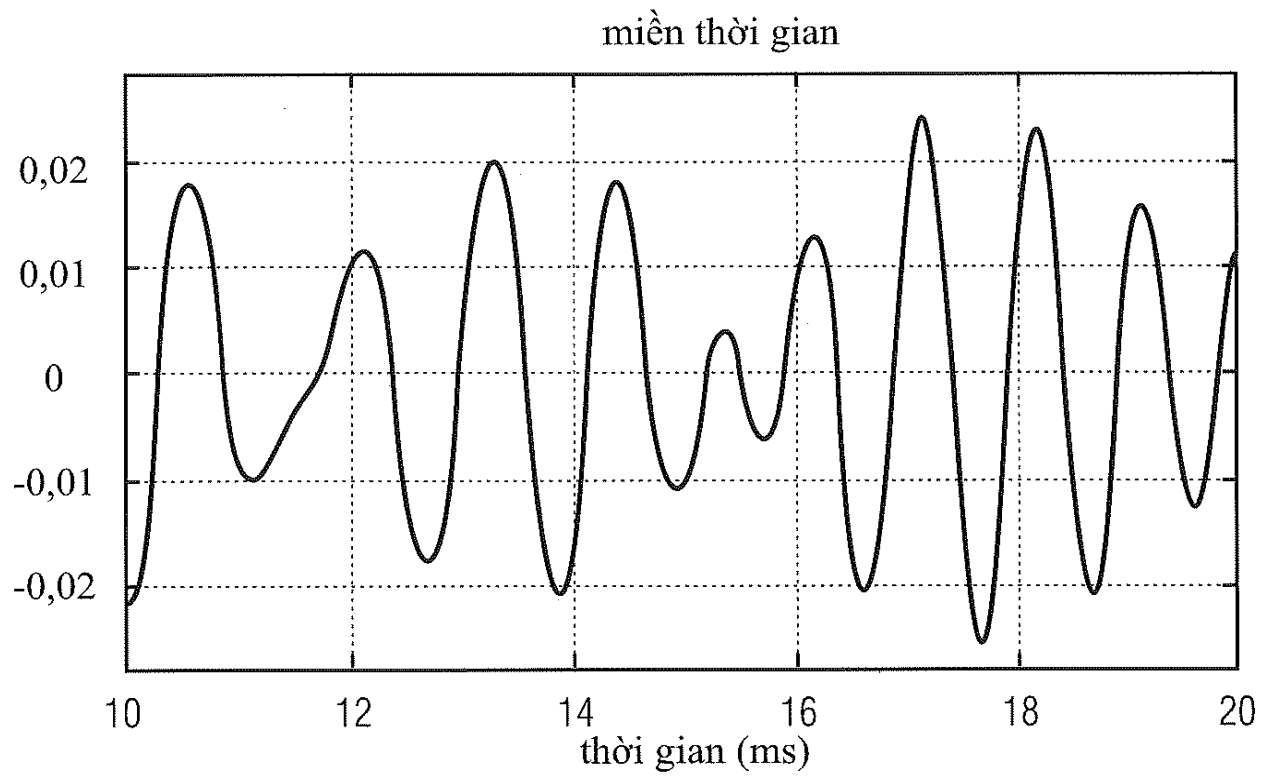


FIG 7

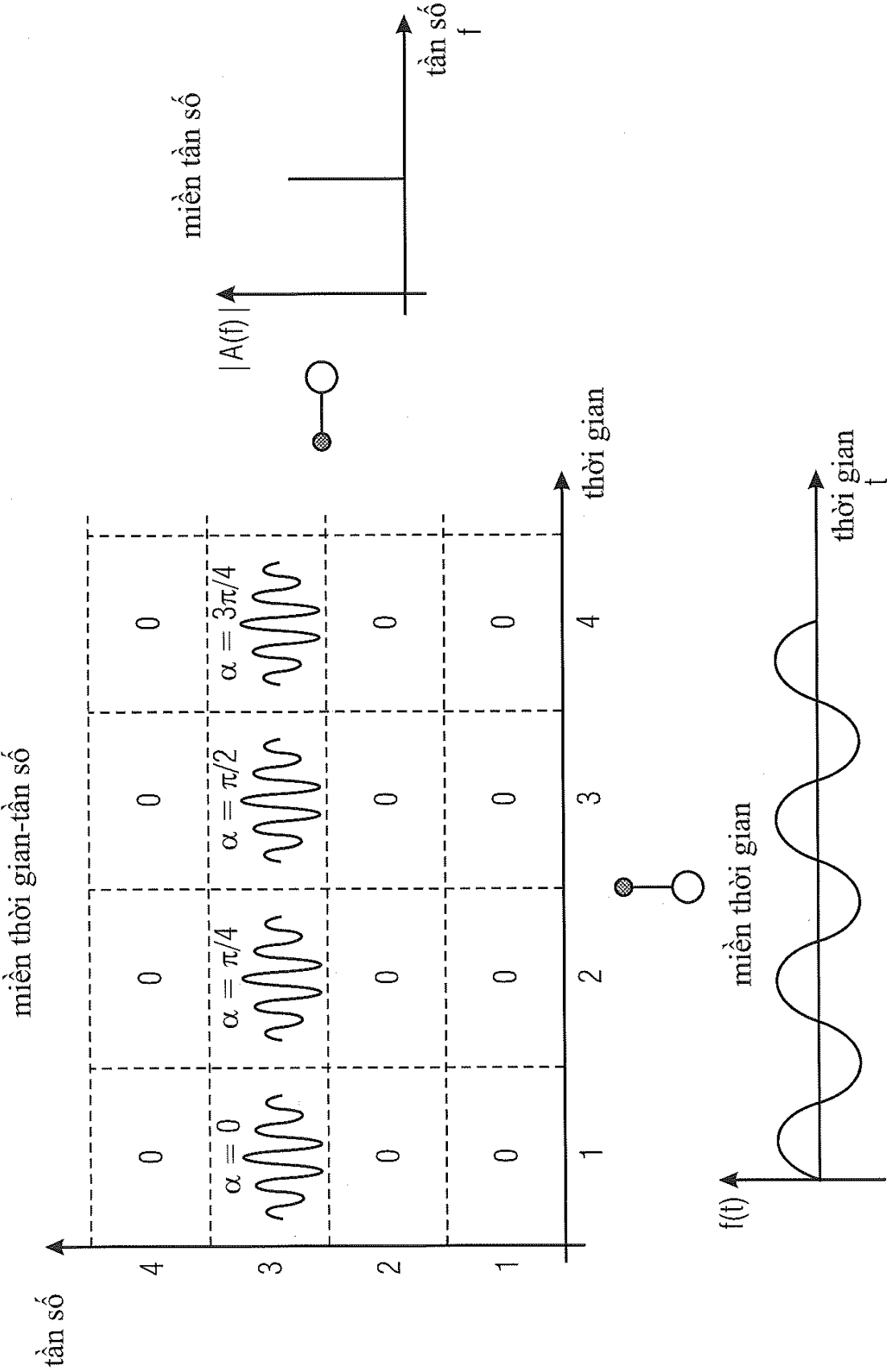


FIG 8

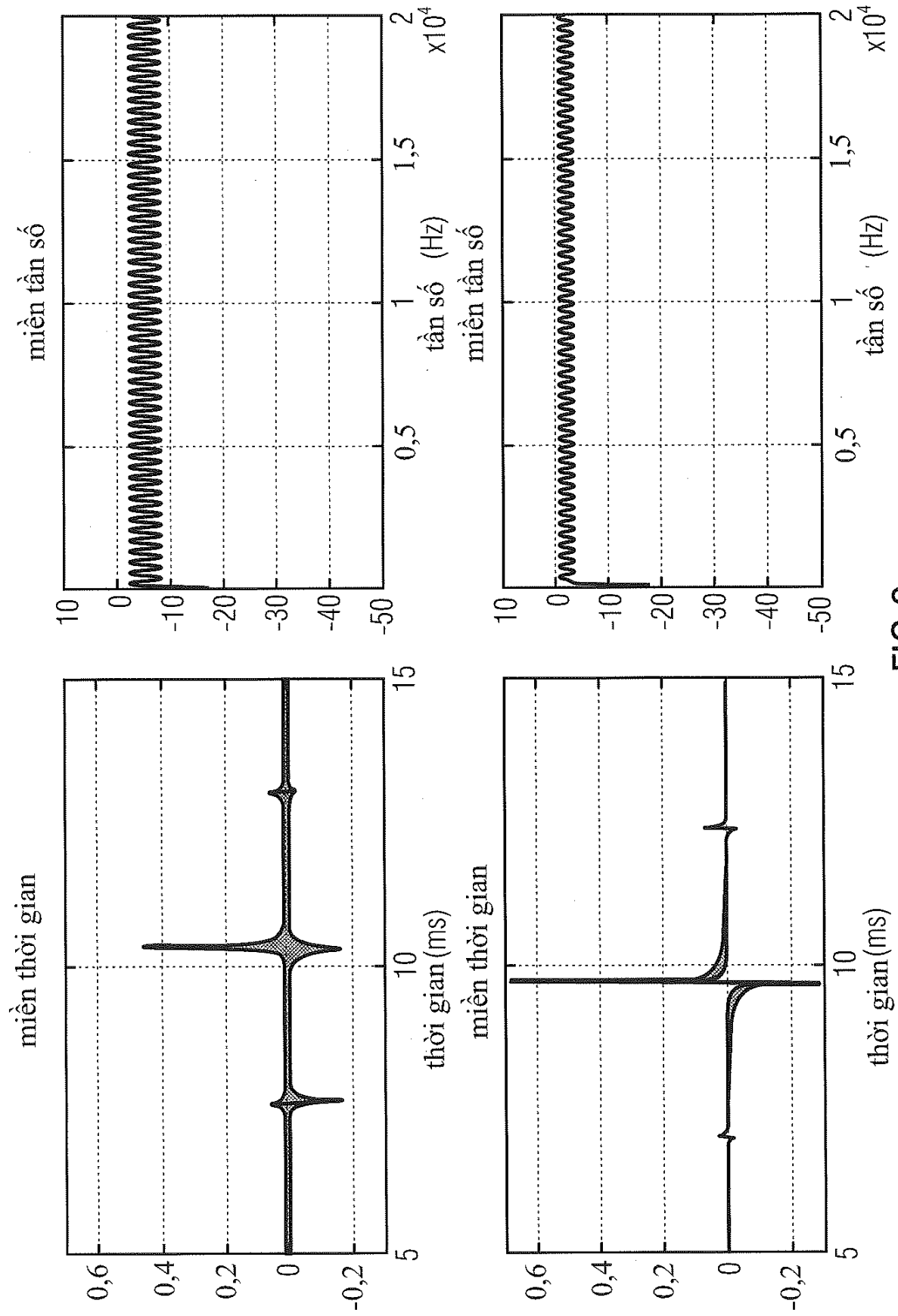


FIG 9

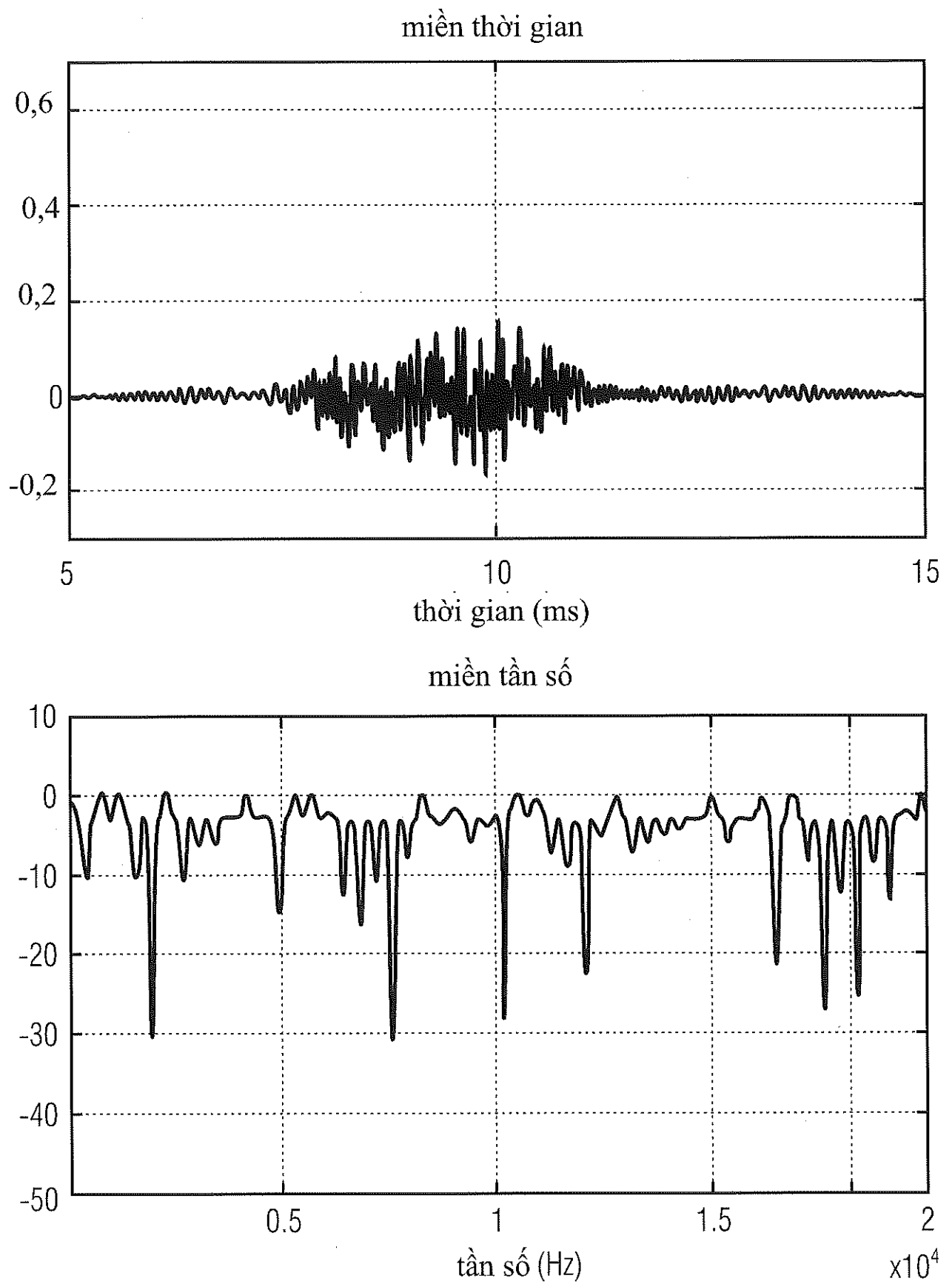


FIG 10

ban đầu: đạo hàm pha theo thời gian (radian)

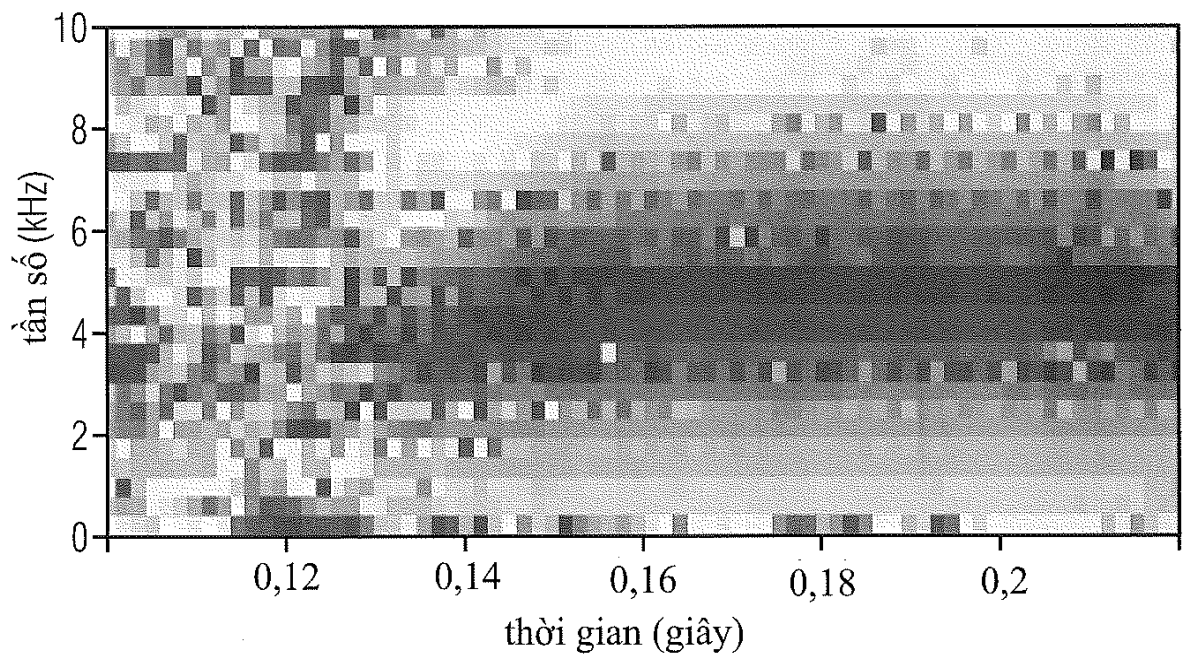


FIG 12A

ban đầu: đạo hàm pha theo tần số (radian)

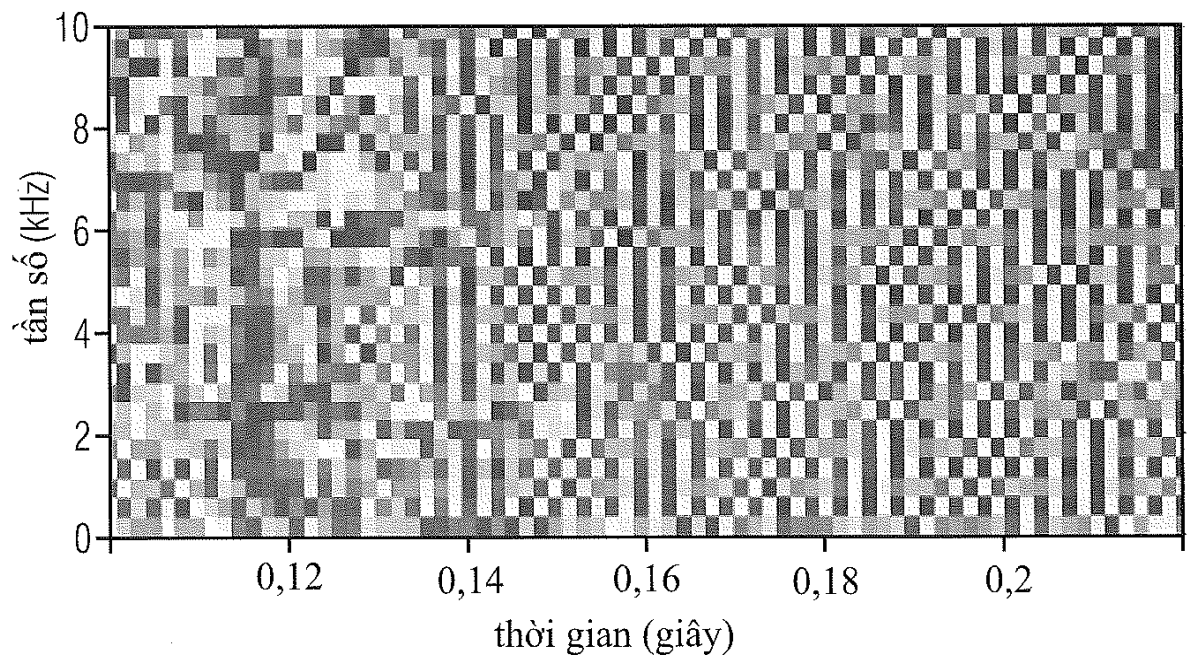


FIG 12B

ban đầu: đạo hàm pha theo thời gian (radian)

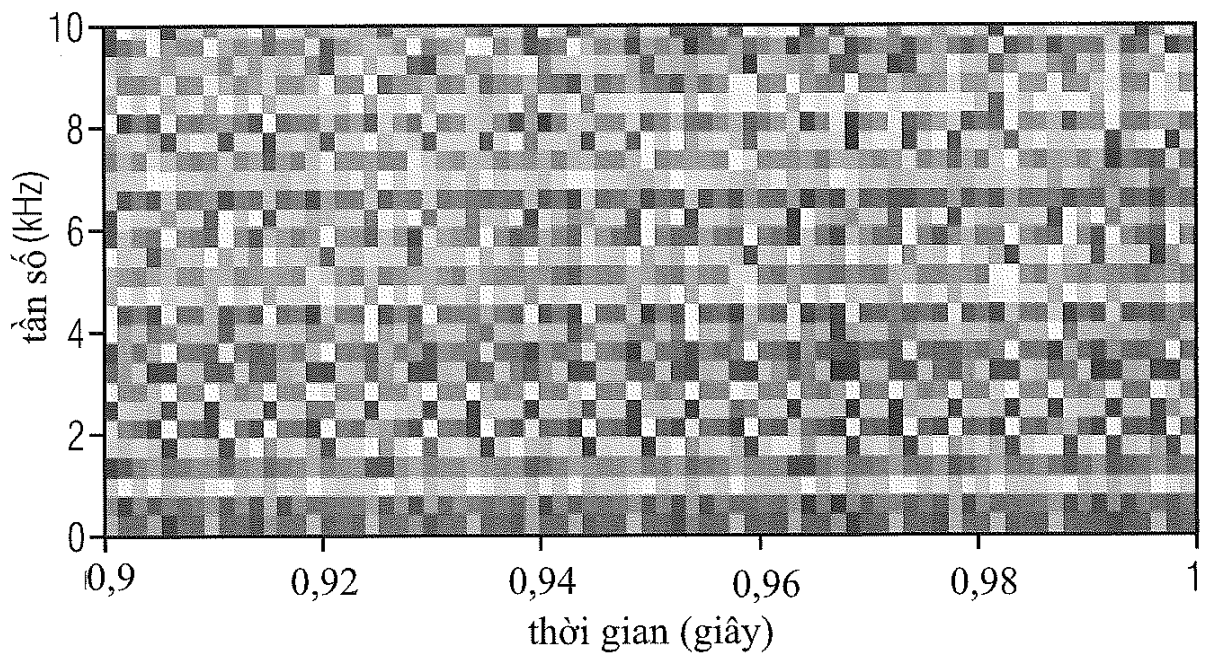


FIG 12C

ban đầu: đạo hàm pha theo tần số (radian)

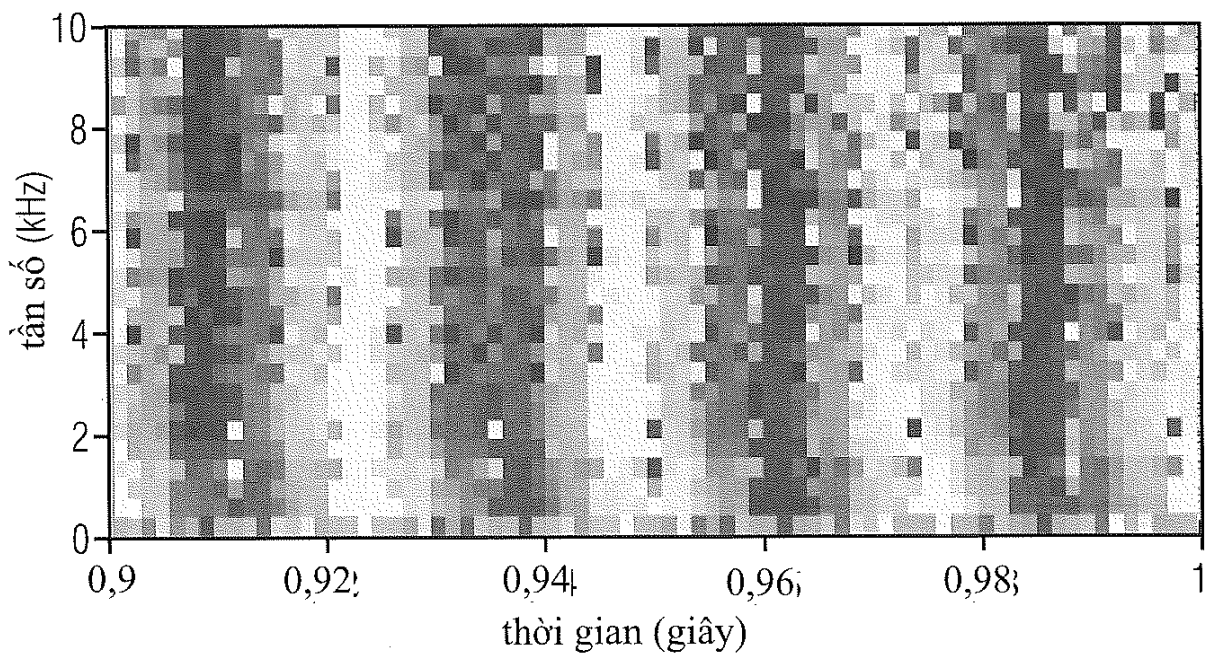


FIG 12D

sao chép: đạo hàm pha theo thời gian (radian)

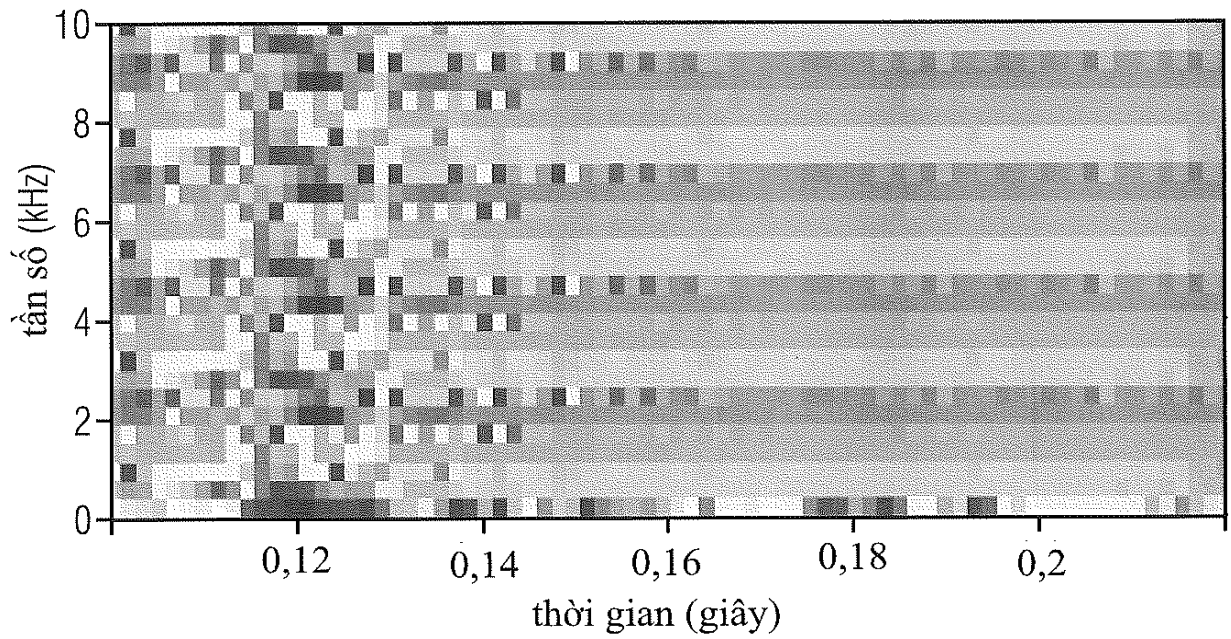


FIG 13A

sao chép: đạo hàm pha theo tần số (radian)

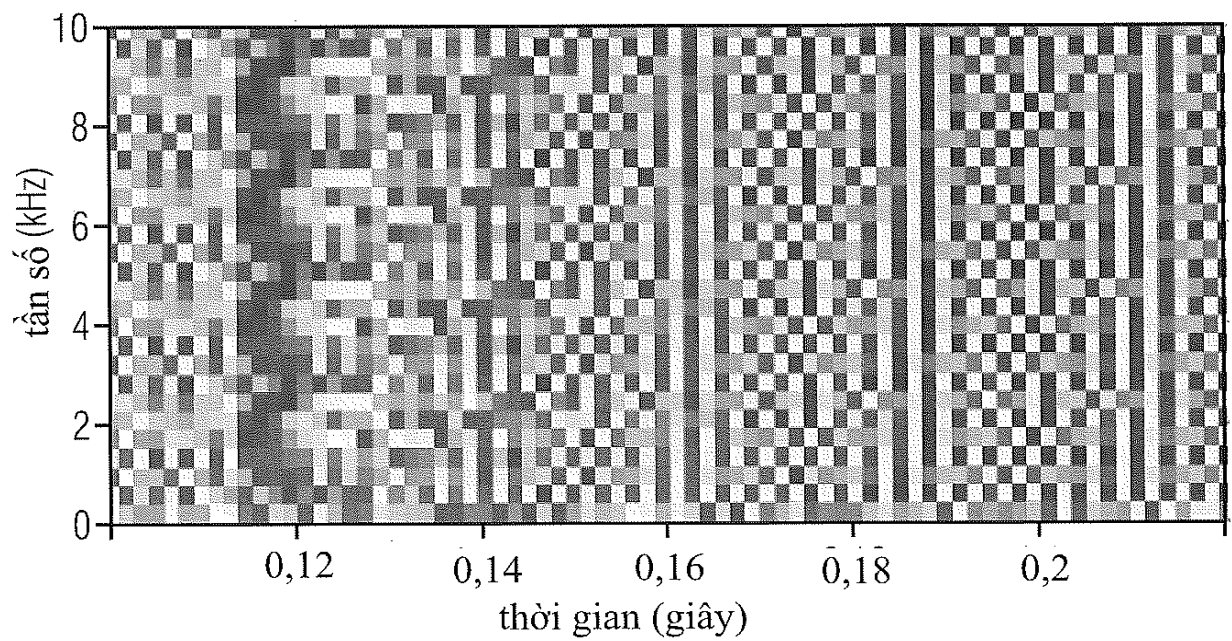


FIG 13B

sao chép: đạo hàm pha theo thời gian (radian)

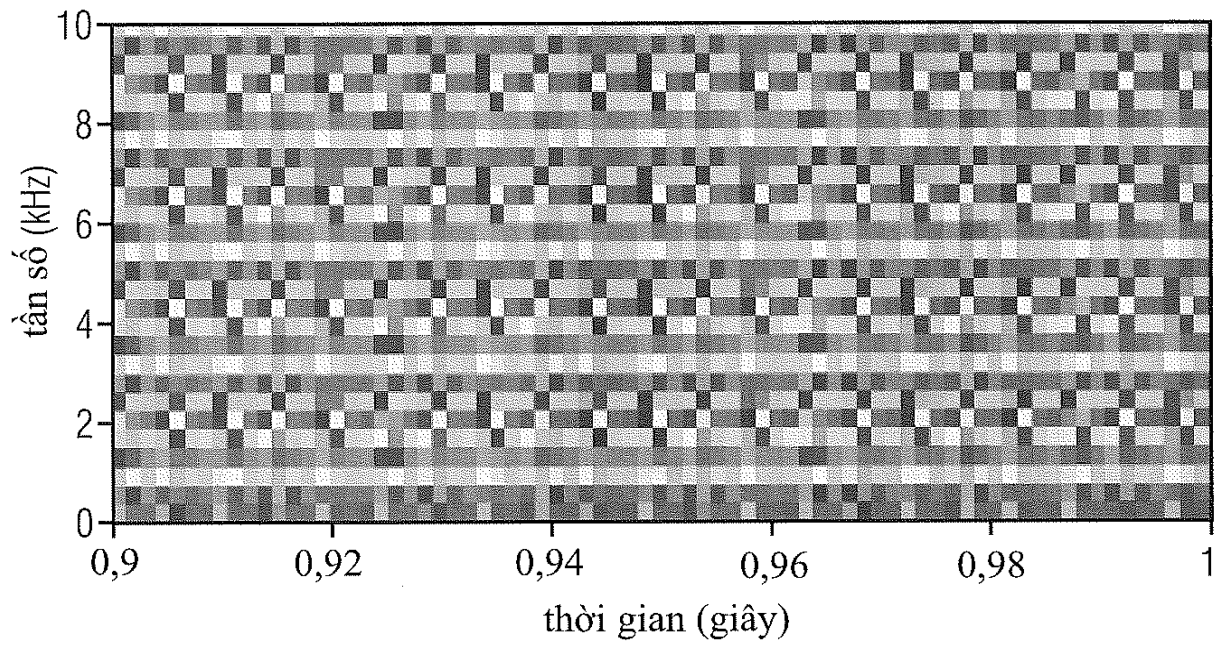


FIG 13C

sao chép: đạo hàm pha theo tần số (radian)

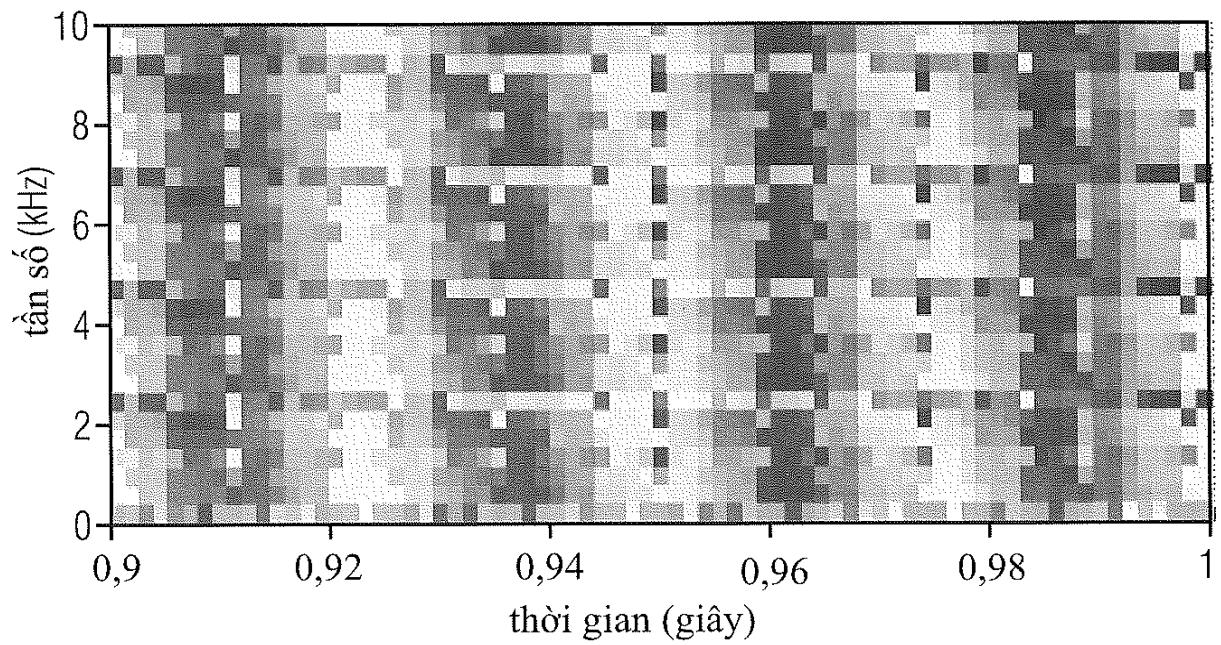


FIG 13D

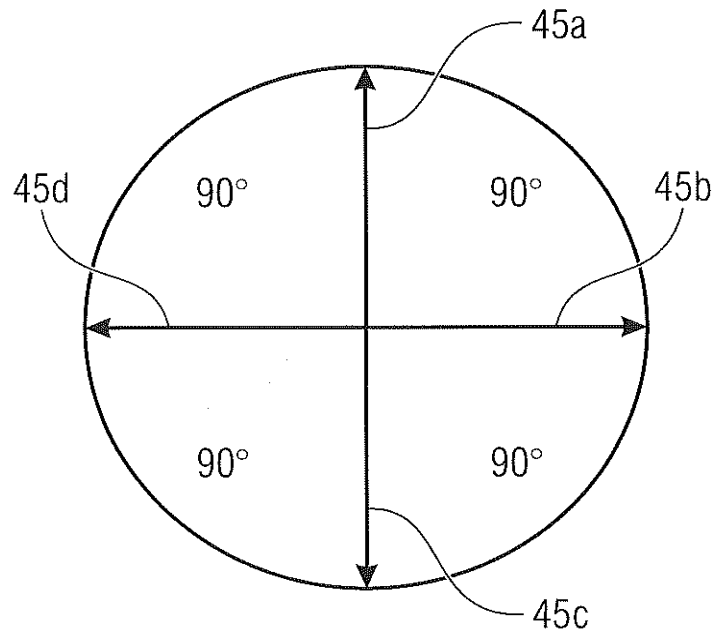


FIG 14A

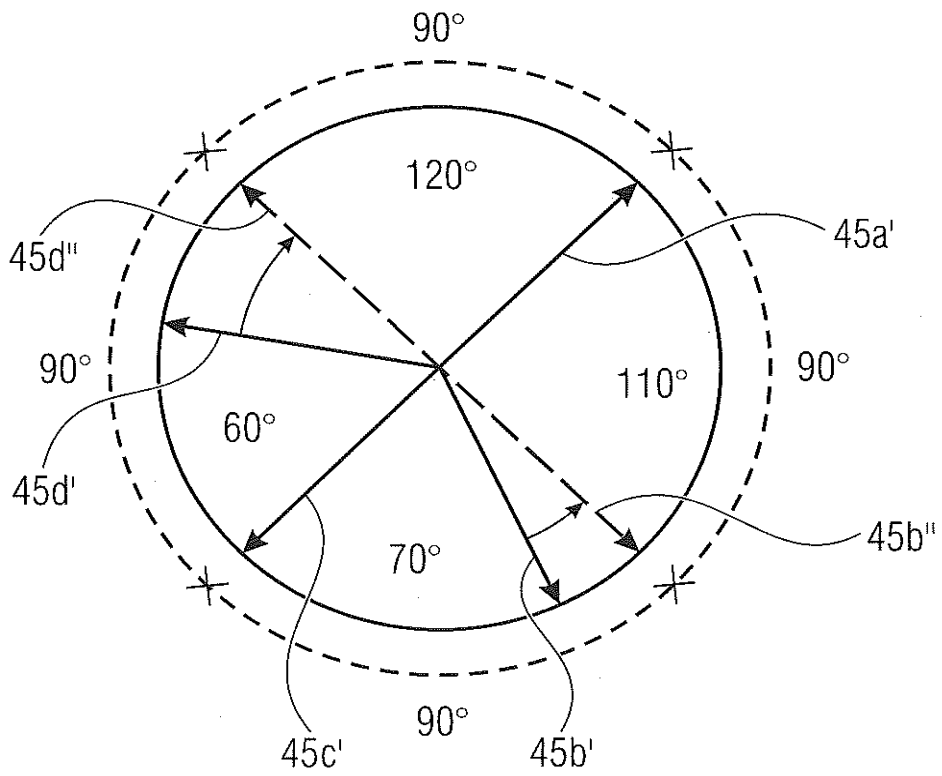


FIG 14B

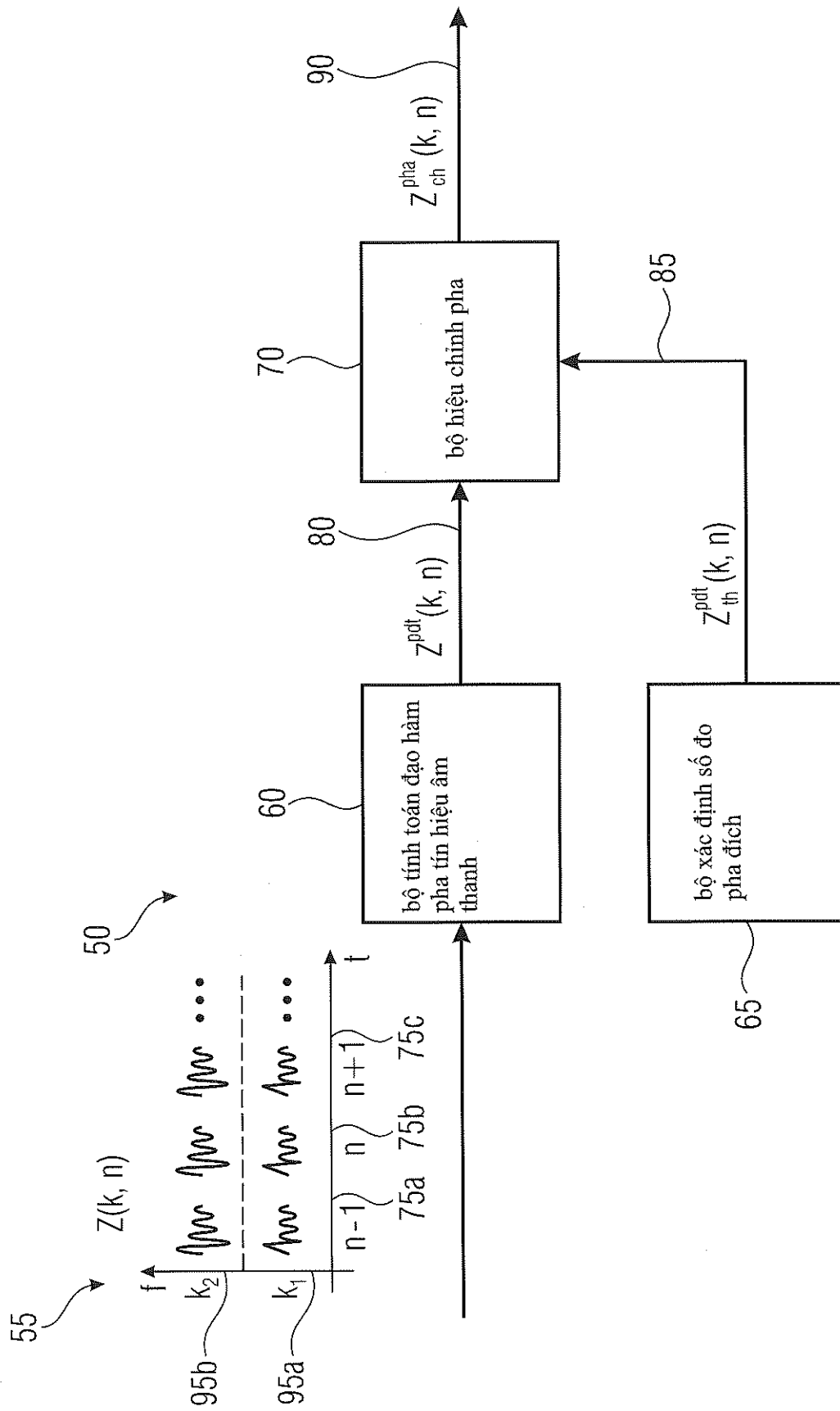


FIG 15

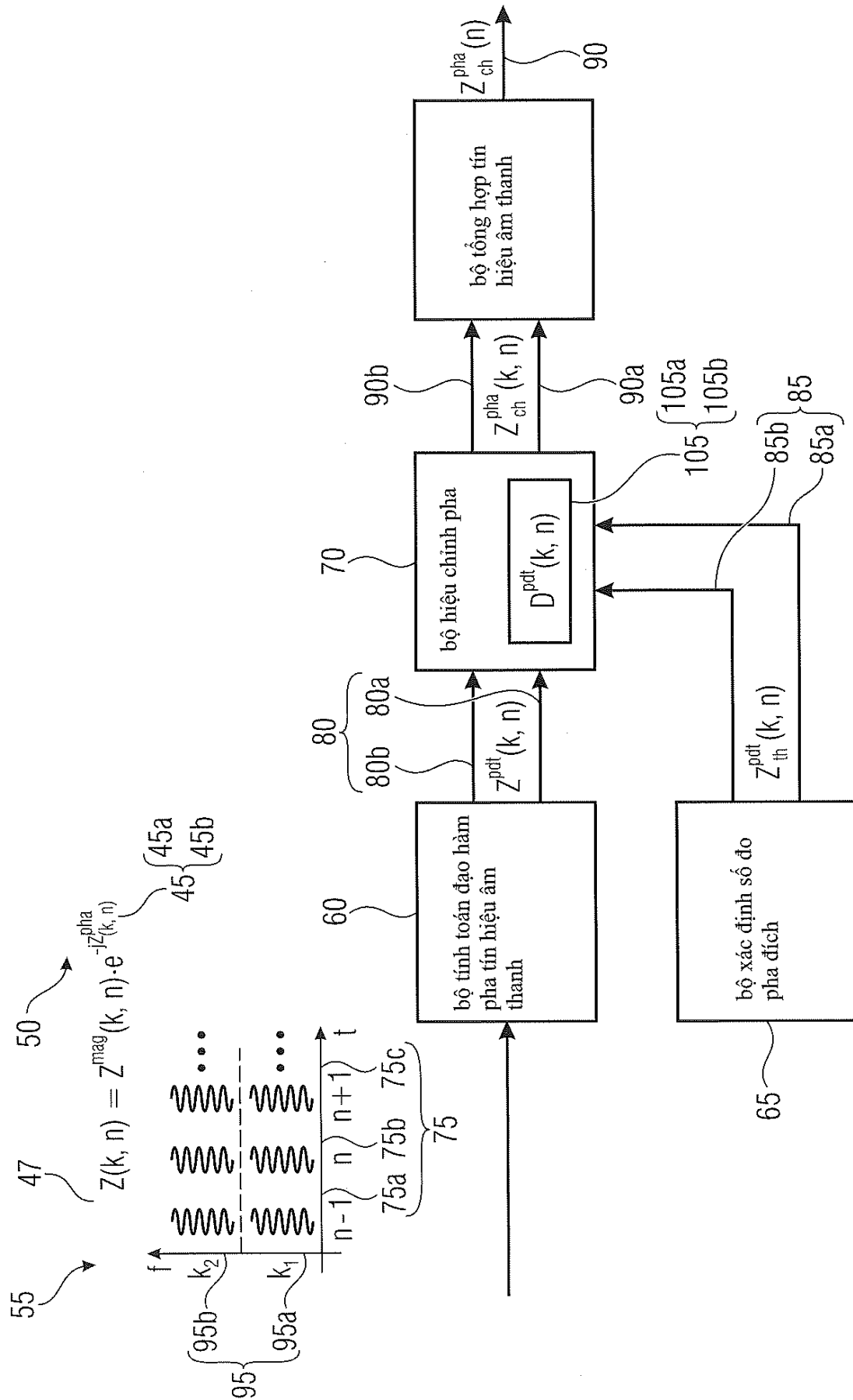


FIG 16

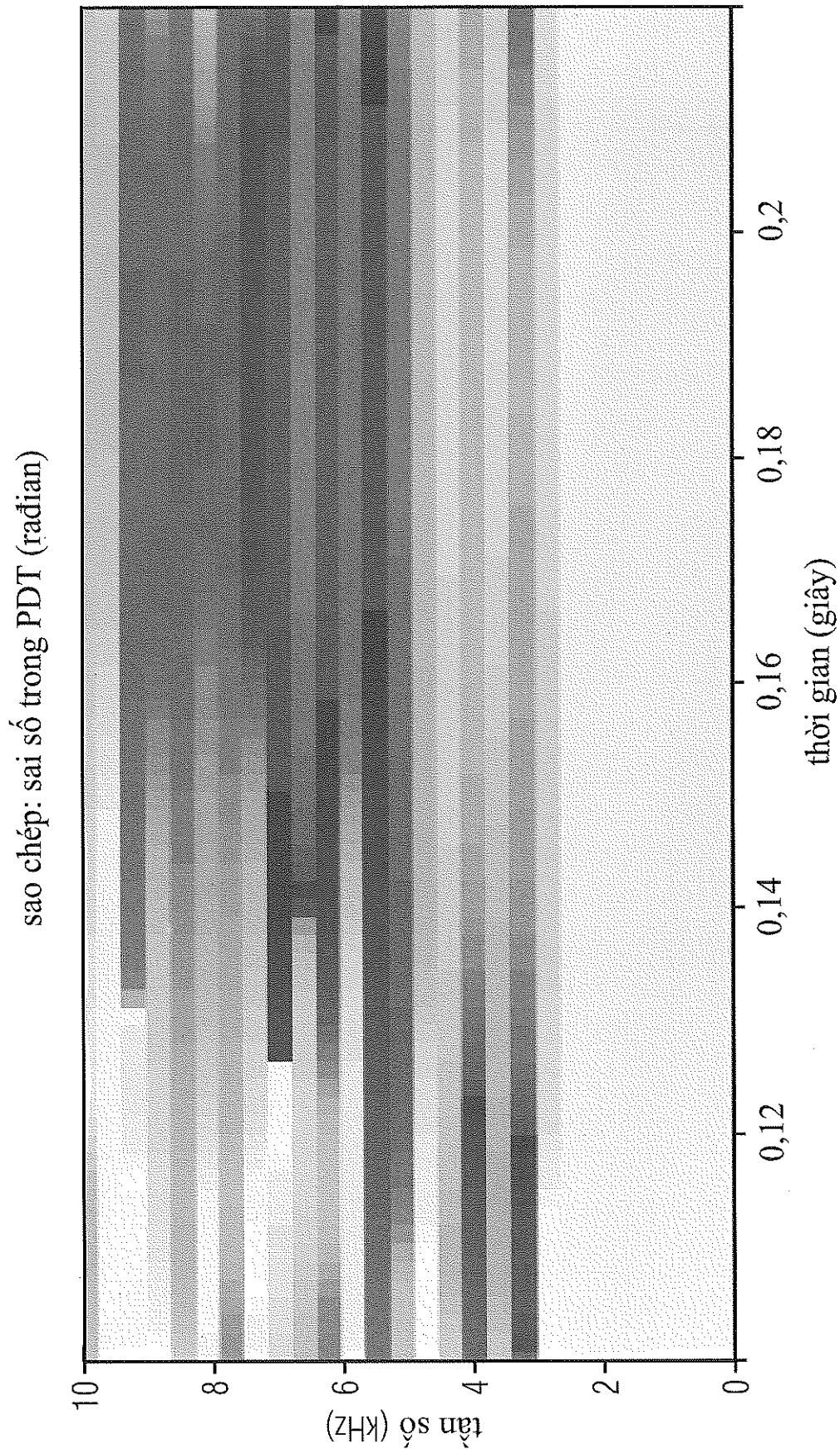


FIG 17

được hiệu chỉnh: sai số trong PDT (radian)

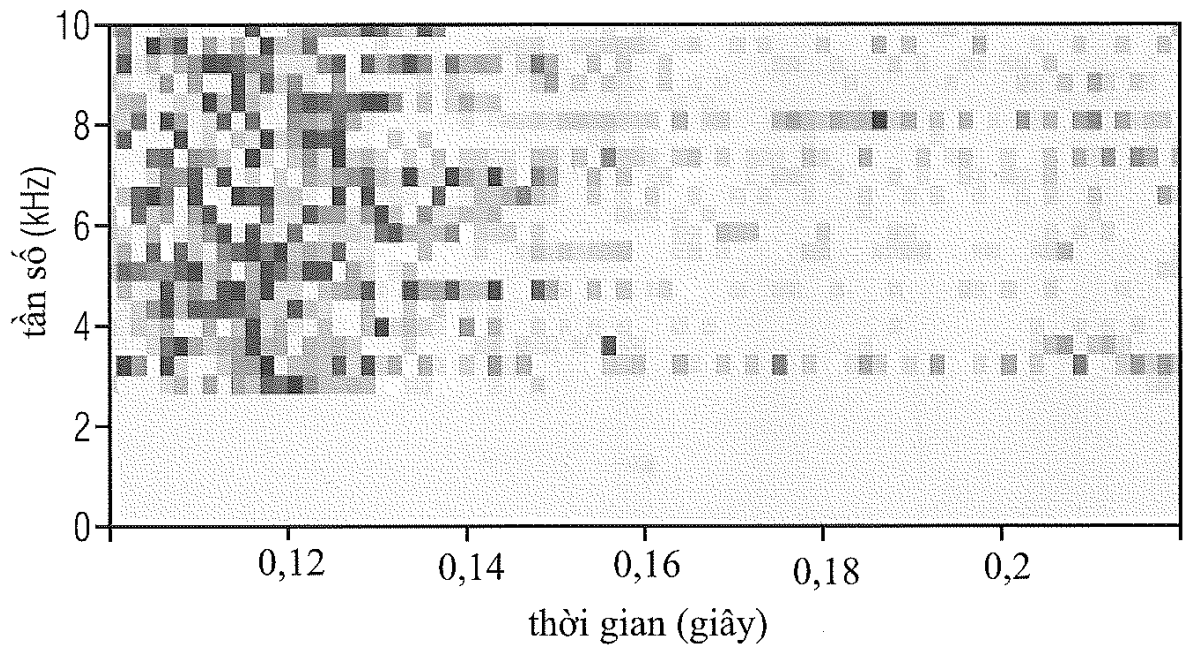


FIG 18A

được hiệu chỉnh: đạo hàm pha theo thời gian (radian)

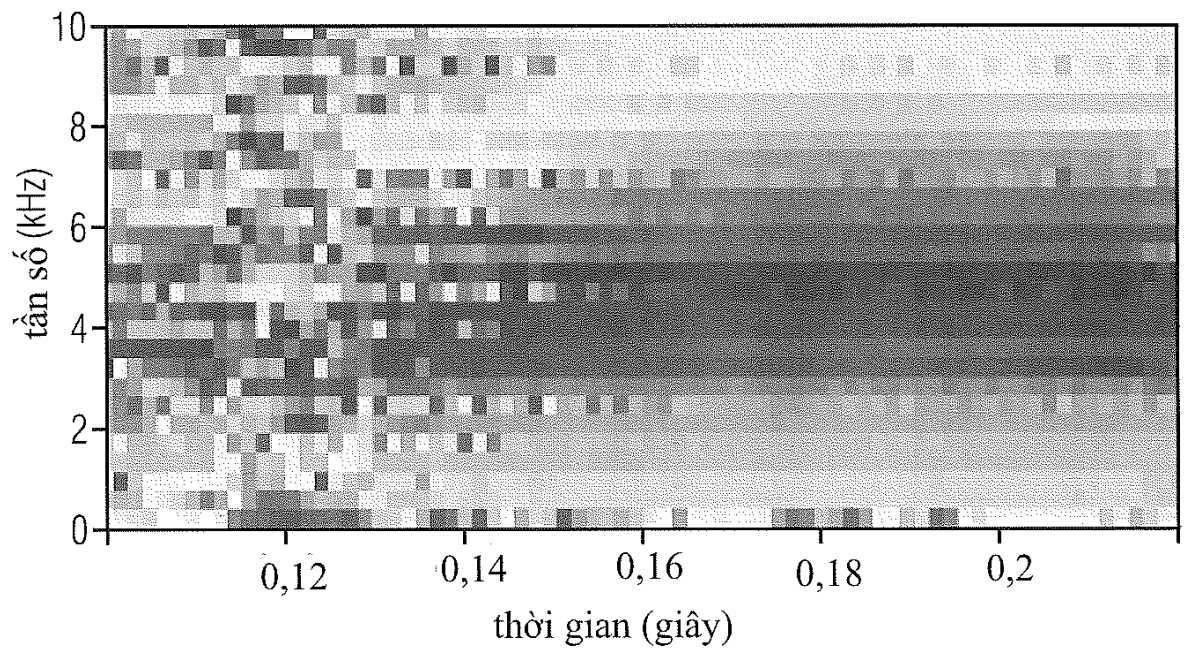


FIG 18B

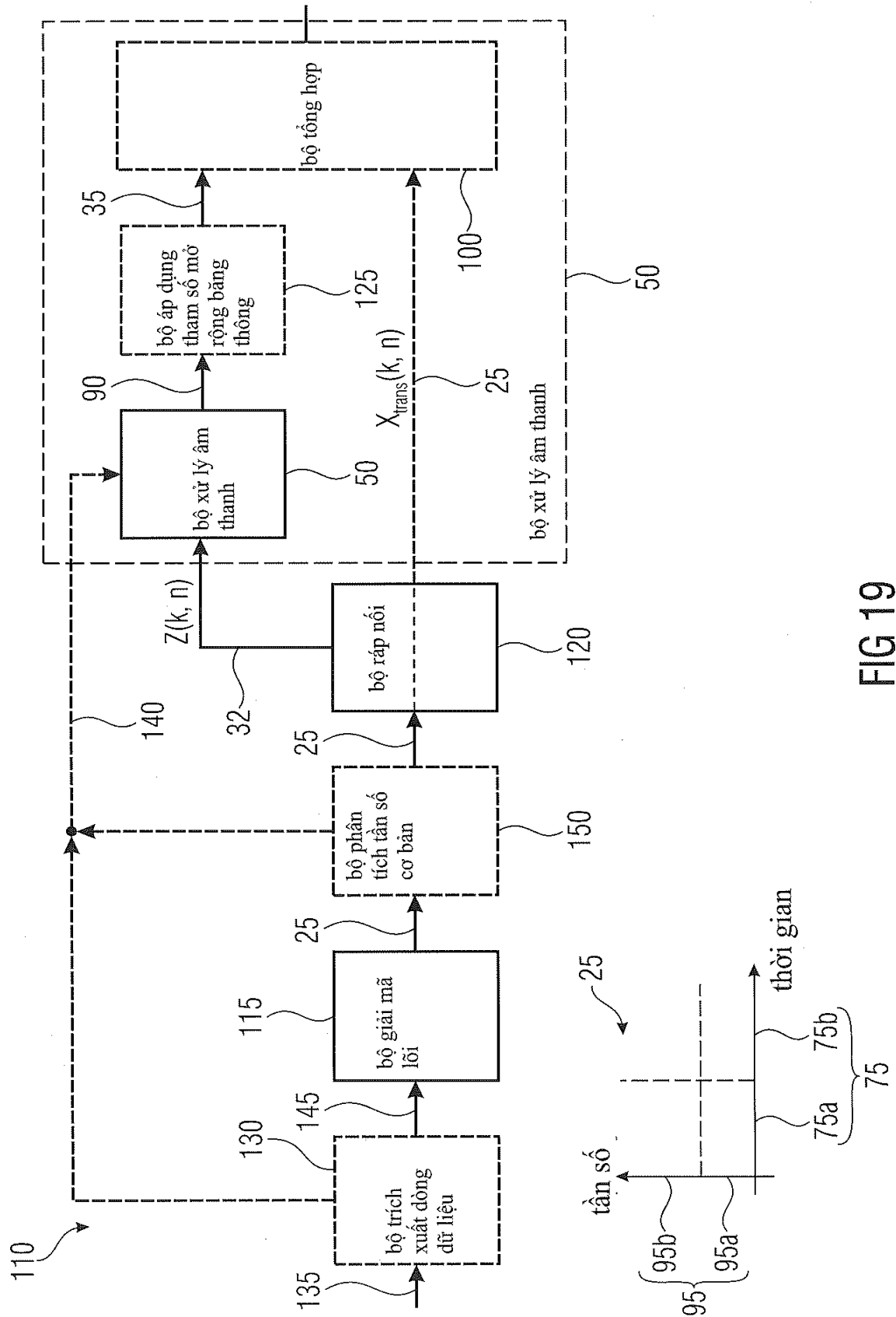


FIG 19

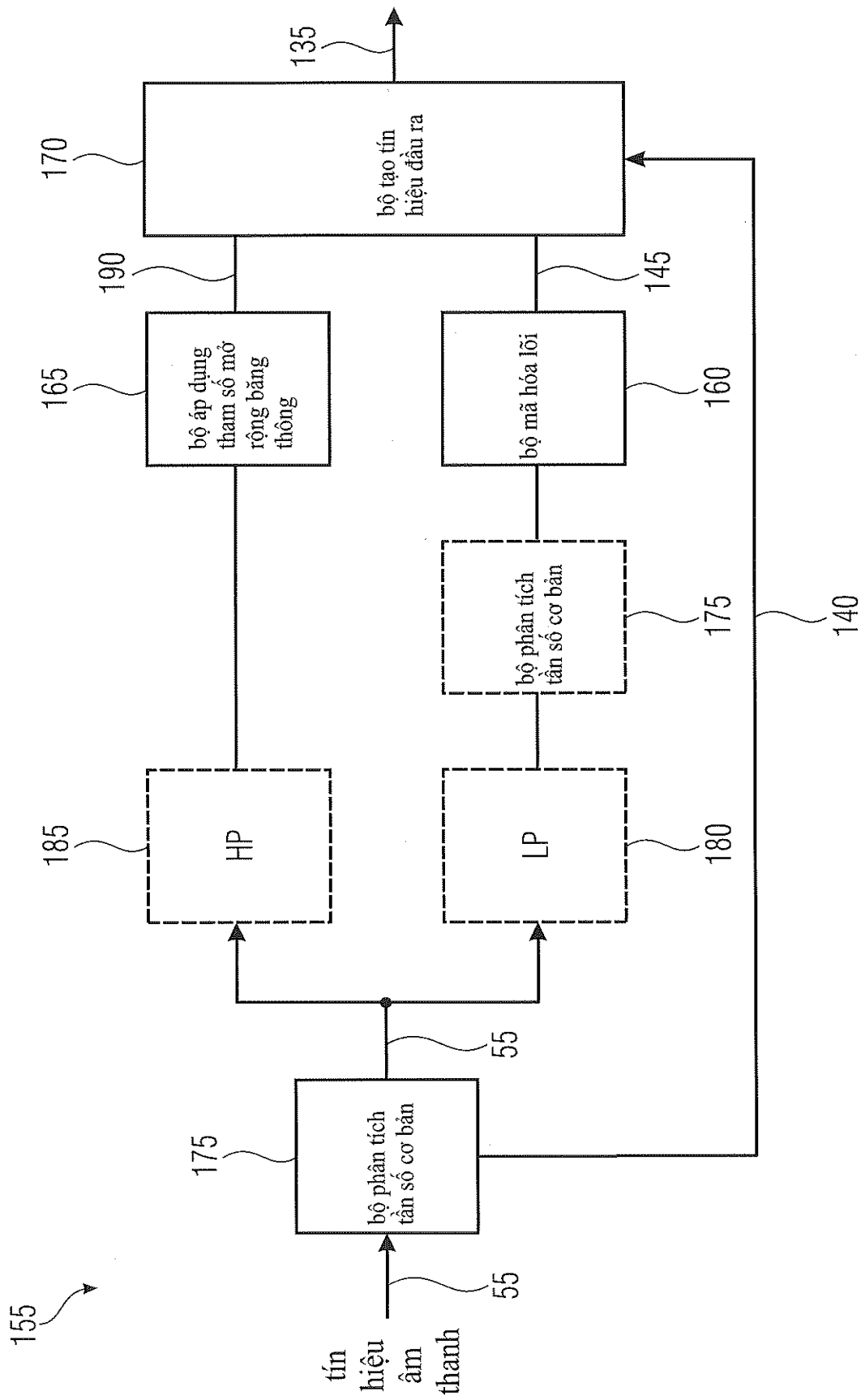


FIG 20

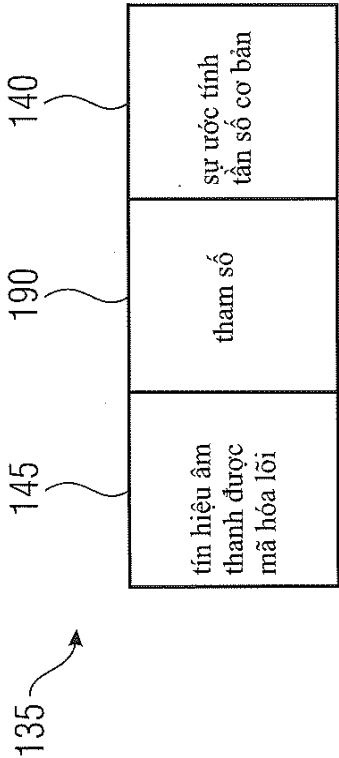


FIG 21

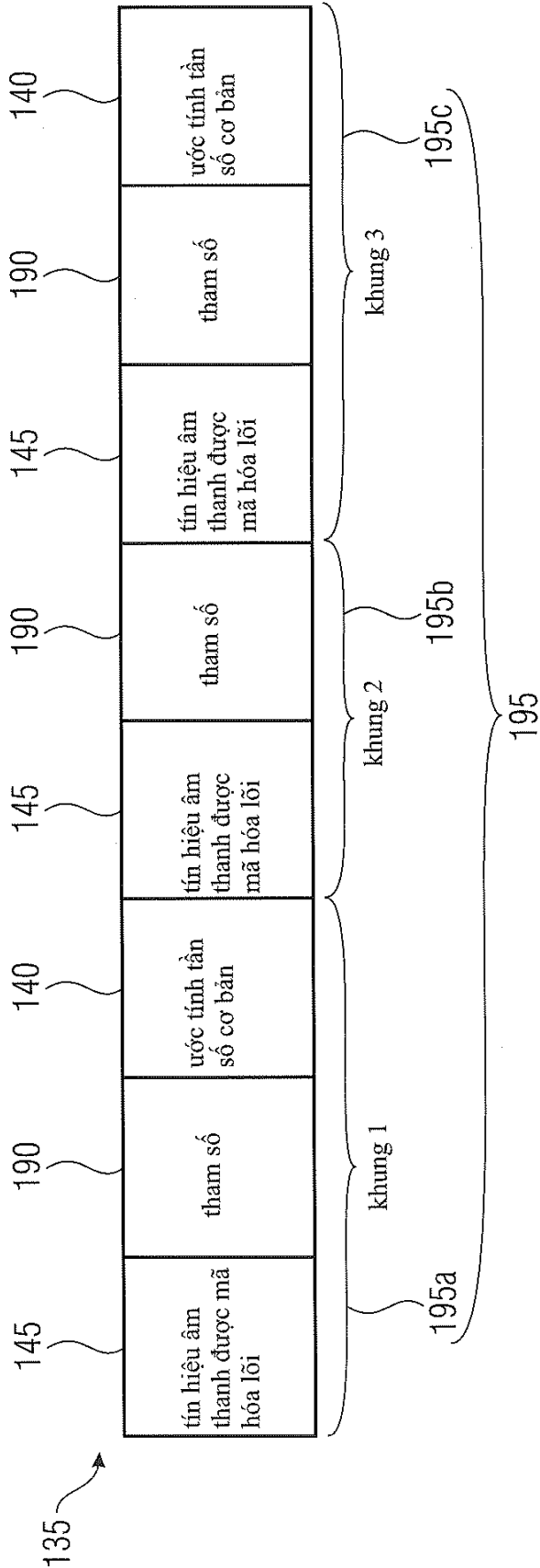


FIG 22

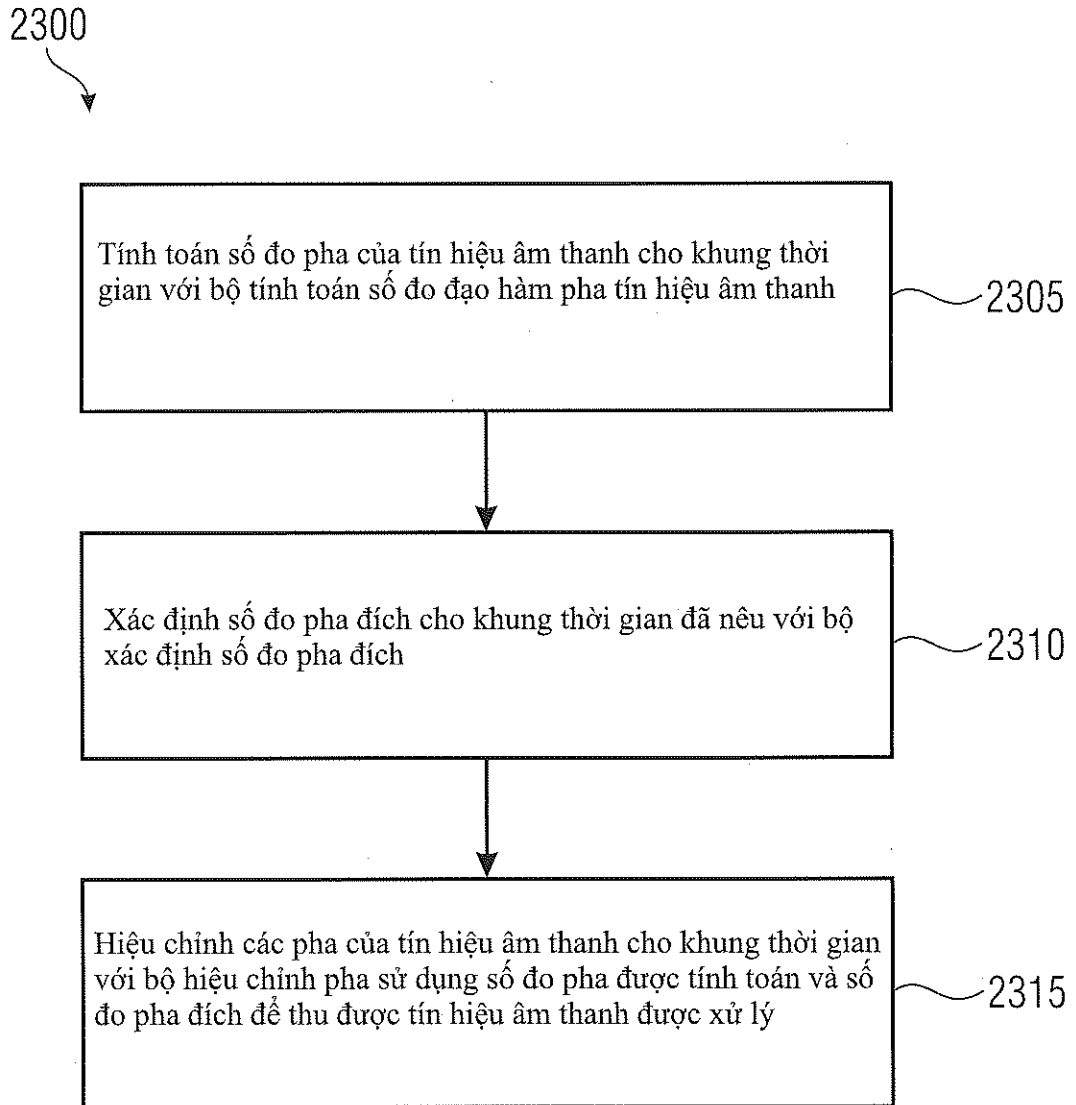


FIG 23

2400

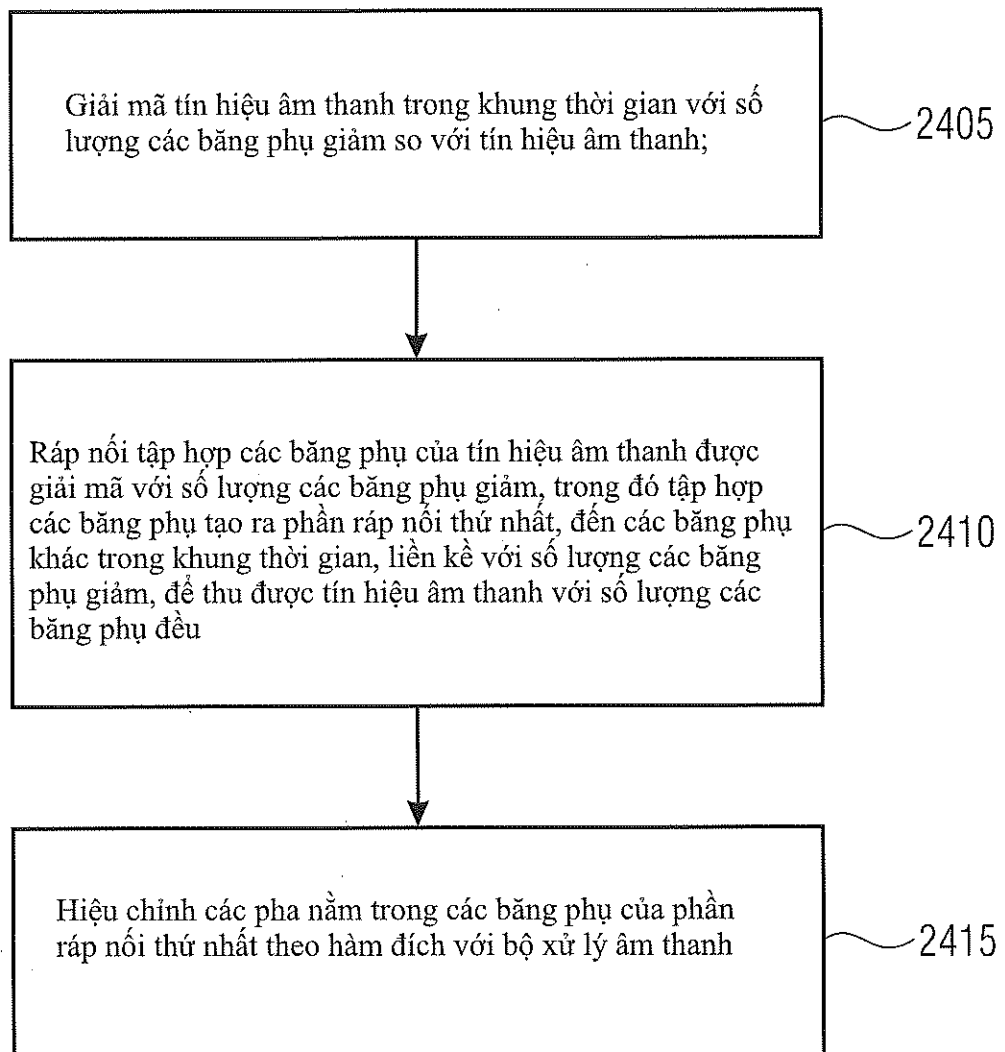


FIG 24

2500

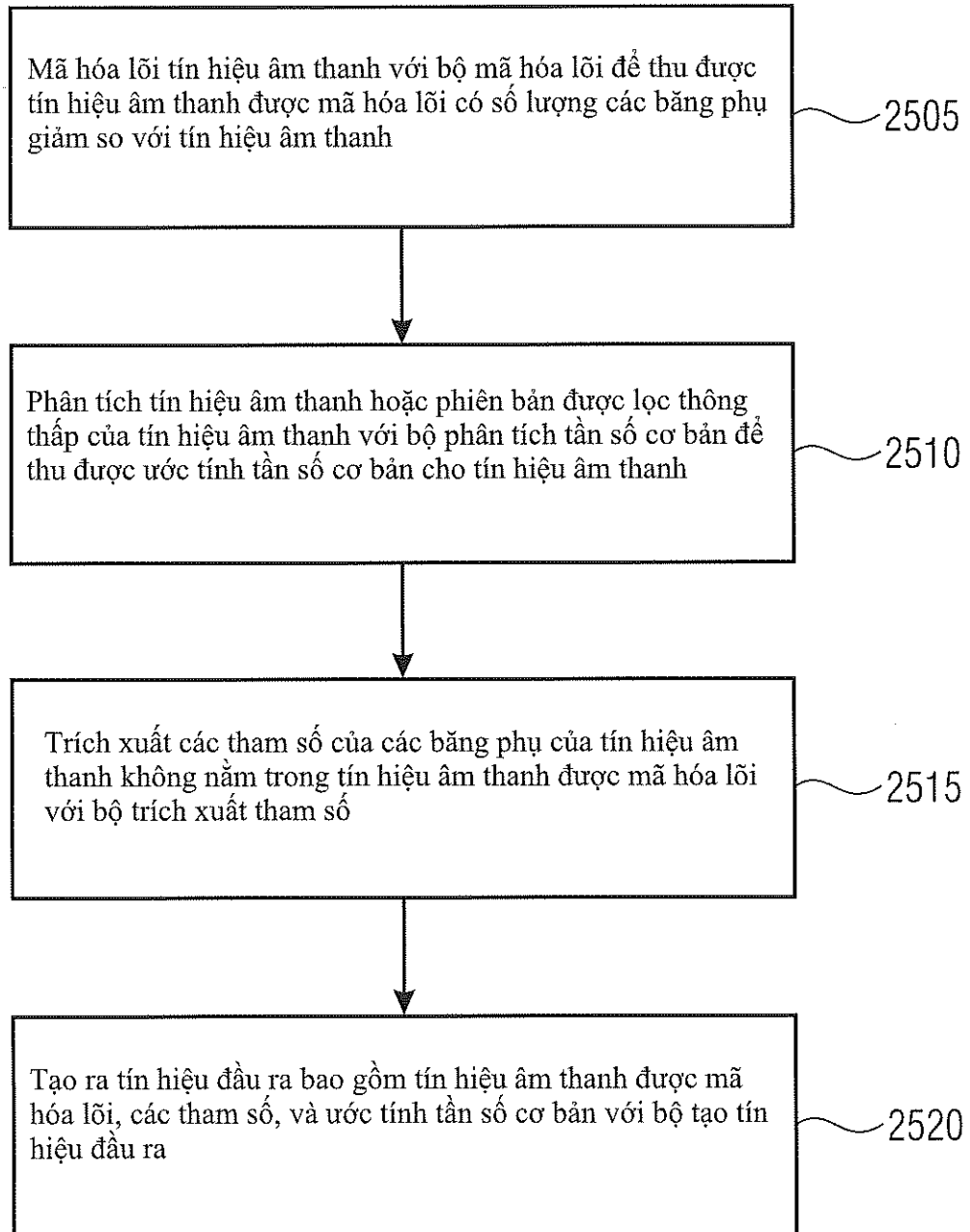


FIG 25

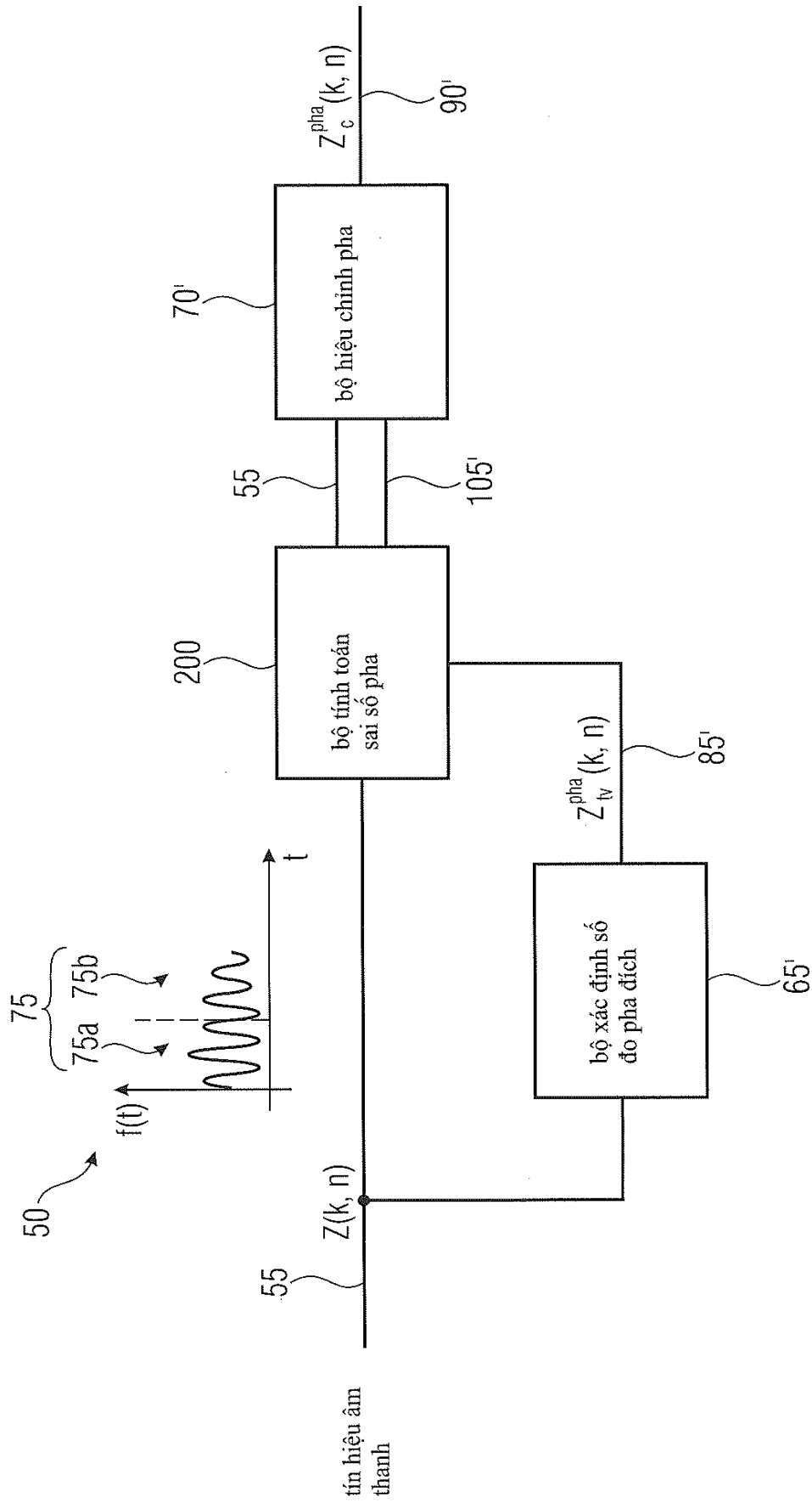


FIG 26

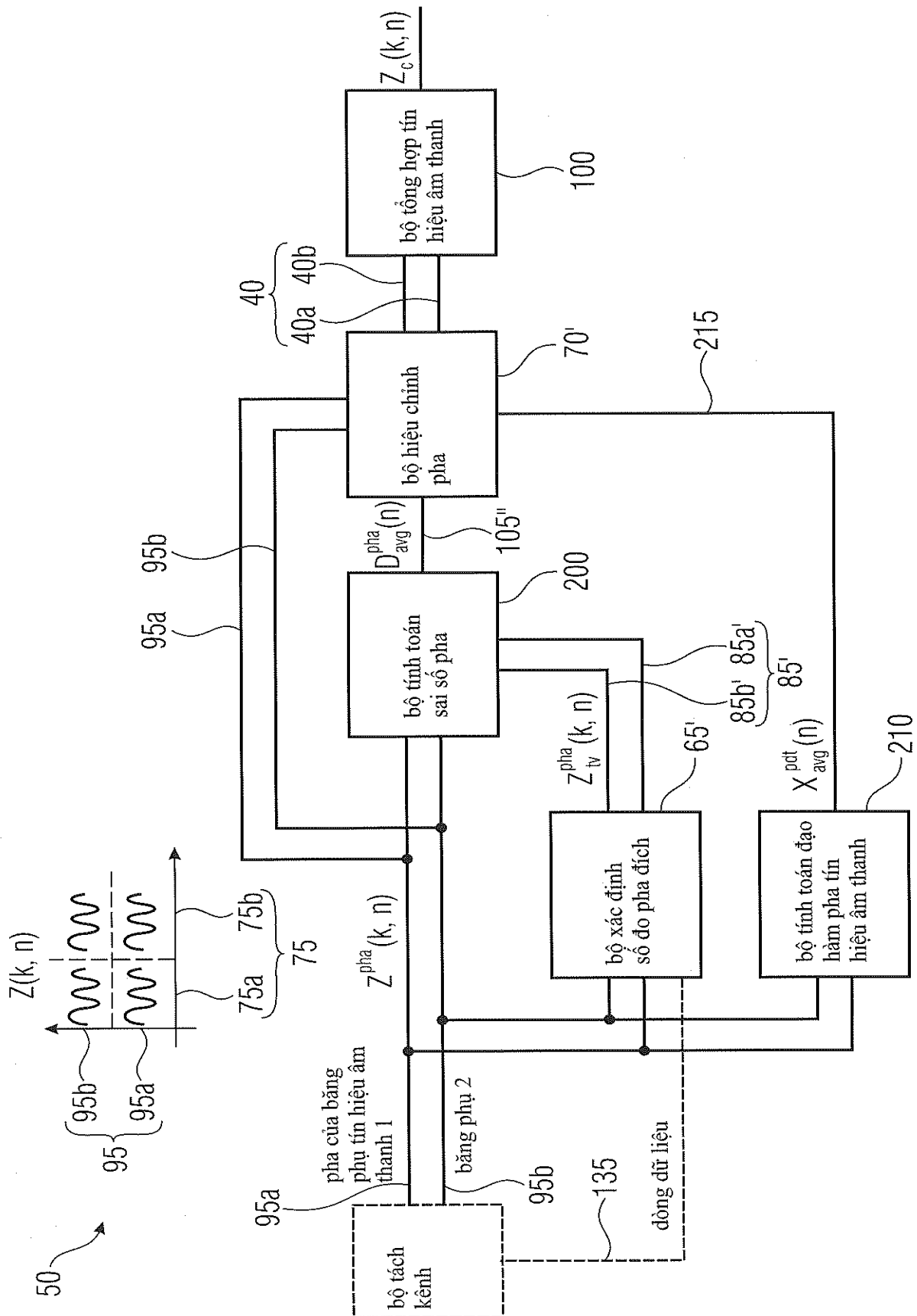


FIG 27

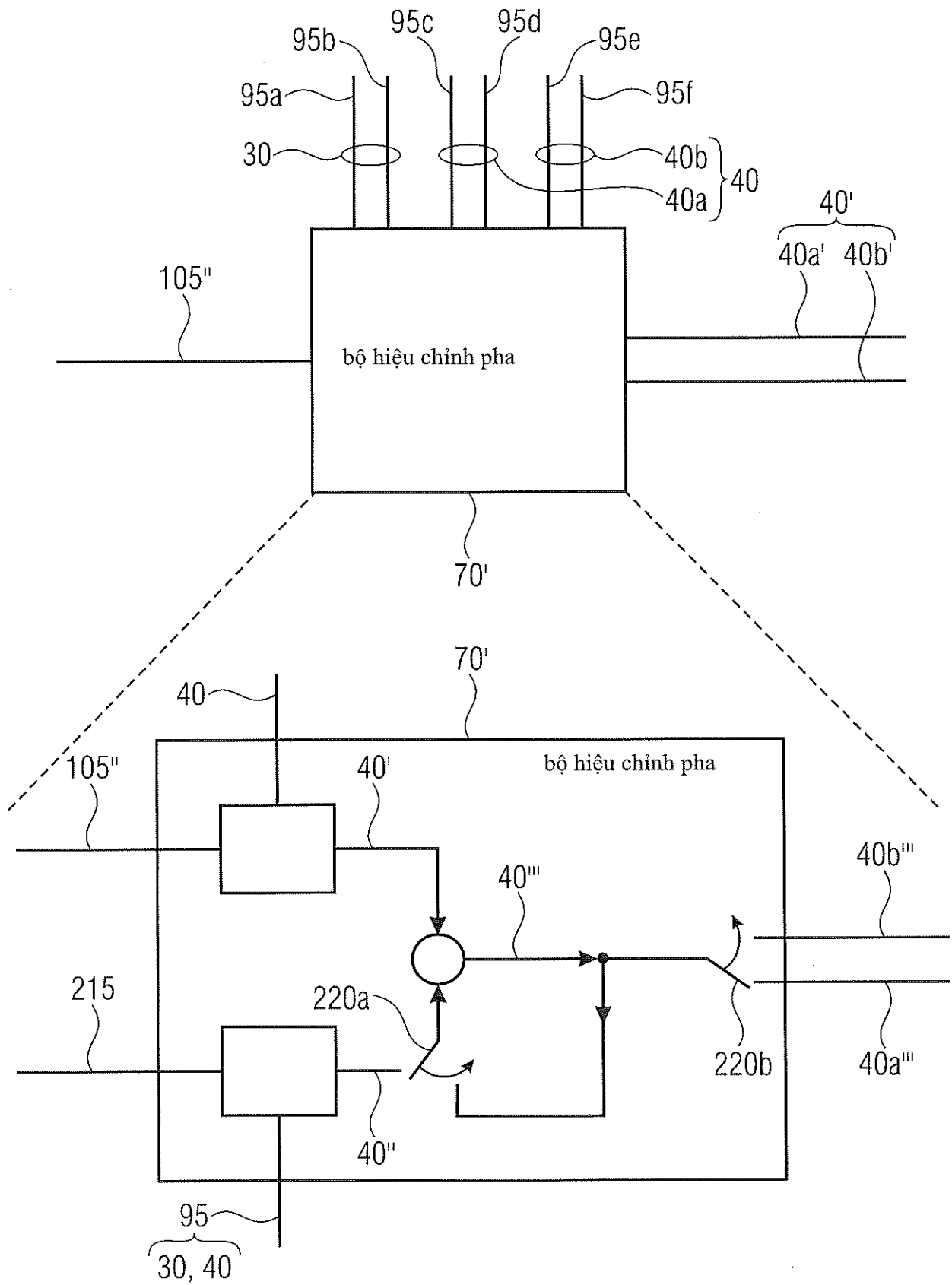


FIG 28A

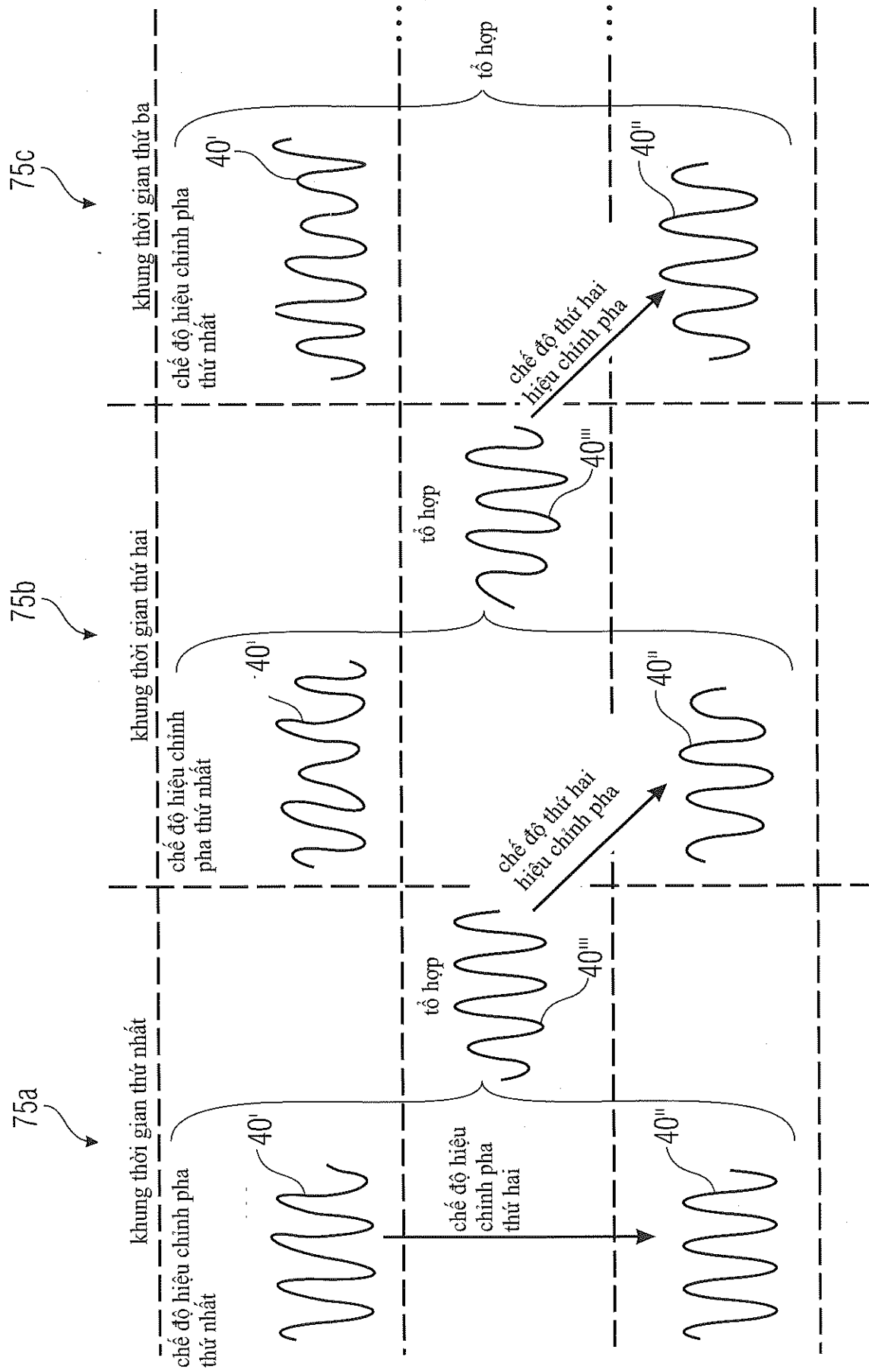


FIG 28B

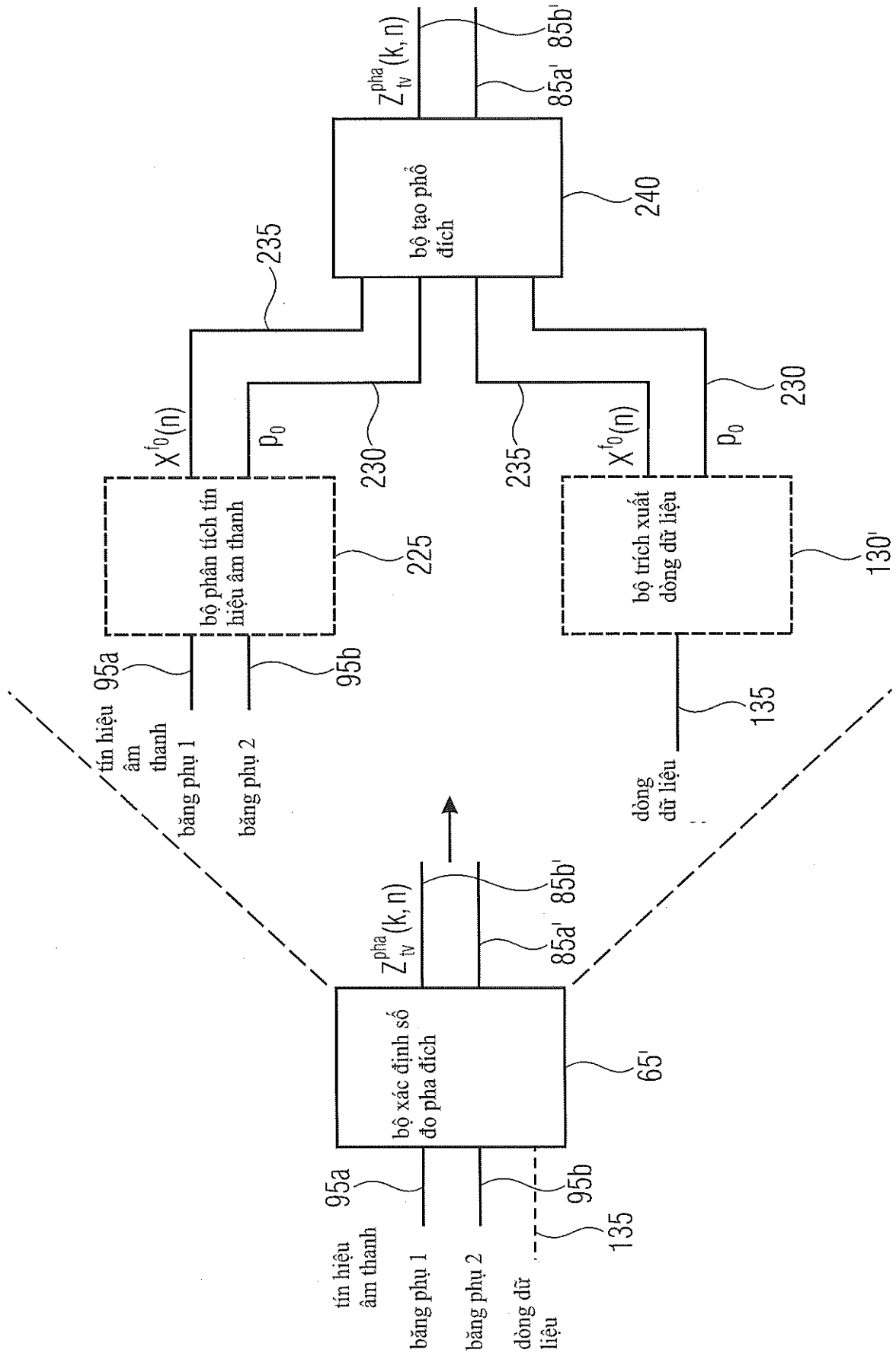


FIG 29

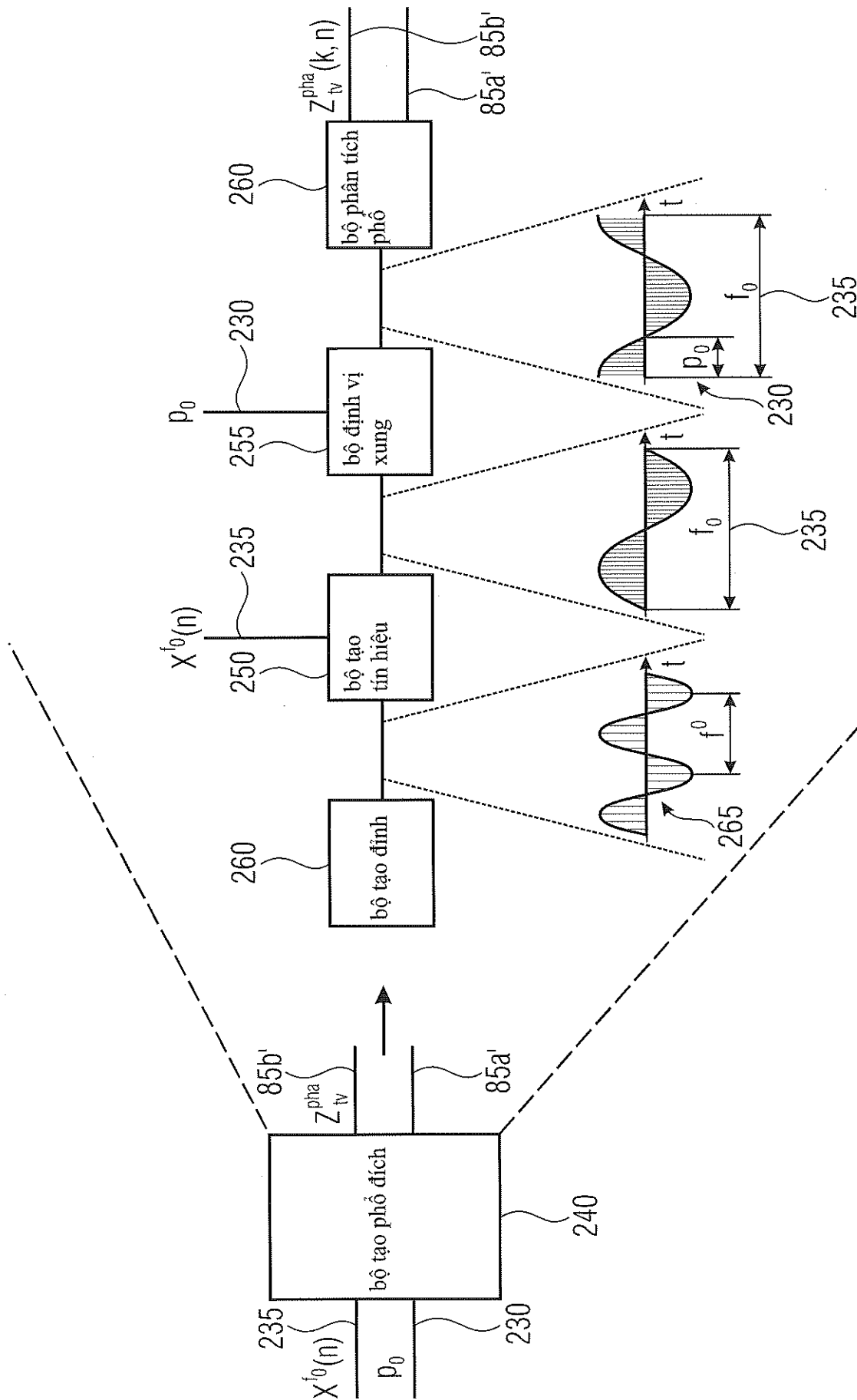


FIG 30

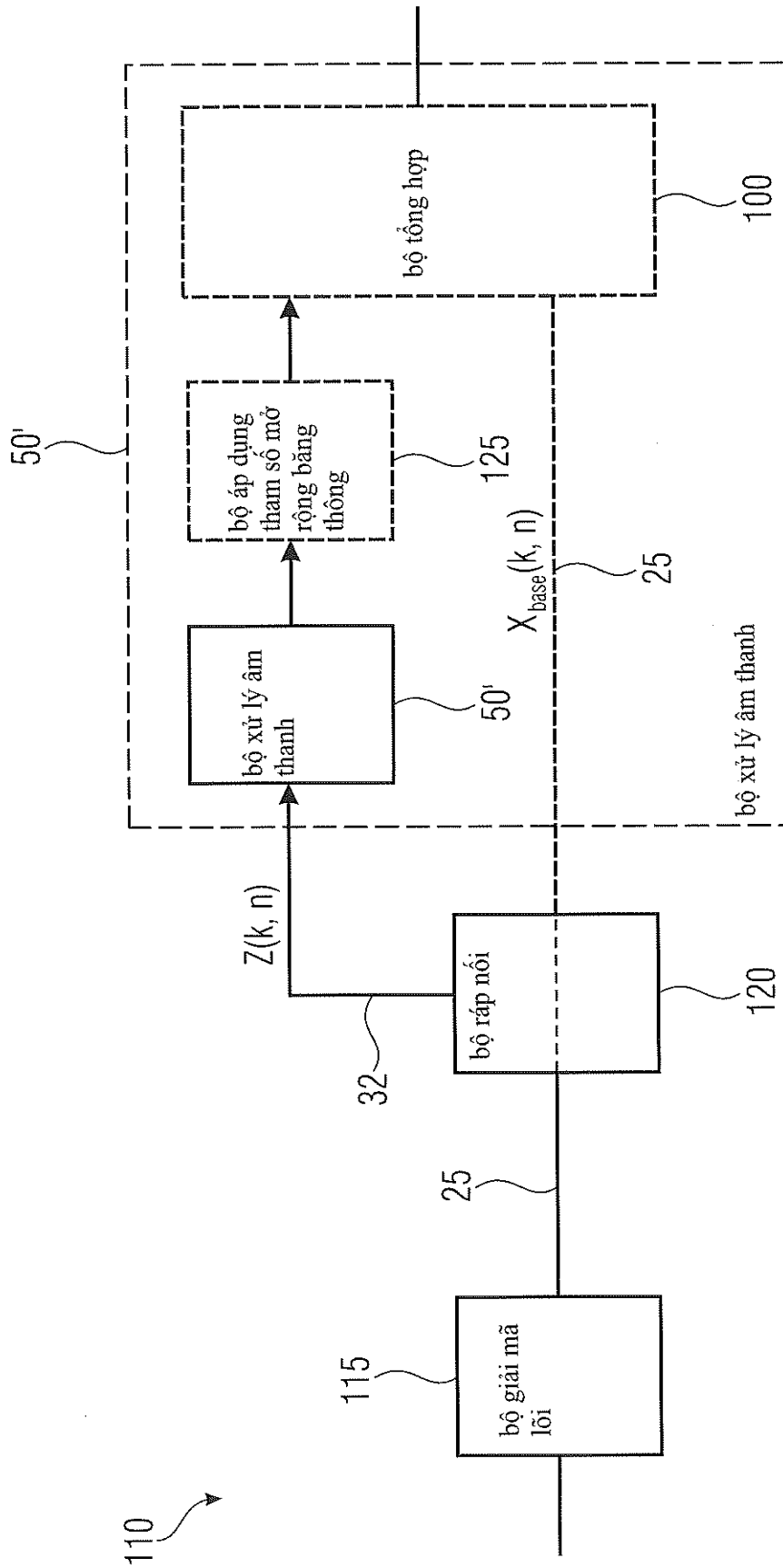


FIG 31

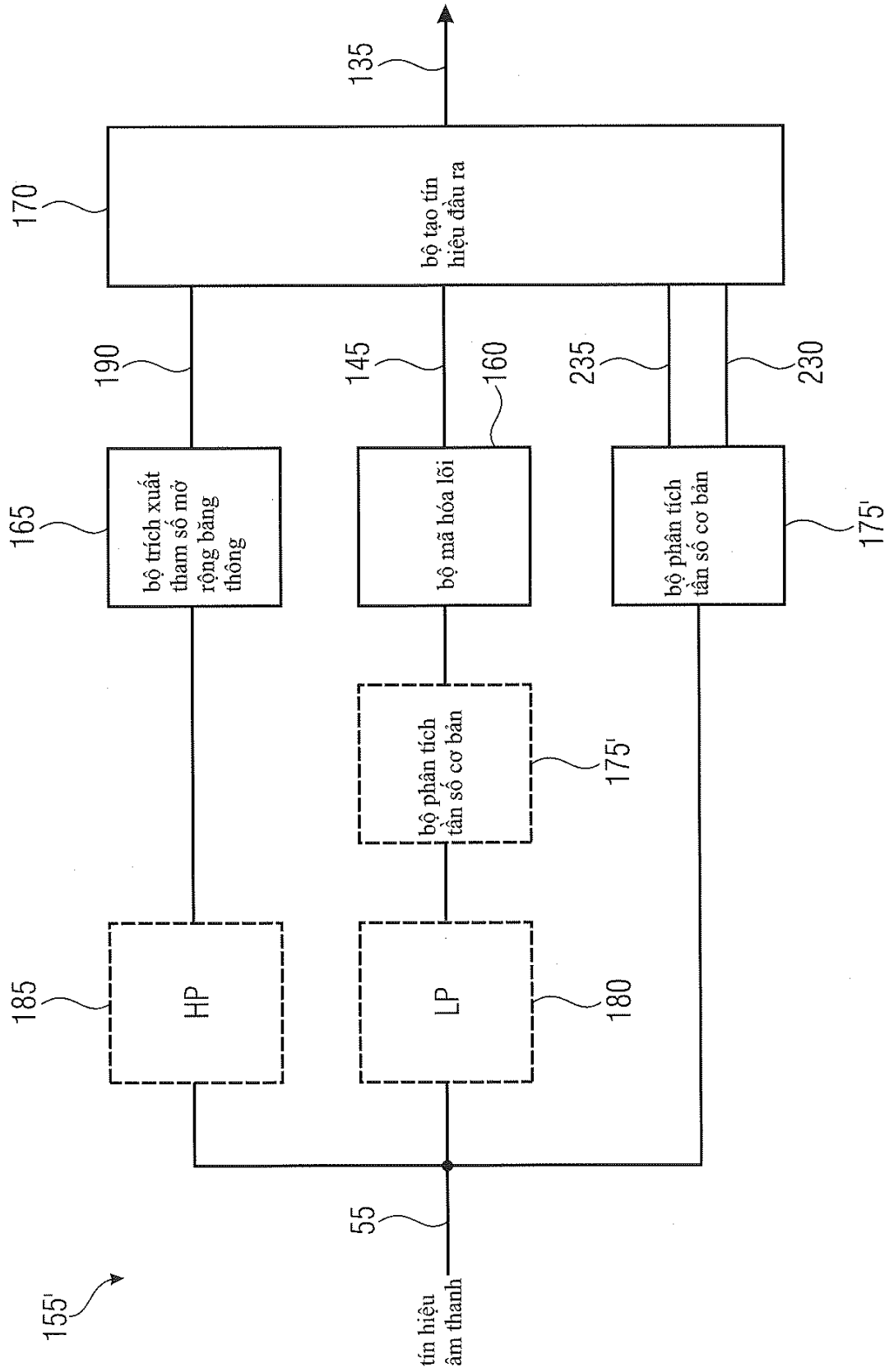


FIG 32

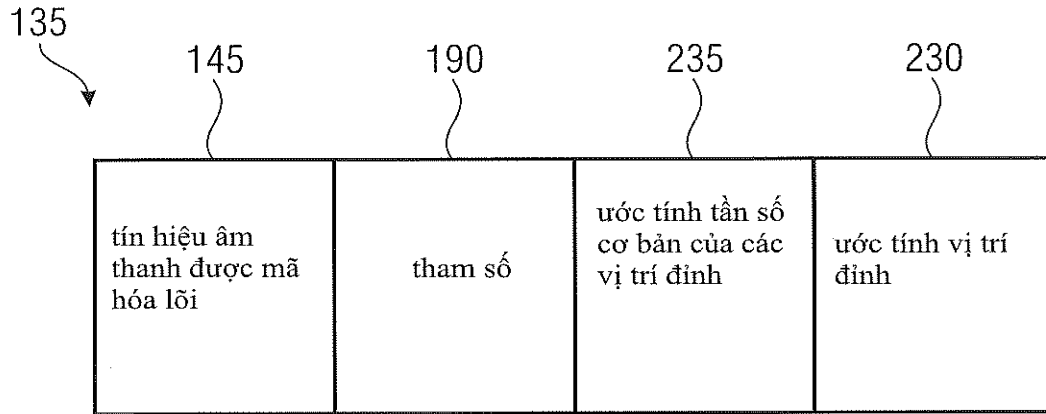


FIG 33

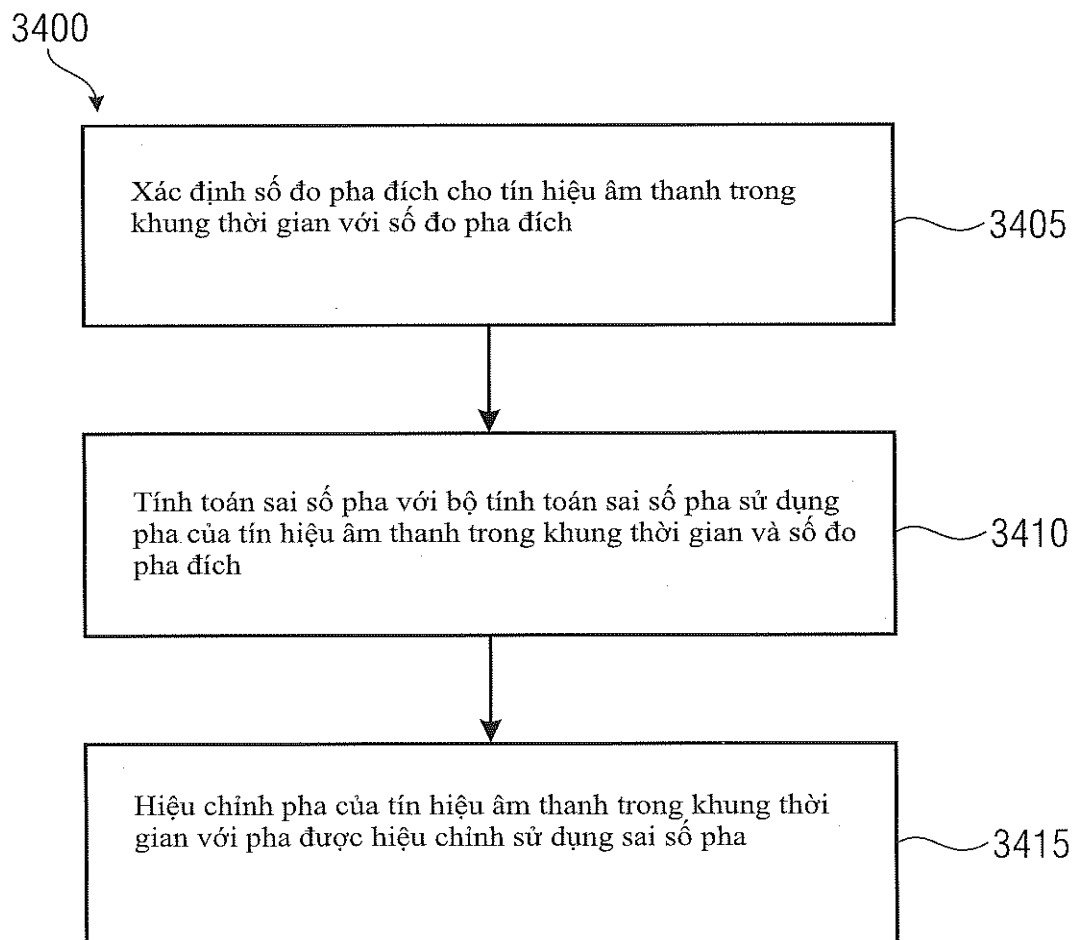


FIG 34

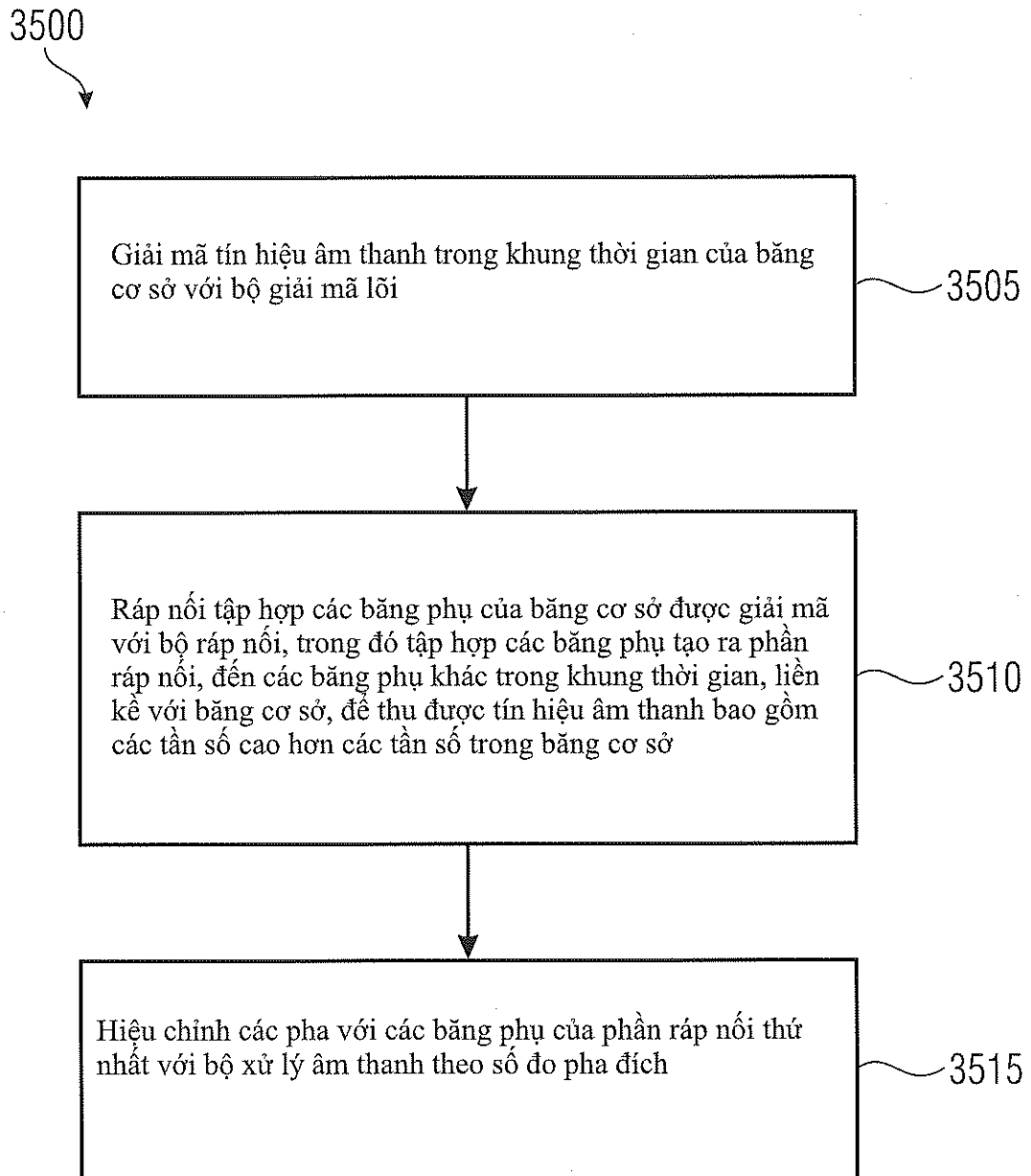


FIG 35

3600

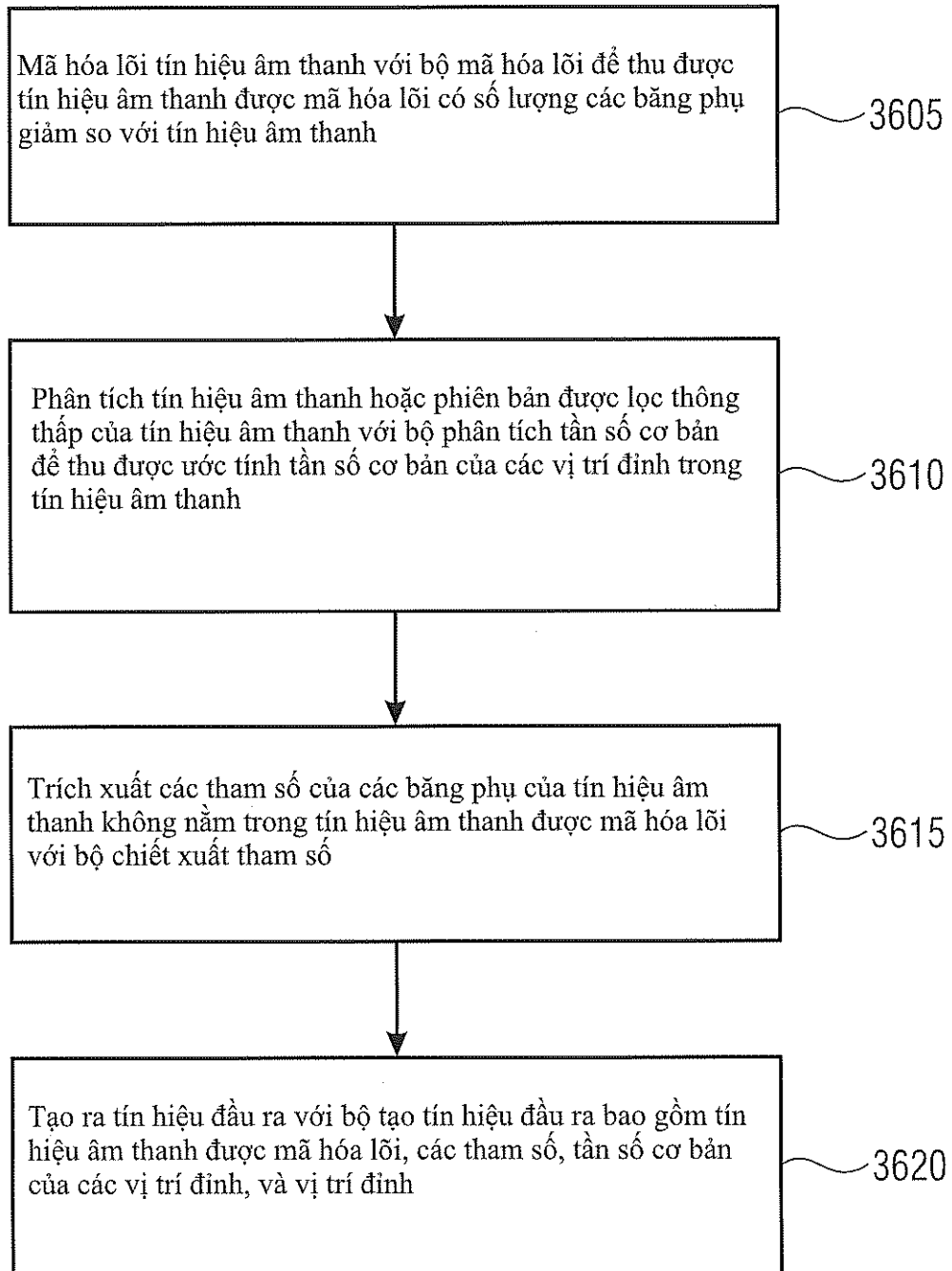


FIG 36

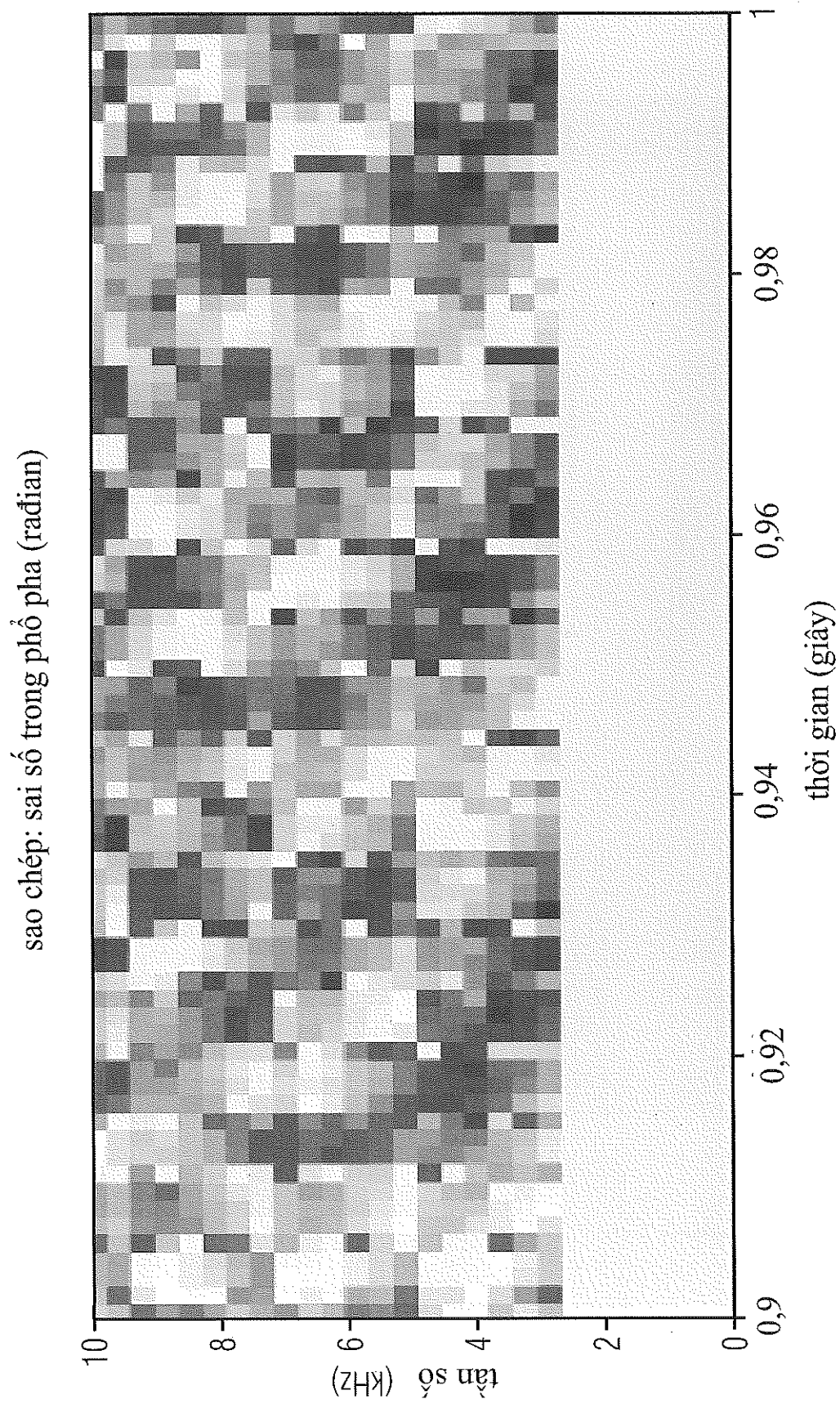


FIG 37

được hiệu chỉnh: sai số trong phổ pha (radian)

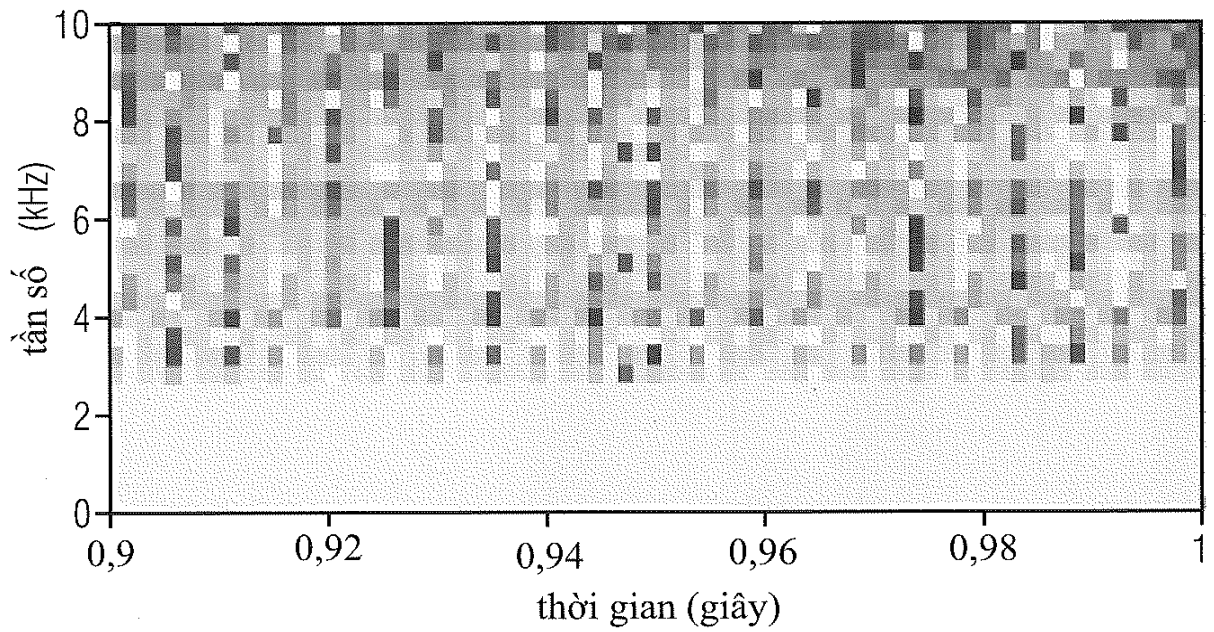


FIG 38A

được hiệu chỉnh: đạo hàm pha theo tần số (radian)

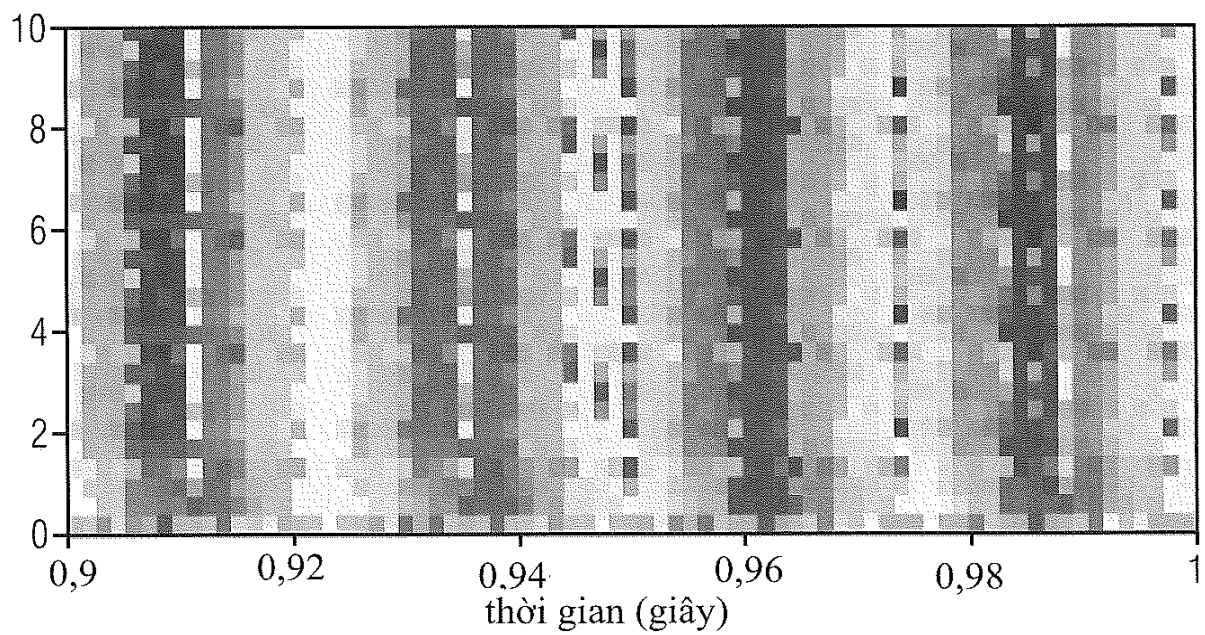


FIG 38B

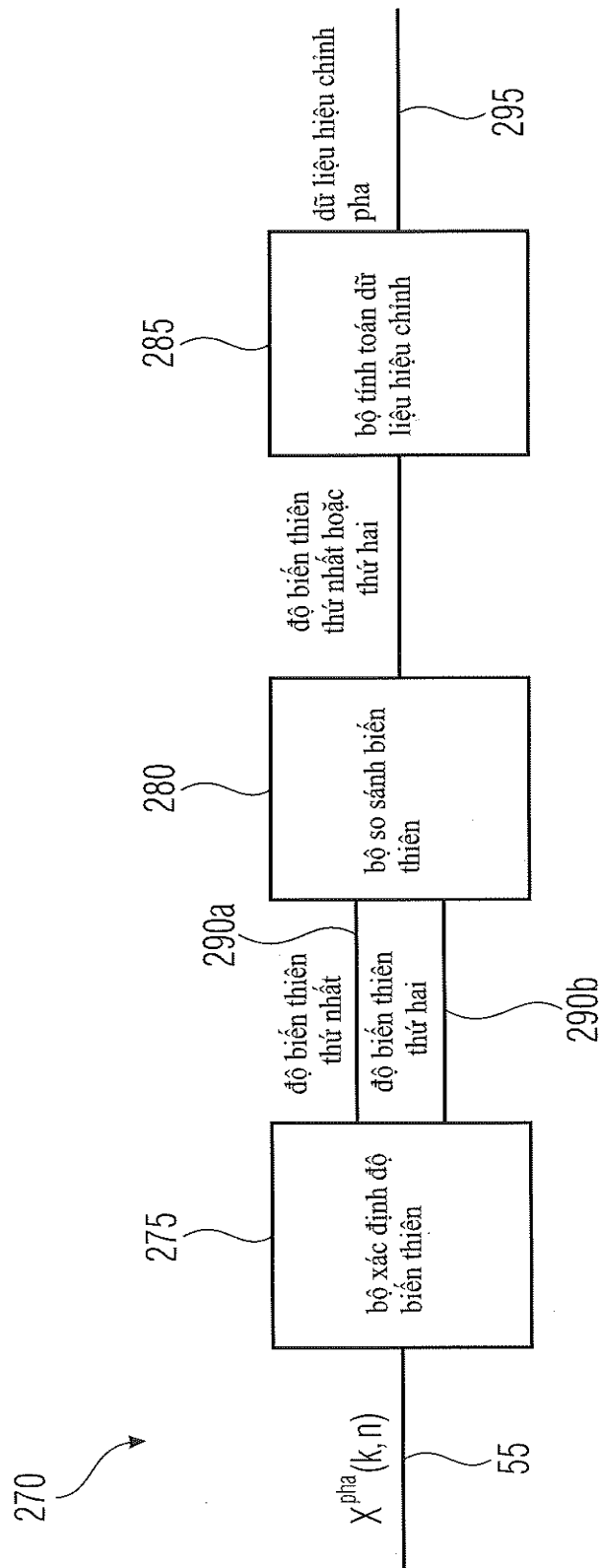


FIG 39

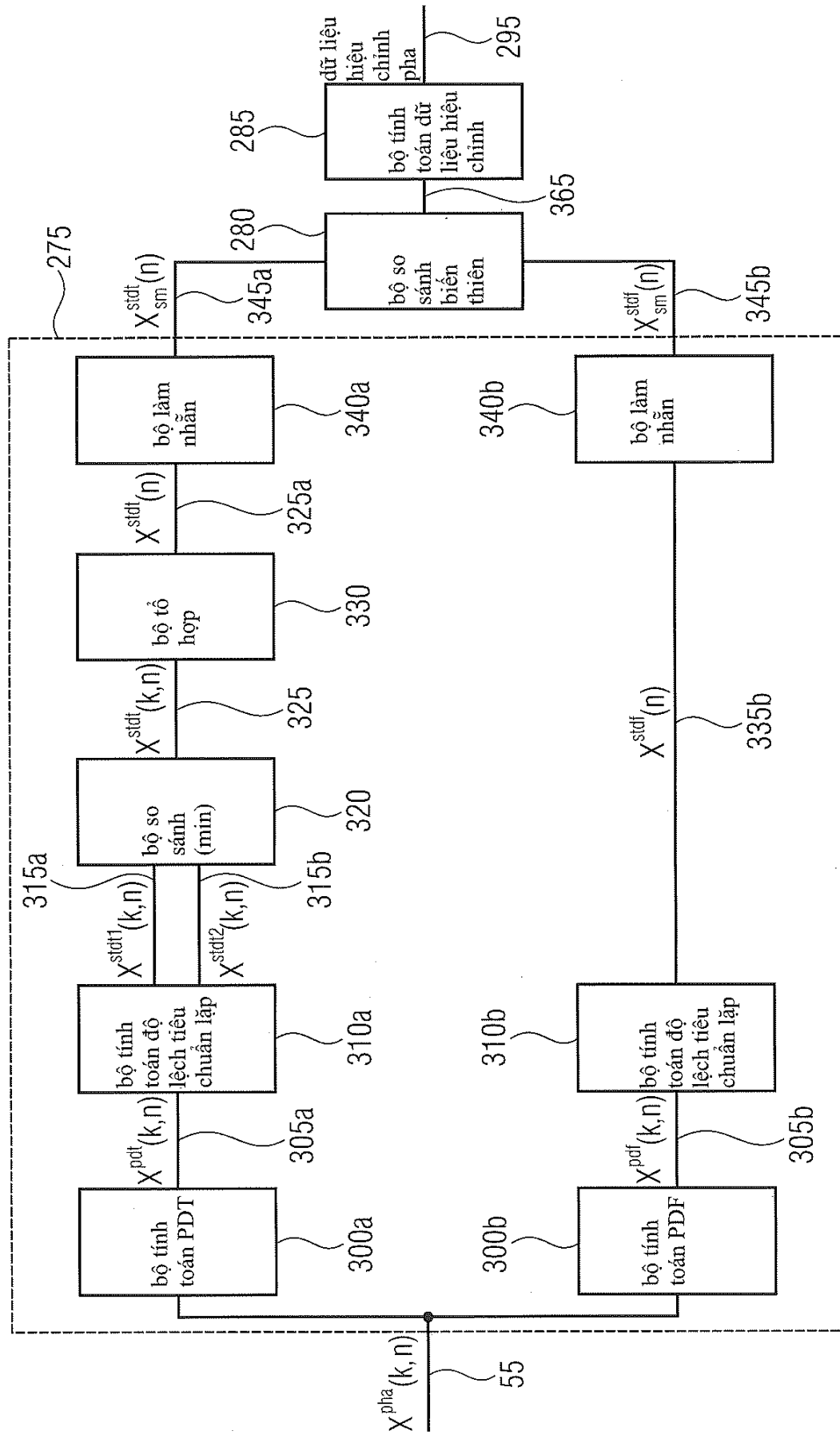


FIG 40

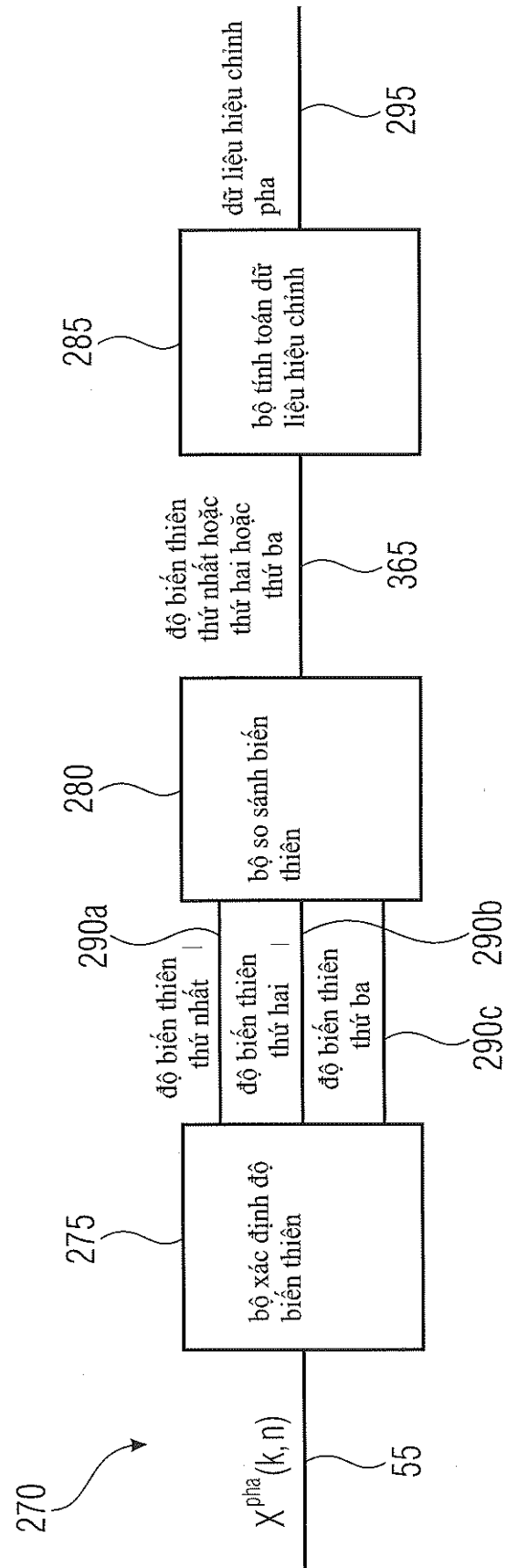


FIG 41

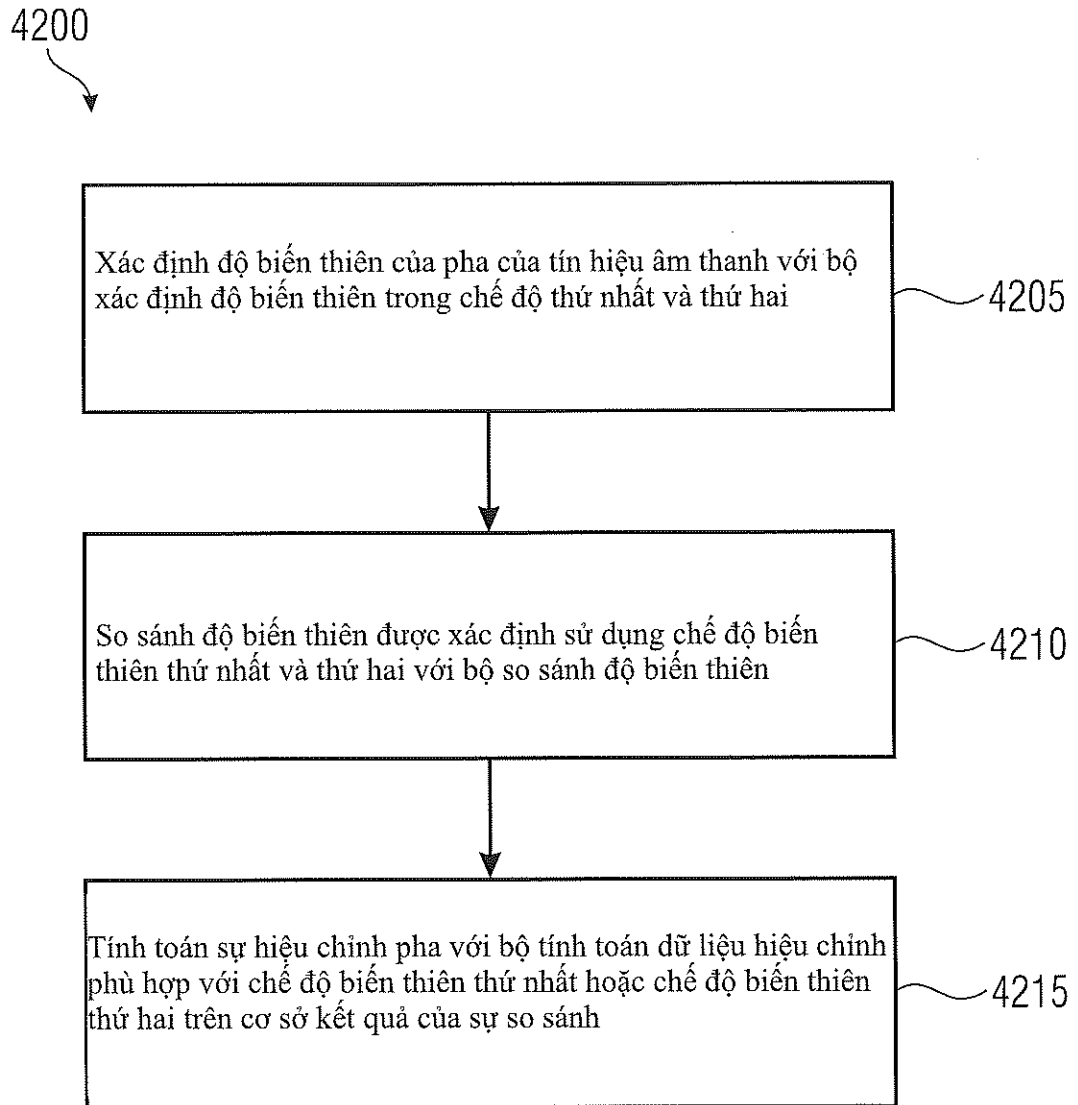


FIG 42

ban đầu: độ lệch chuẩn của PDT (radian)

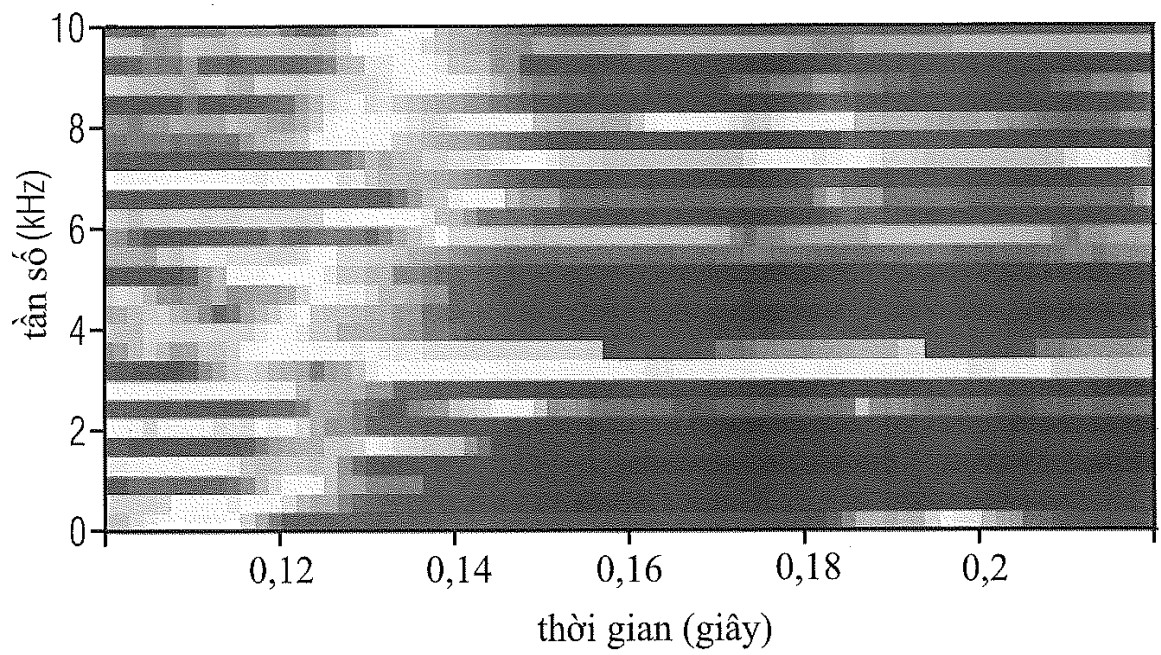


FIG 43A

ban đầu: độ lệch chuẩn của PDF (radian)

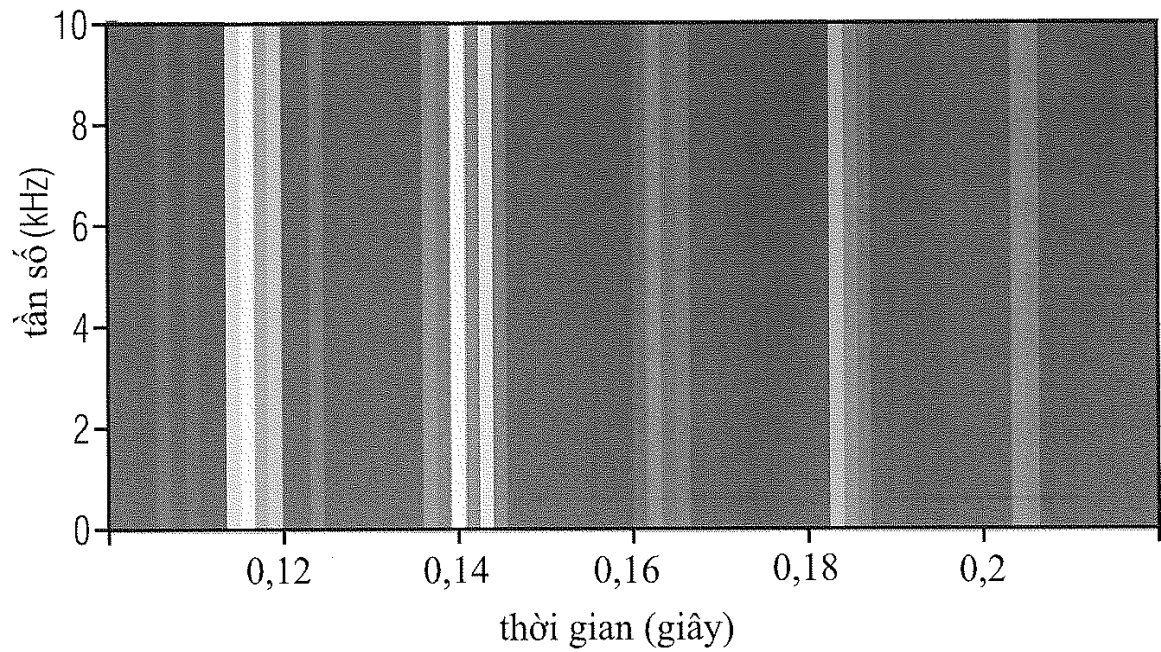


FIG 43B

ban đầu: độ lệch chuẩn của PDT (radian)

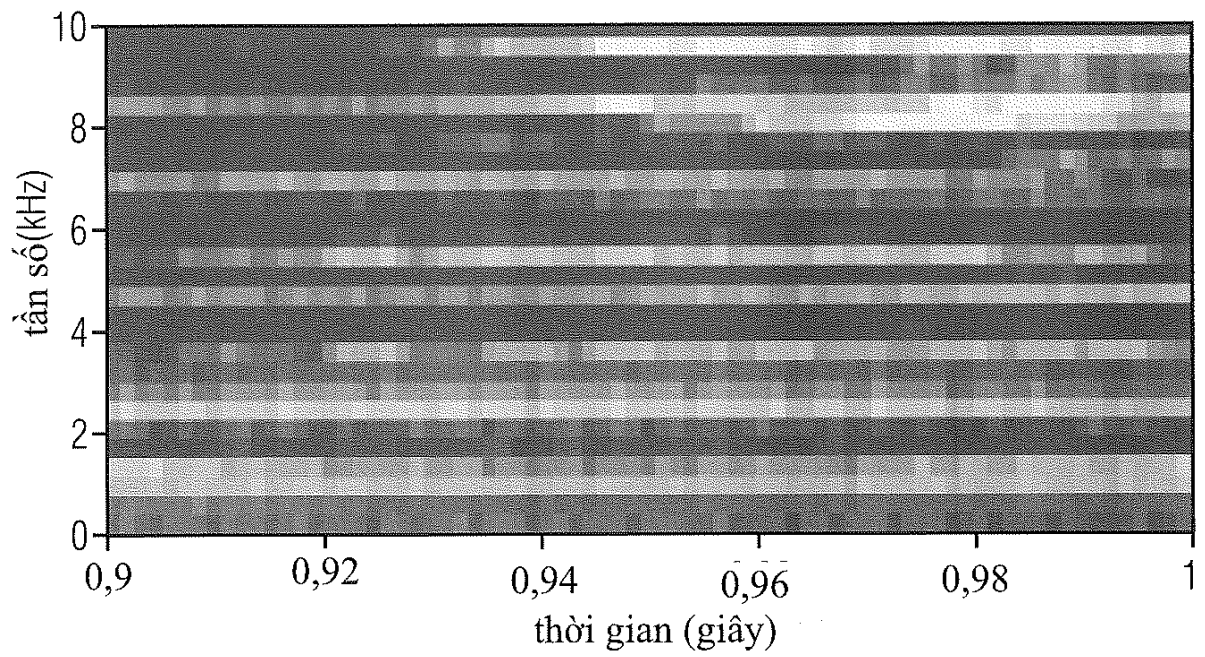


FIG 43C

ban đầu: độ lệch chuẩn của PDF (radian)

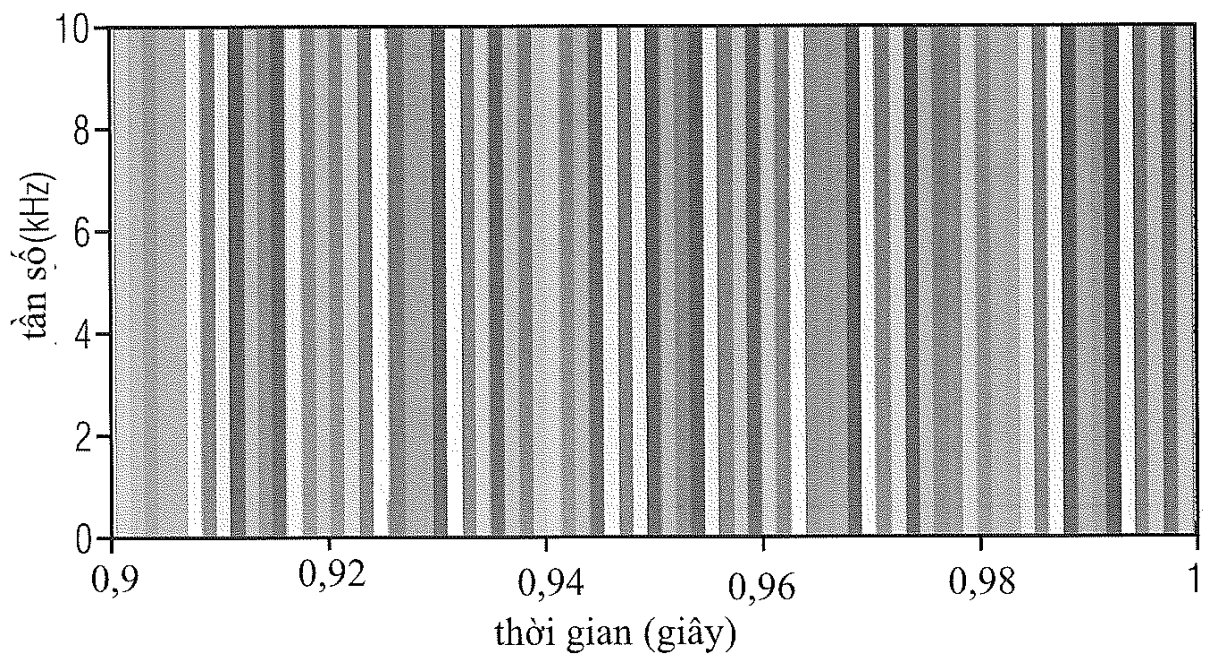


FIG 43D

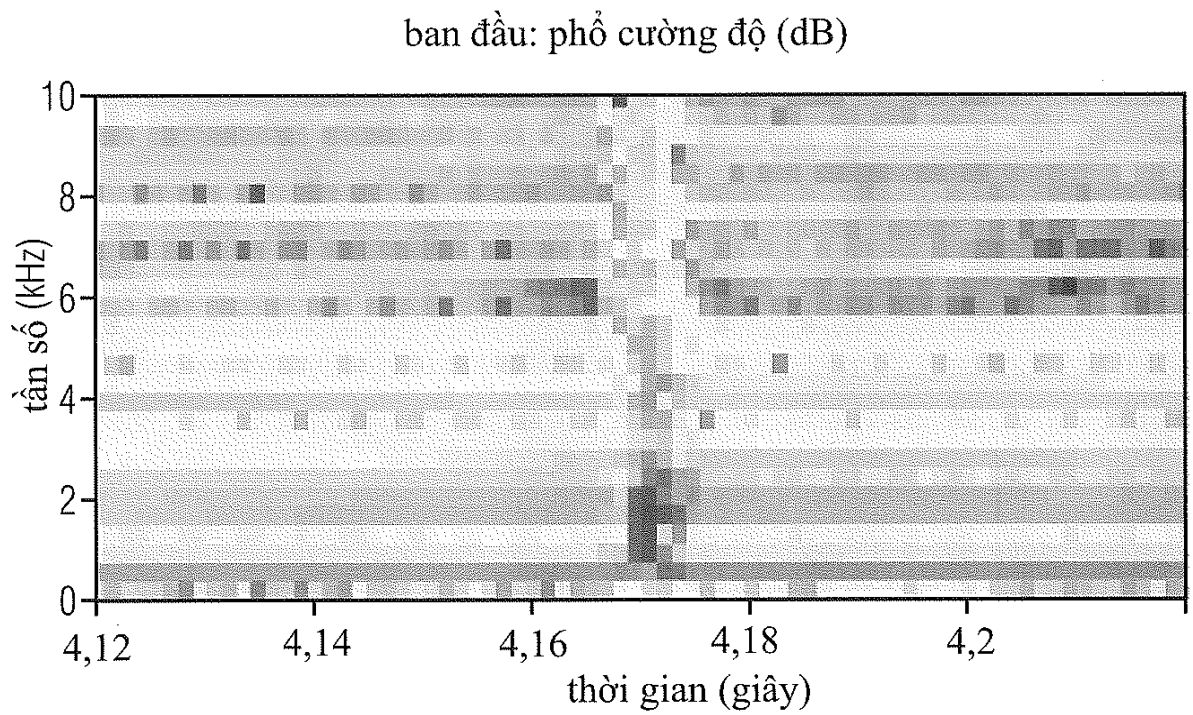


FIG 44A

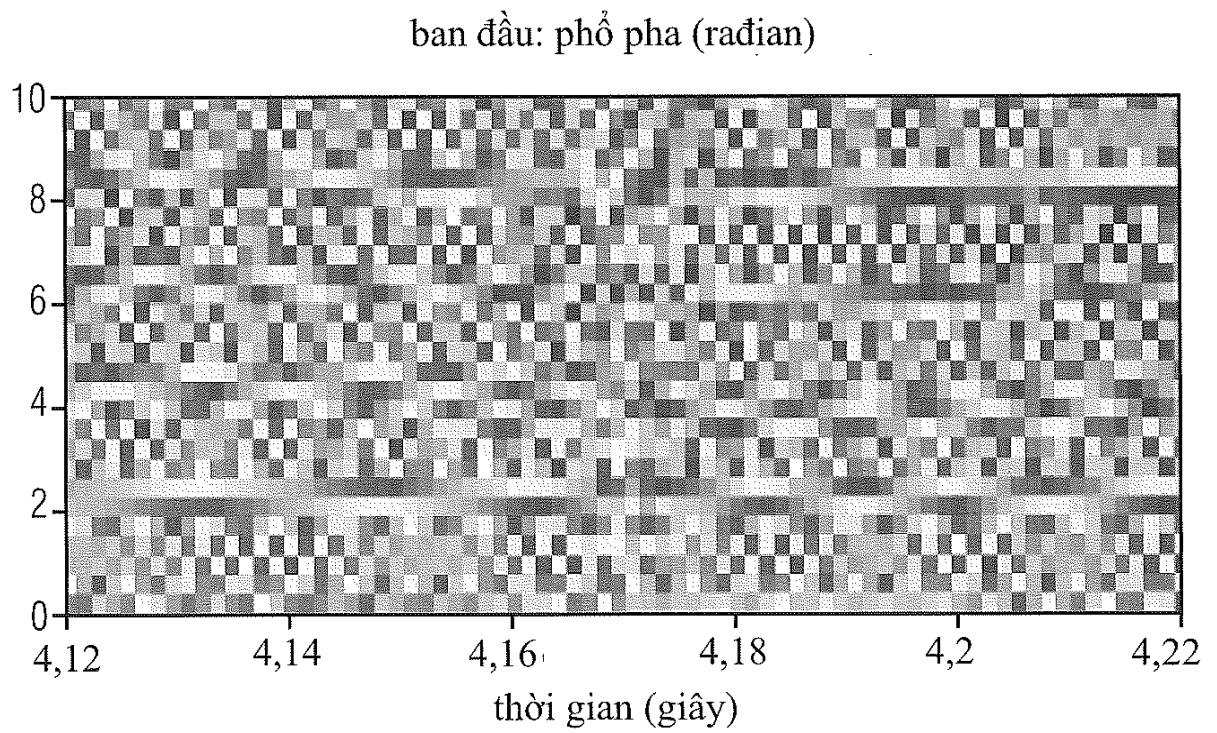


FIG 44B

ban đầu: đạo hàm pha theo thời gian (radian)

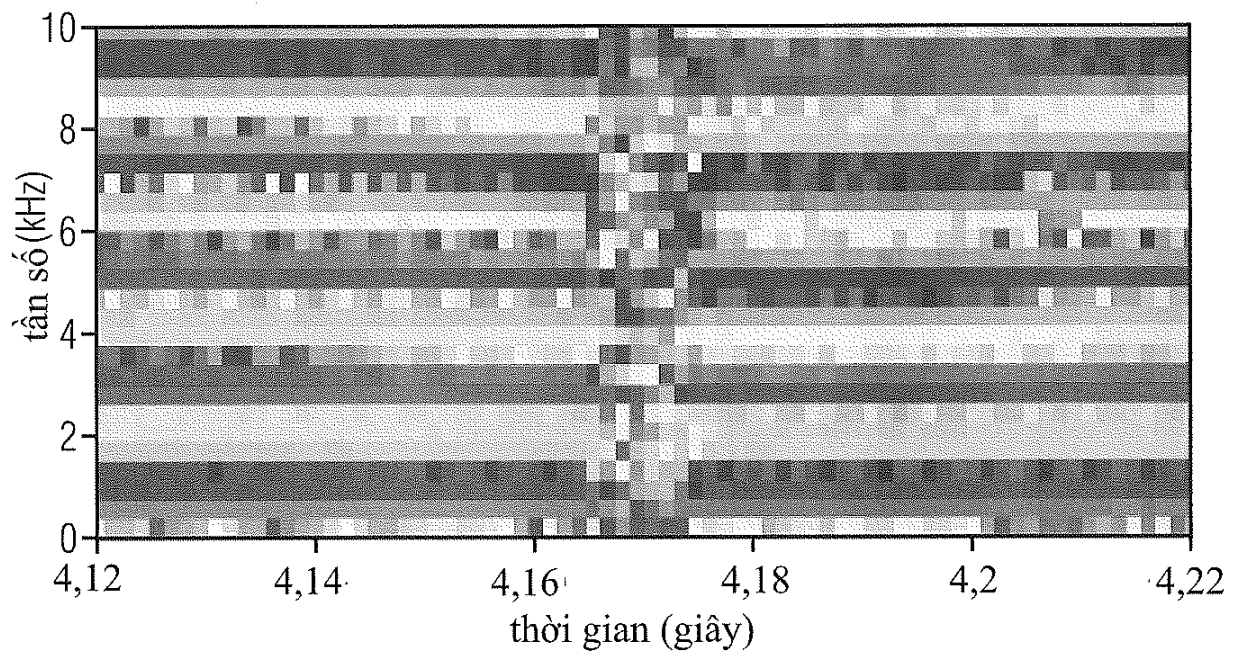


FIG 45A

ban đầu: đạo hàm pha theo tần số (radian)

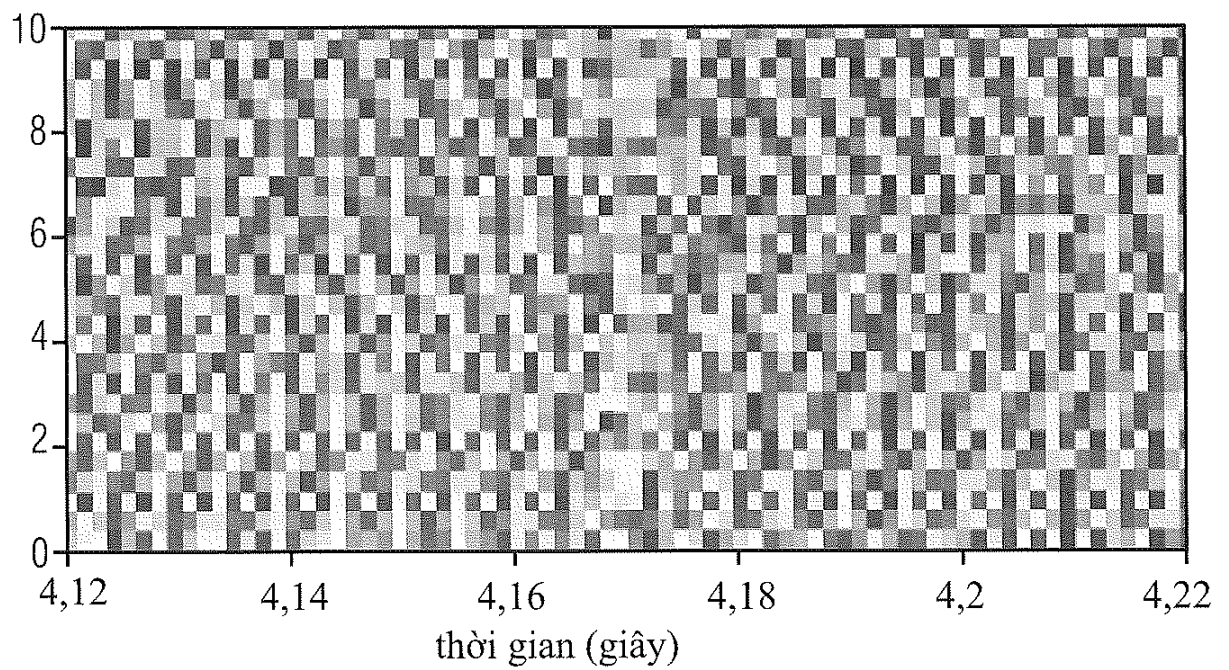


FIG 45B

được hiệu chỉnh: đạo hàm pha theo thời gian (radian)

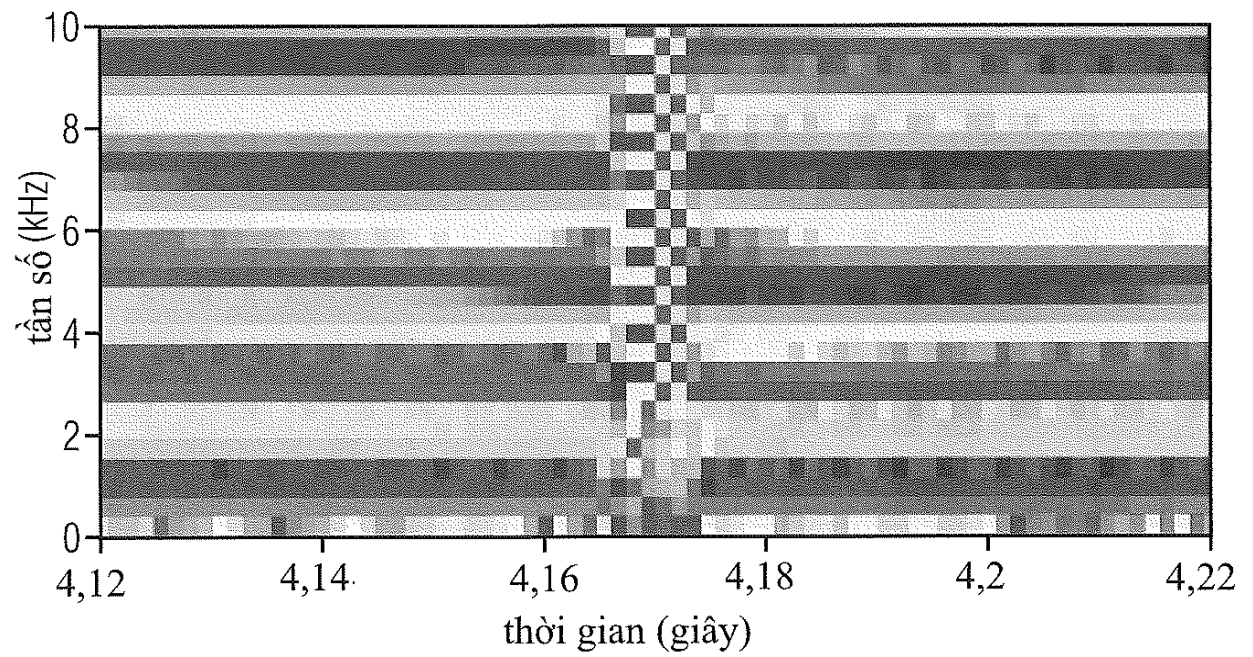


FIG 46A

được hiệu chỉnh: đạo hàm pha theo tần số (radian)

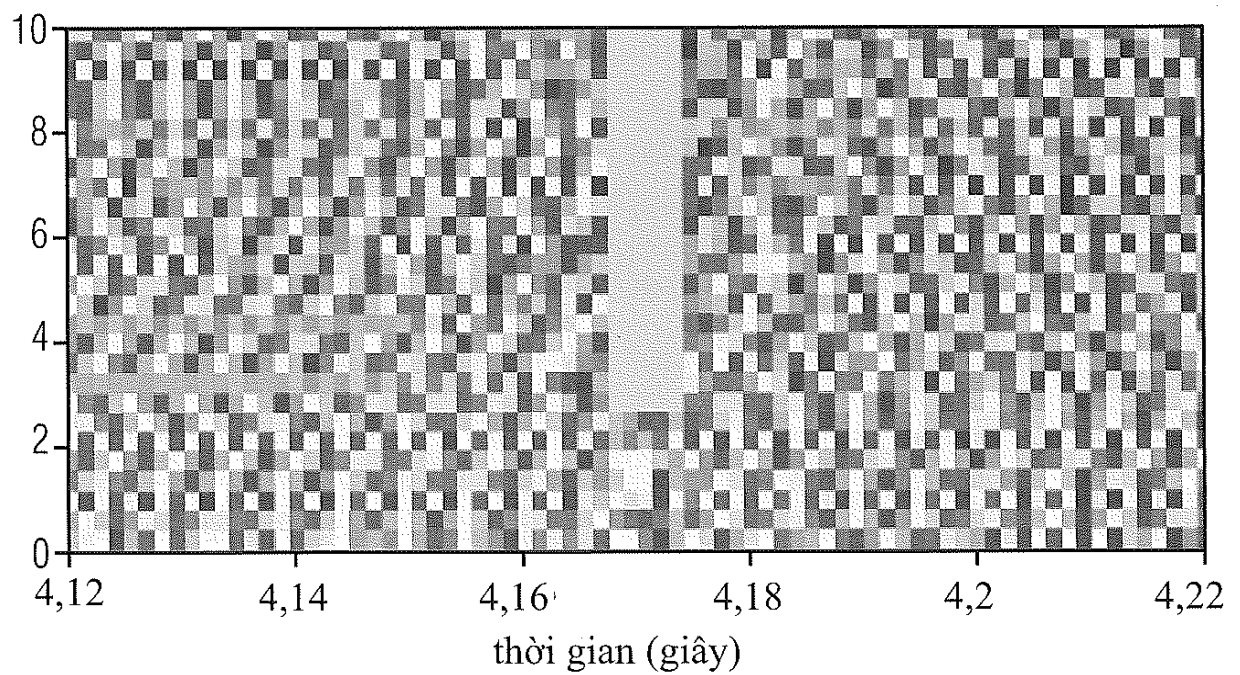


FIG 46B

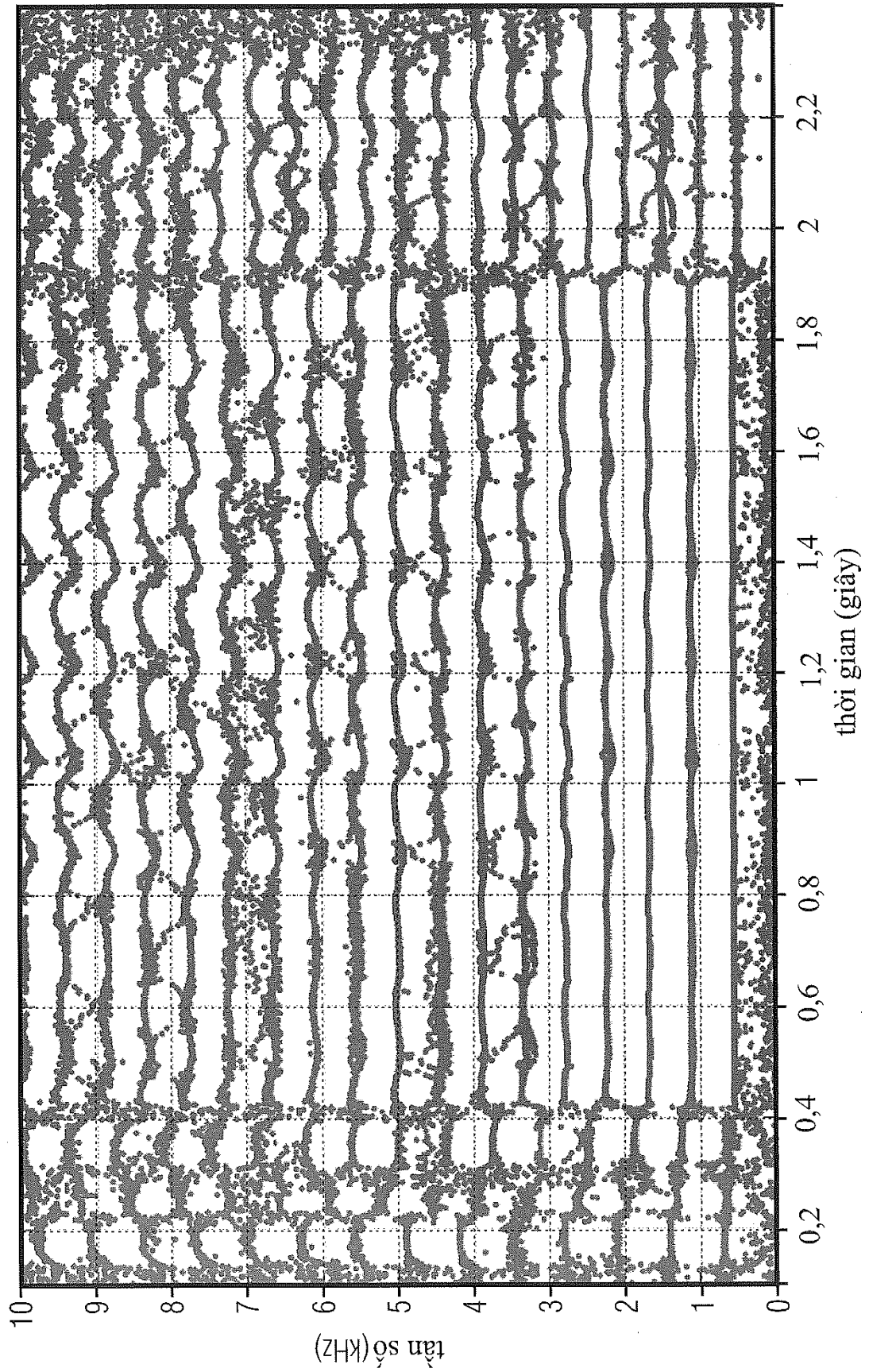


FIG 47

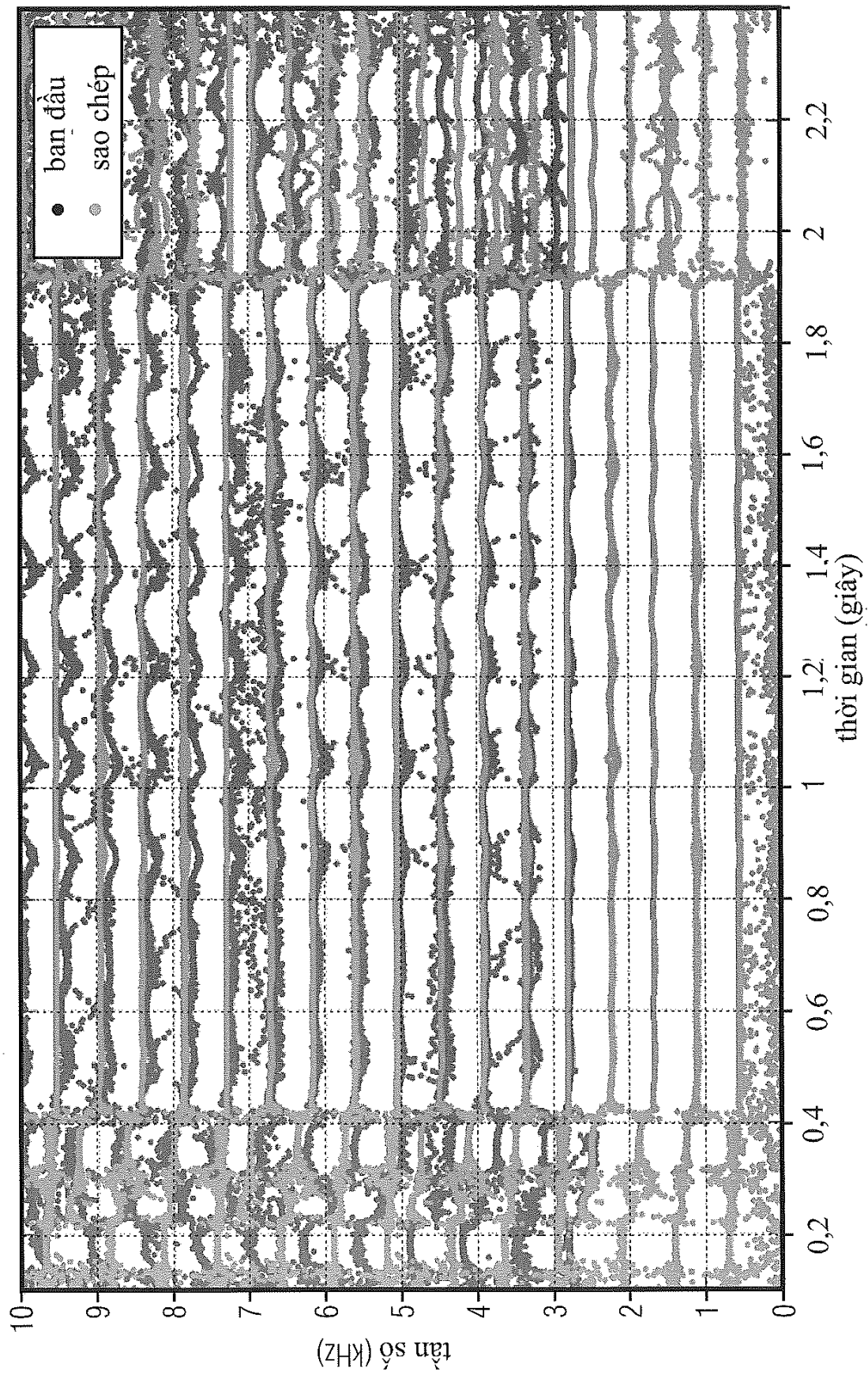


FIG 48A

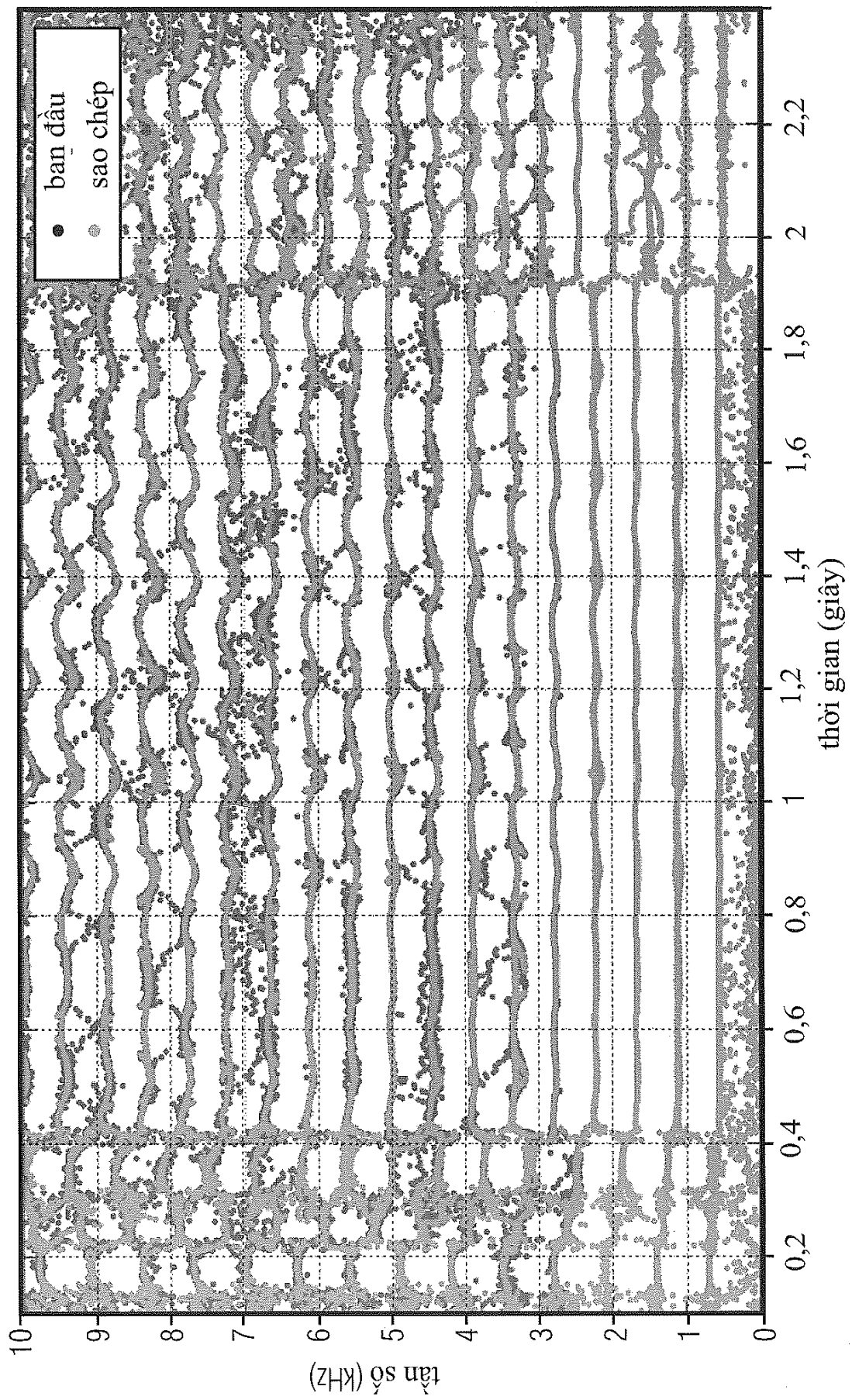


FIG 48B

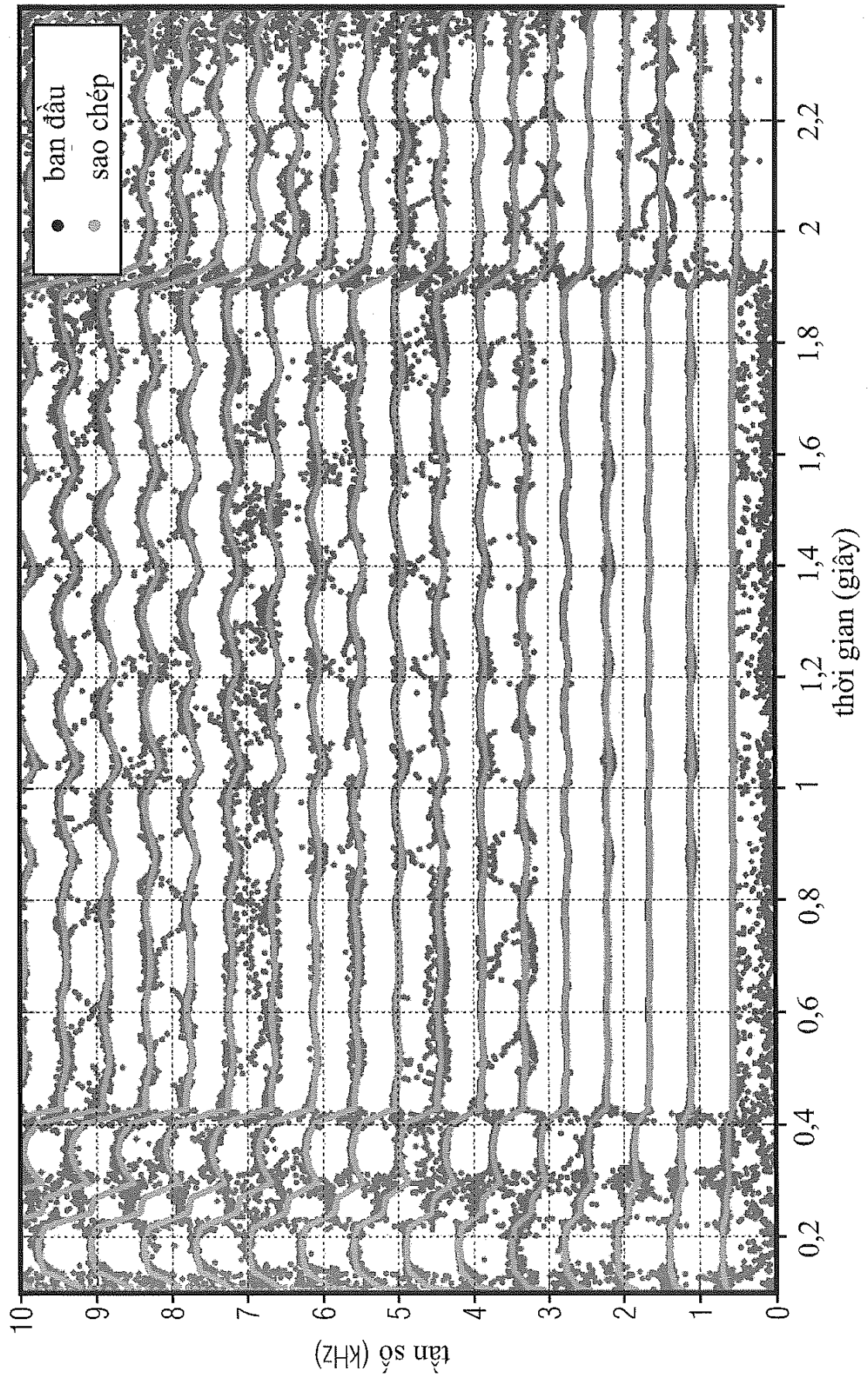


FIG 49

được hiệu chỉnh: sai số trong PDT (radian)

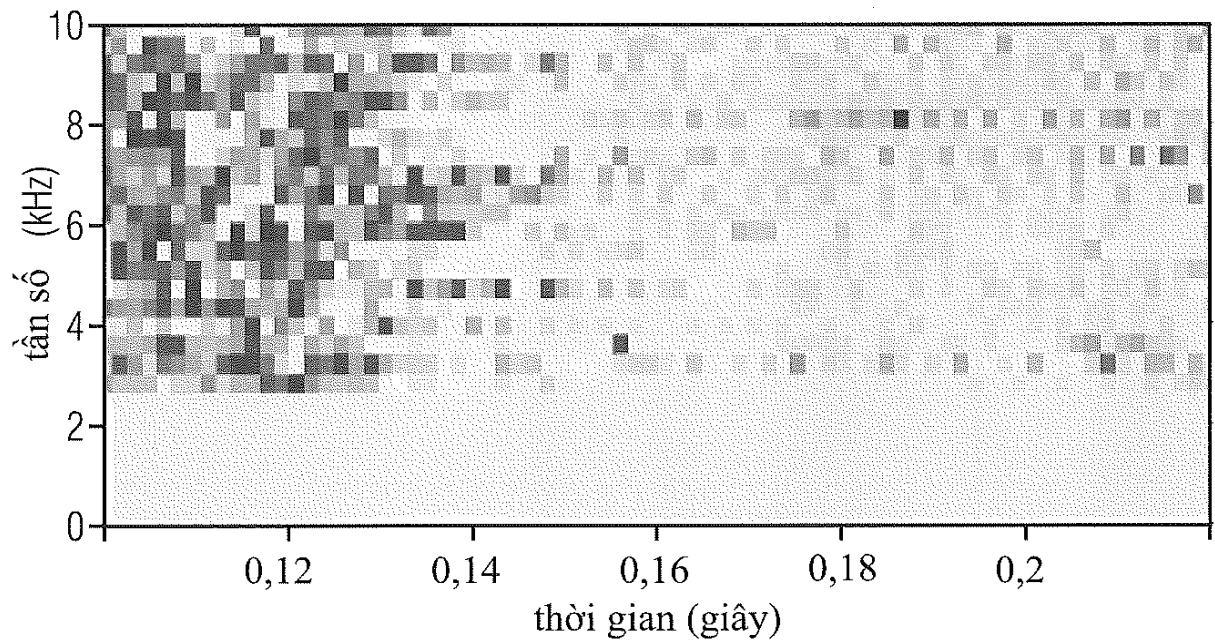


FIG 50A

được hiệu chỉnh: đạo hàm pha theo thời gian (radian)

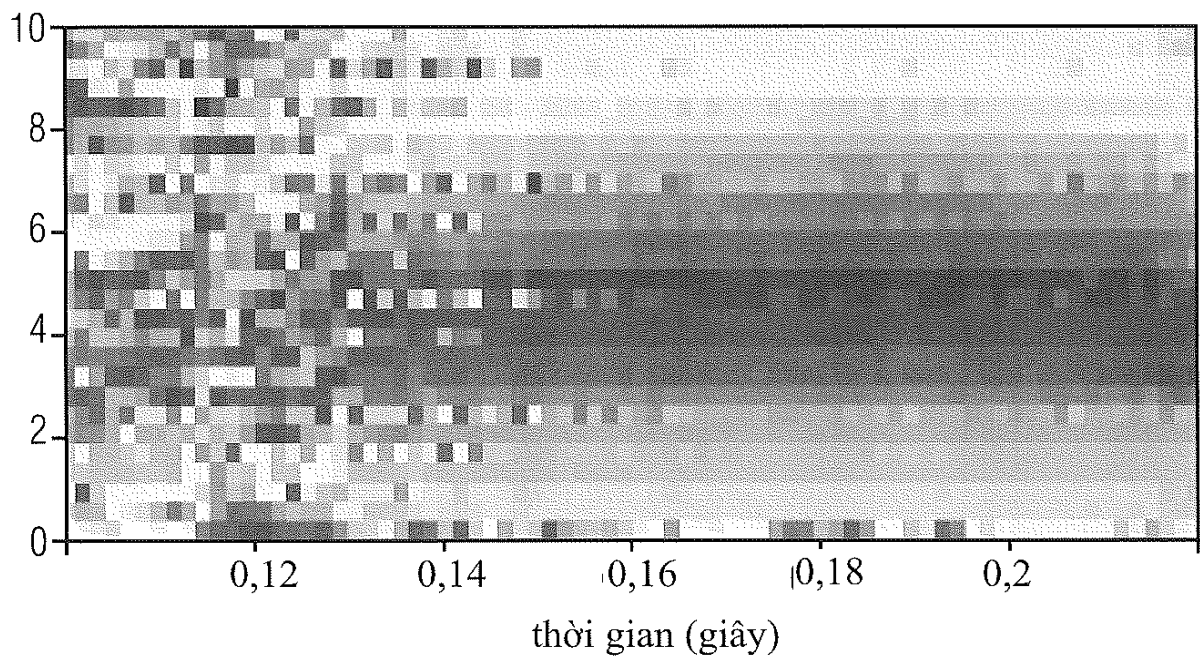


FIG 50B

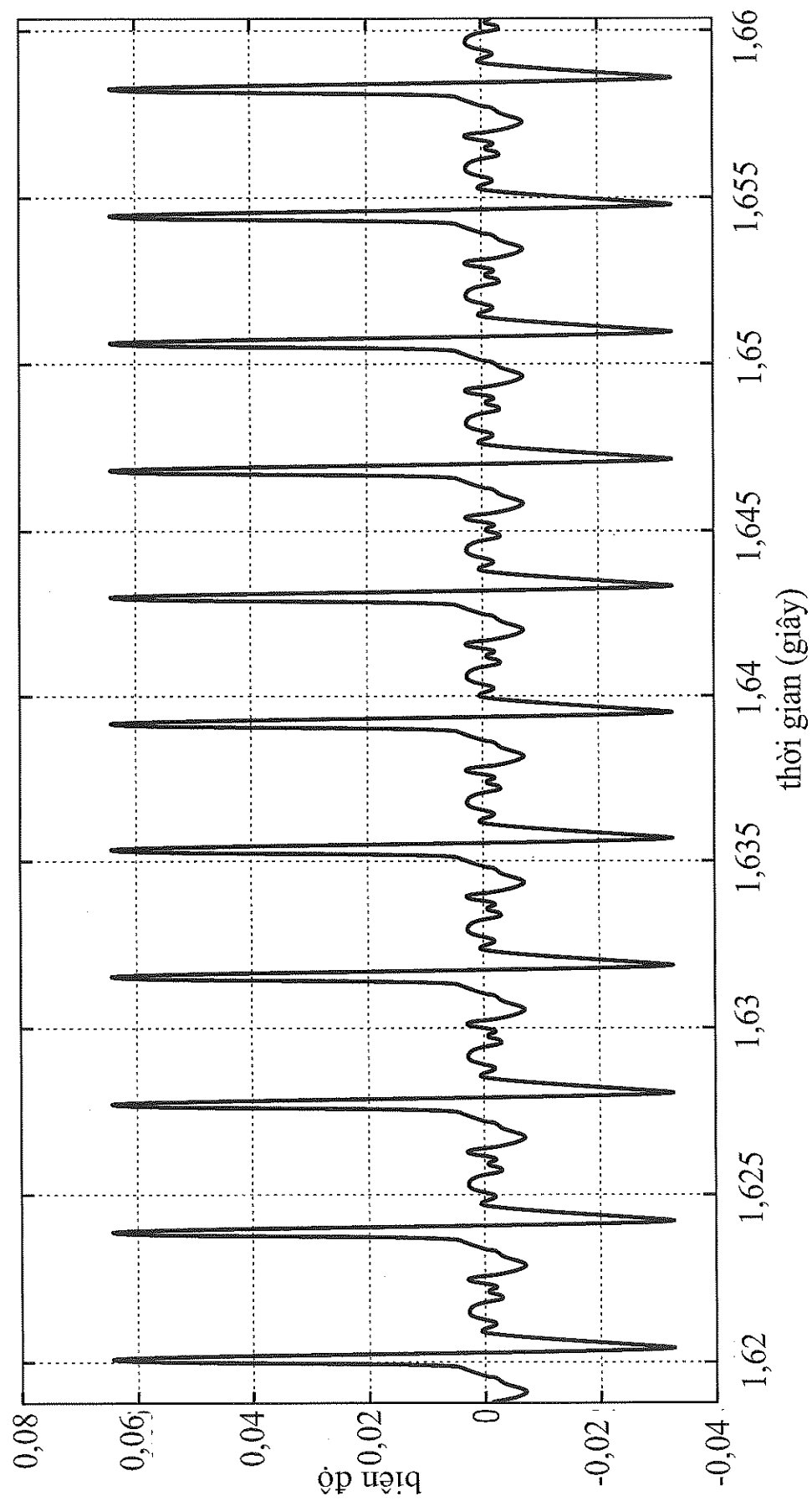


FIG 51A

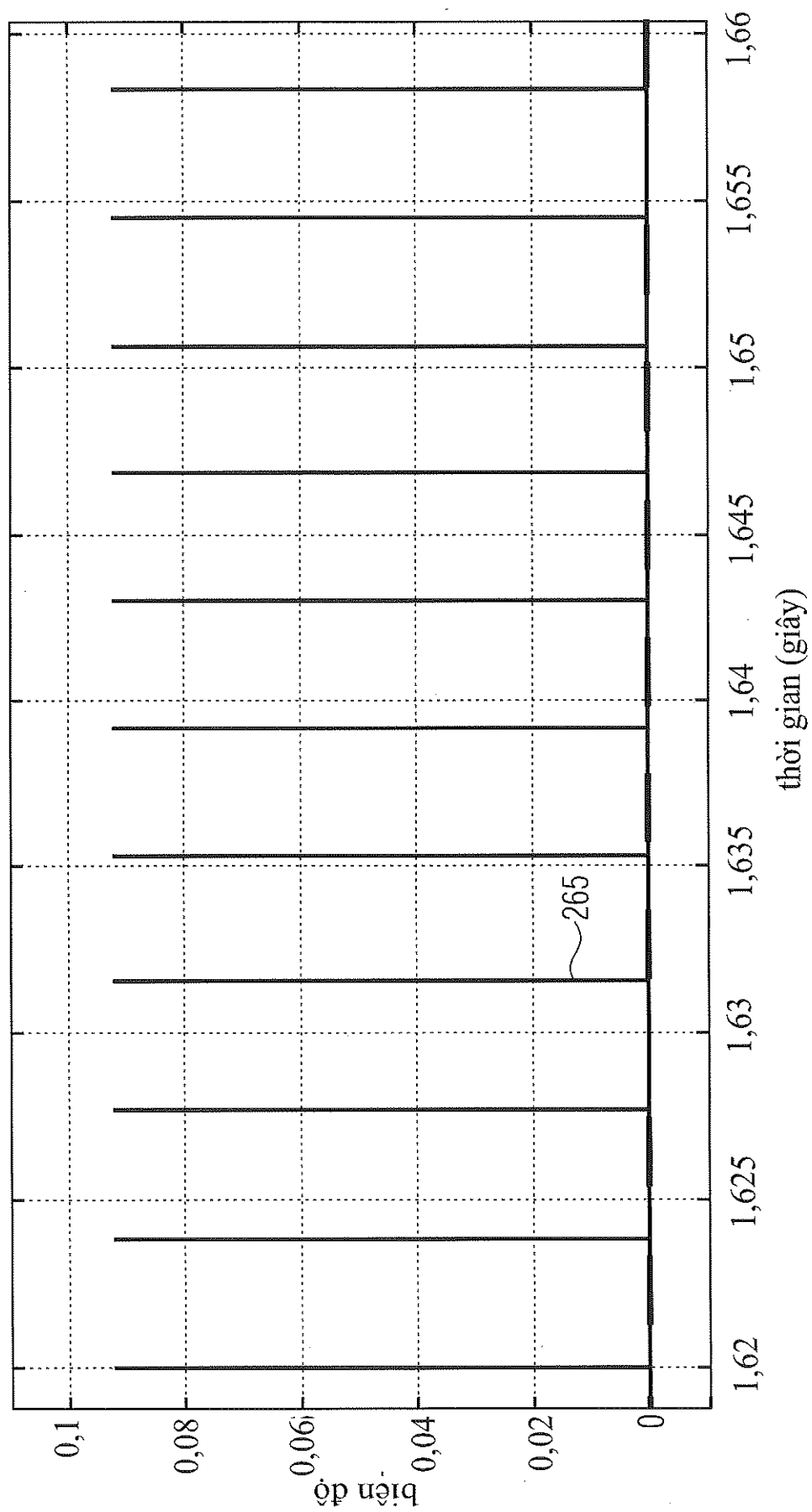


FIG 51B

được hiệu chỉnh: sai số trong phổ pha (radian)

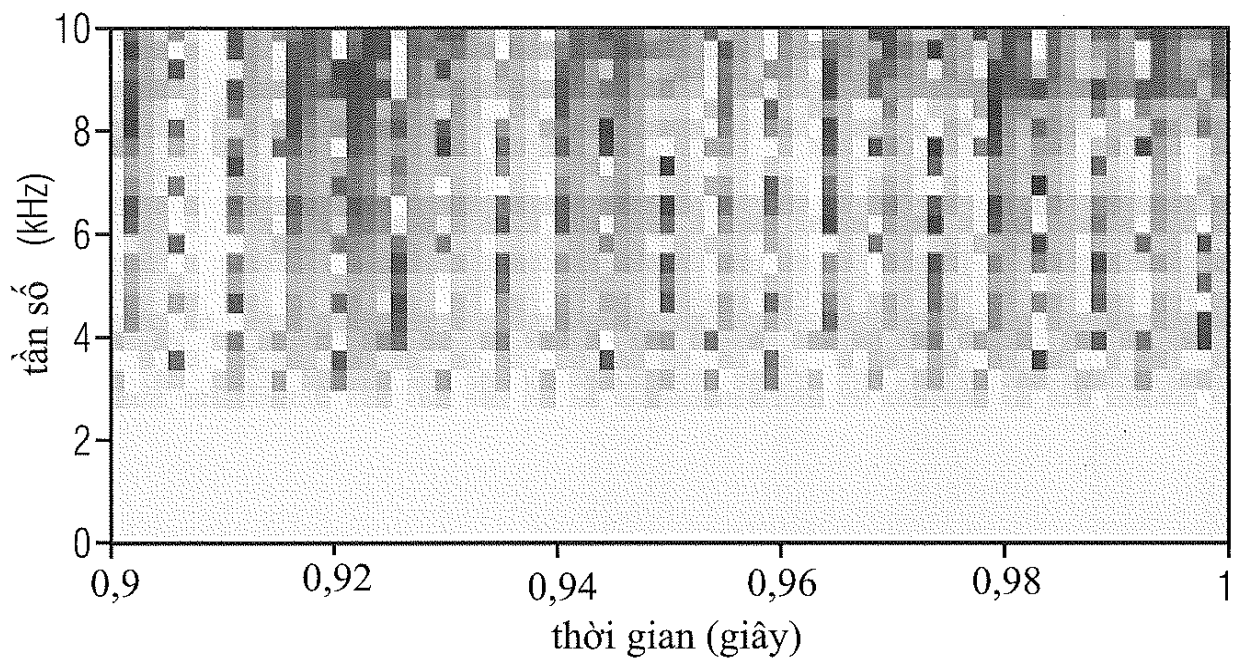


FIG 52A

được hiệu chỉnh: đạo hàm pha theo tần số (radian)

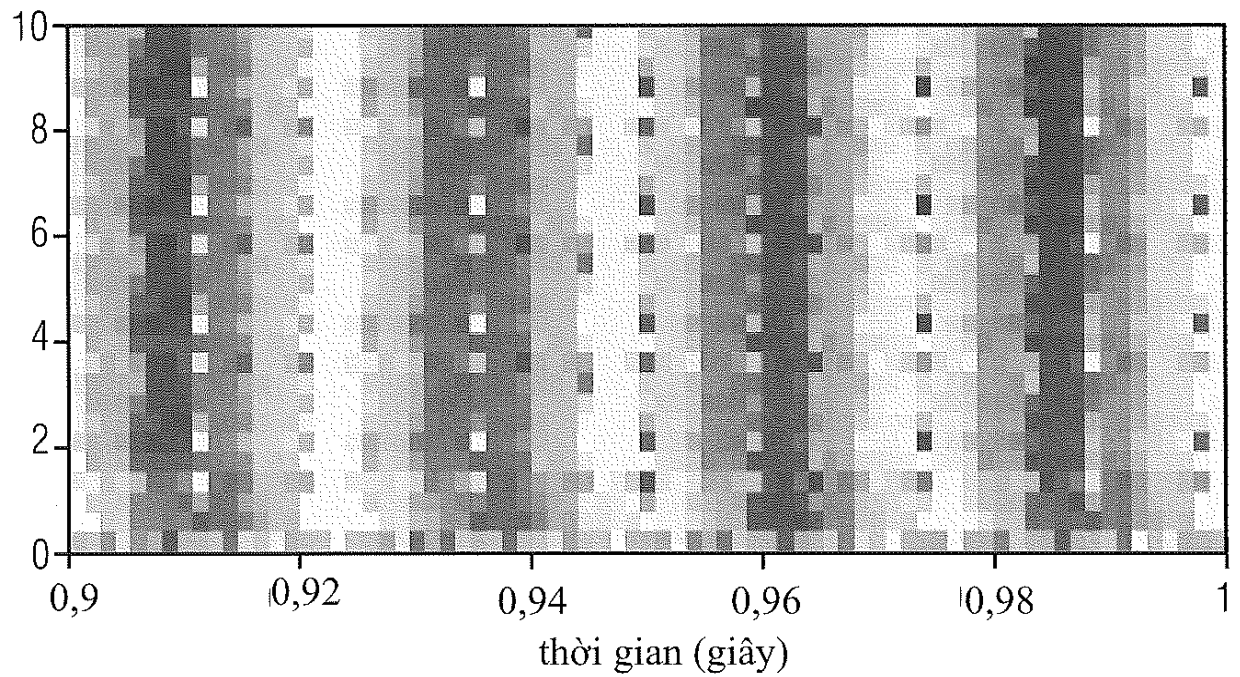


FIG 52B

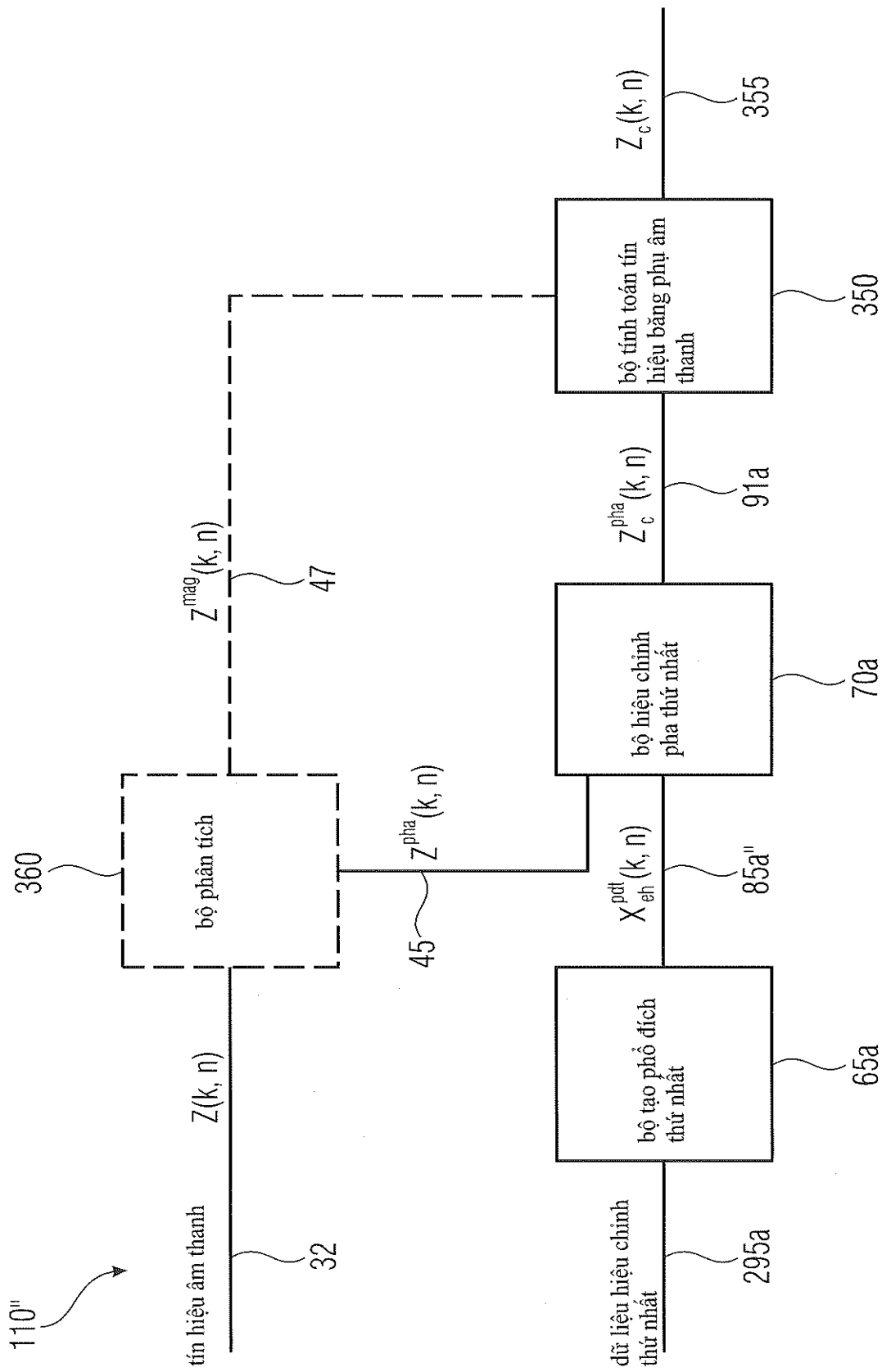


FIG 53

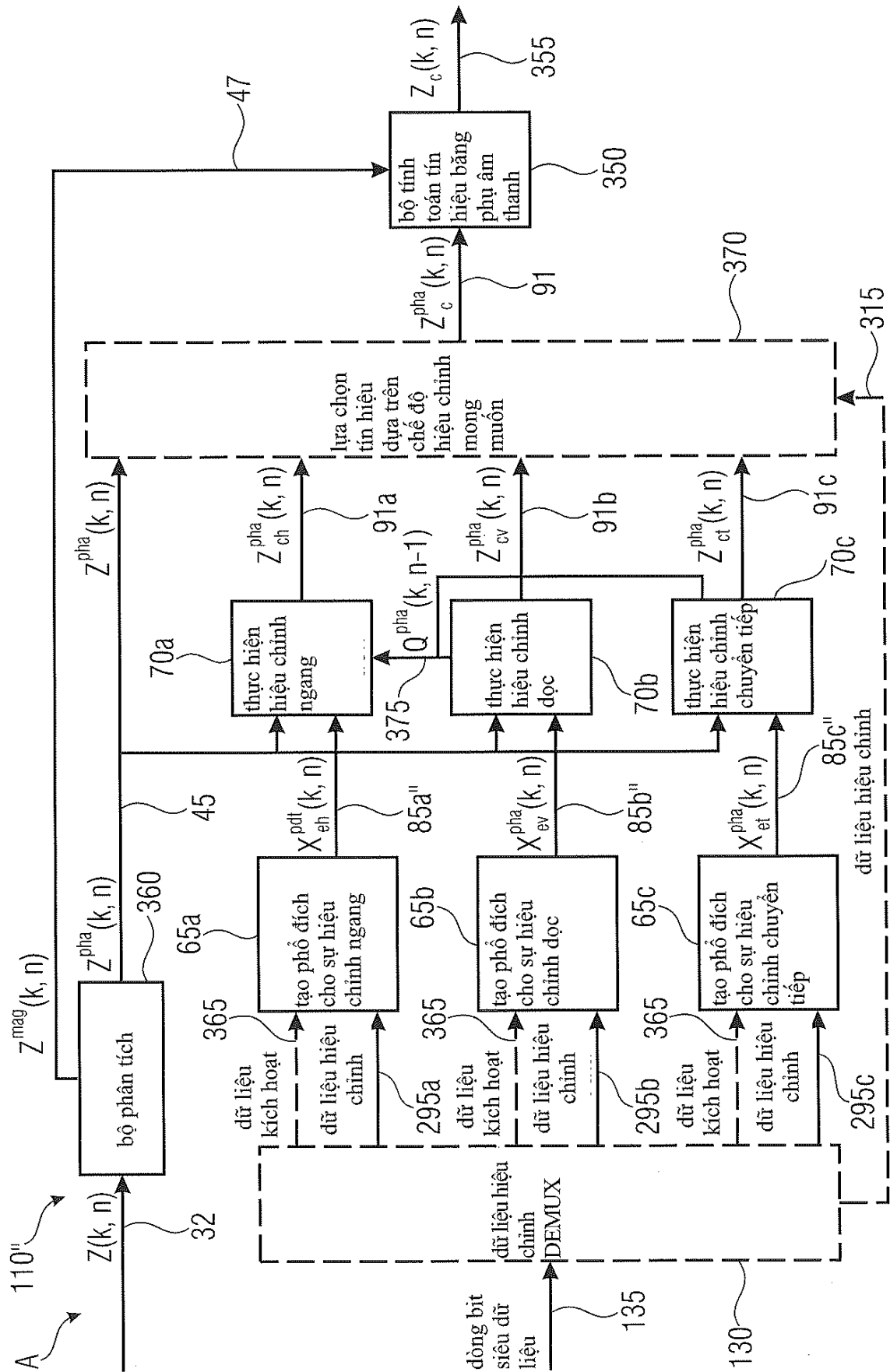


FIG 54

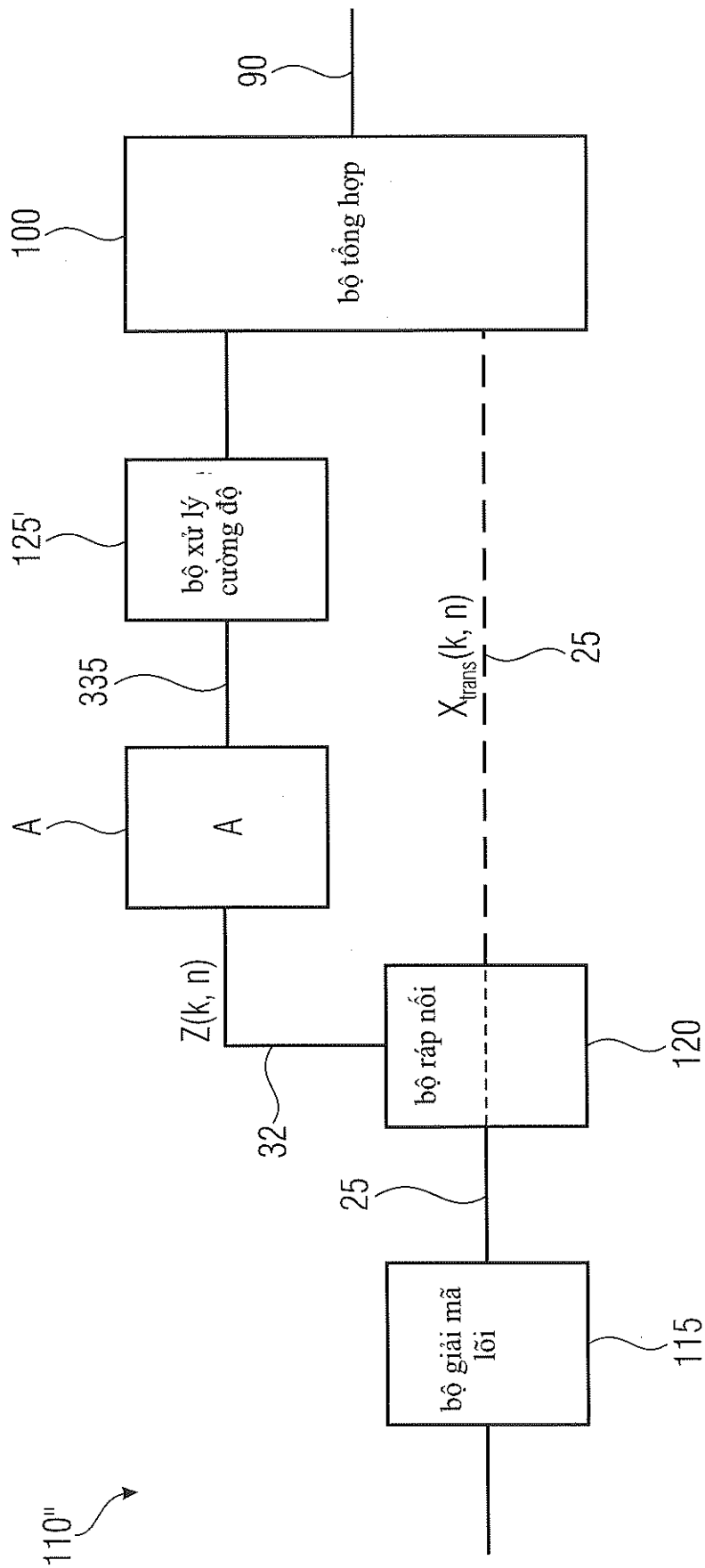


FIG 55

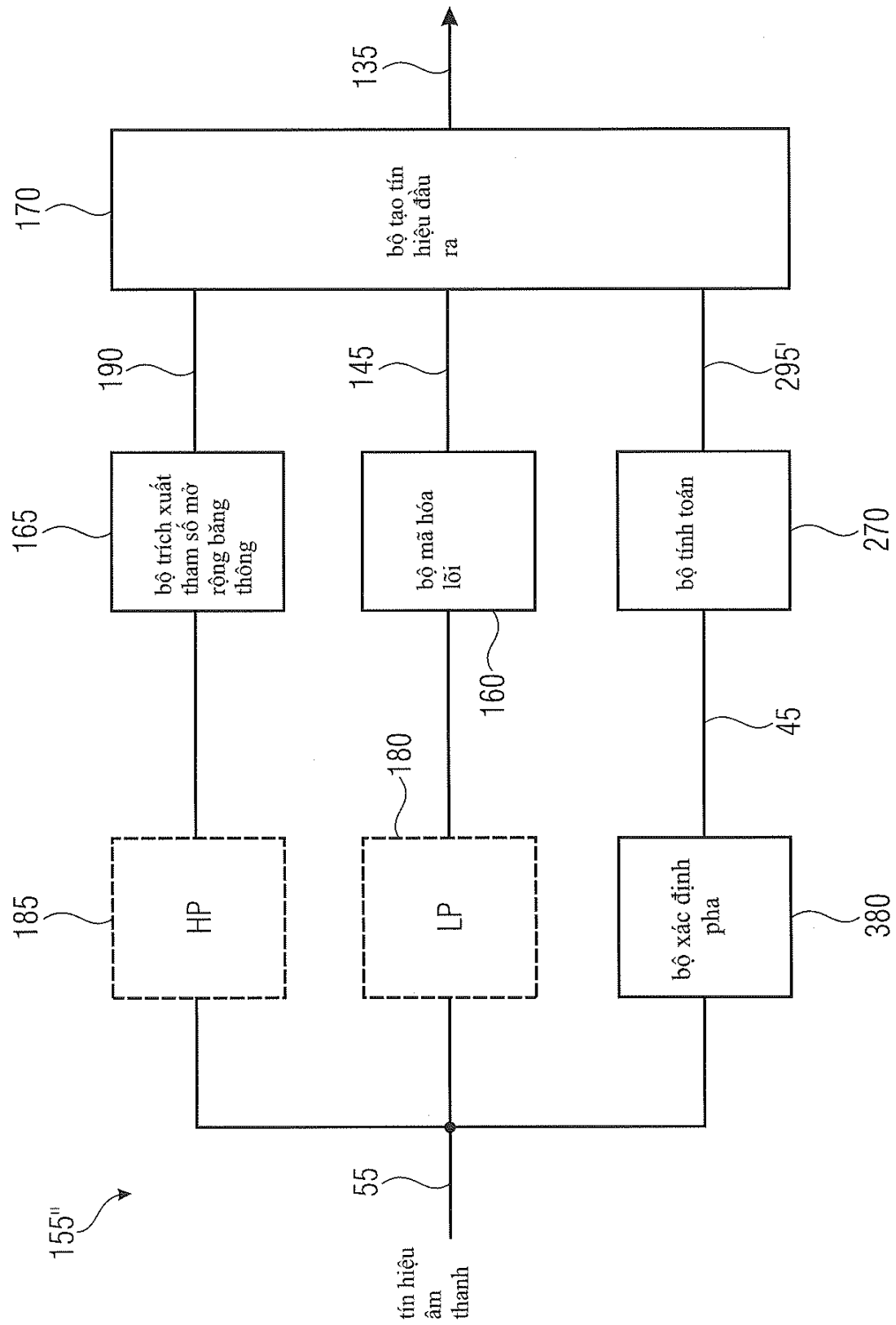


FIG 56

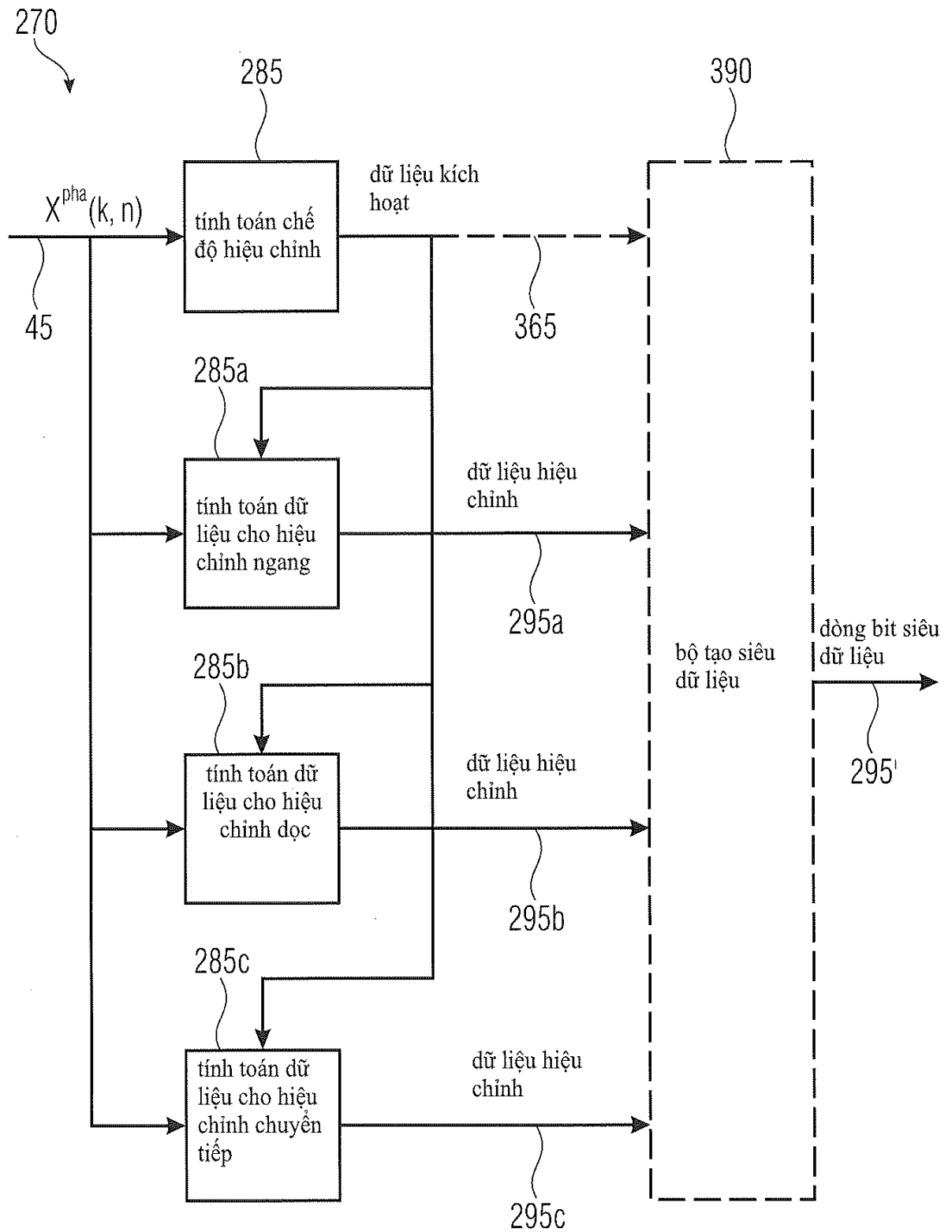


FIG 57

5800

Tạo ra phổ đích cho khung thời gian thứ nhất của tín hiệu băng phụ của tín hiệu âm thanh với bộ tạo phổ đích thứ nhất sử dụng dữ liệu hiệu chỉnh thứ nhất

5805

Hiệu chỉnh pha của tín hiệu băng phụ trong khung thời gian thứ nhất của tín hiệu âm thanh với bộ hiệu chỉnh pha thứ nhất được xác định với thuật toán hiệu chỉnh pha, trong đó sự hiệu chỉnh được thực hiện bằng việc giảm sự chênh lệch giữa số đo của tín hiệu băng phụ trong khung thời gian thứ nhất của tín hiệu âm thanh và phổ đích

5810

Tính toán tín hiệu băng phụ âm thanh cho khung thời gian thứ nhất với bộ tính toán tín hiệu băng phụ âm thanh sử dụng pha được hiệu chỉnh của khung thời gian và để tính toán các tín hiệu băng phụ âm thanh cho khung thời gian thứ hai khác với khung thời gian thứ nhất sử dụng số đo của tín hiệu băng phụ trong khung thời gian thứ hai hoặc sử dụng sự tính toán pha được hiệu chỉnh phù hợp với thuật toán hiệu chỉnh pha nữa khác với thuật toán hiệu chỉnh pha

5815

FIG 58

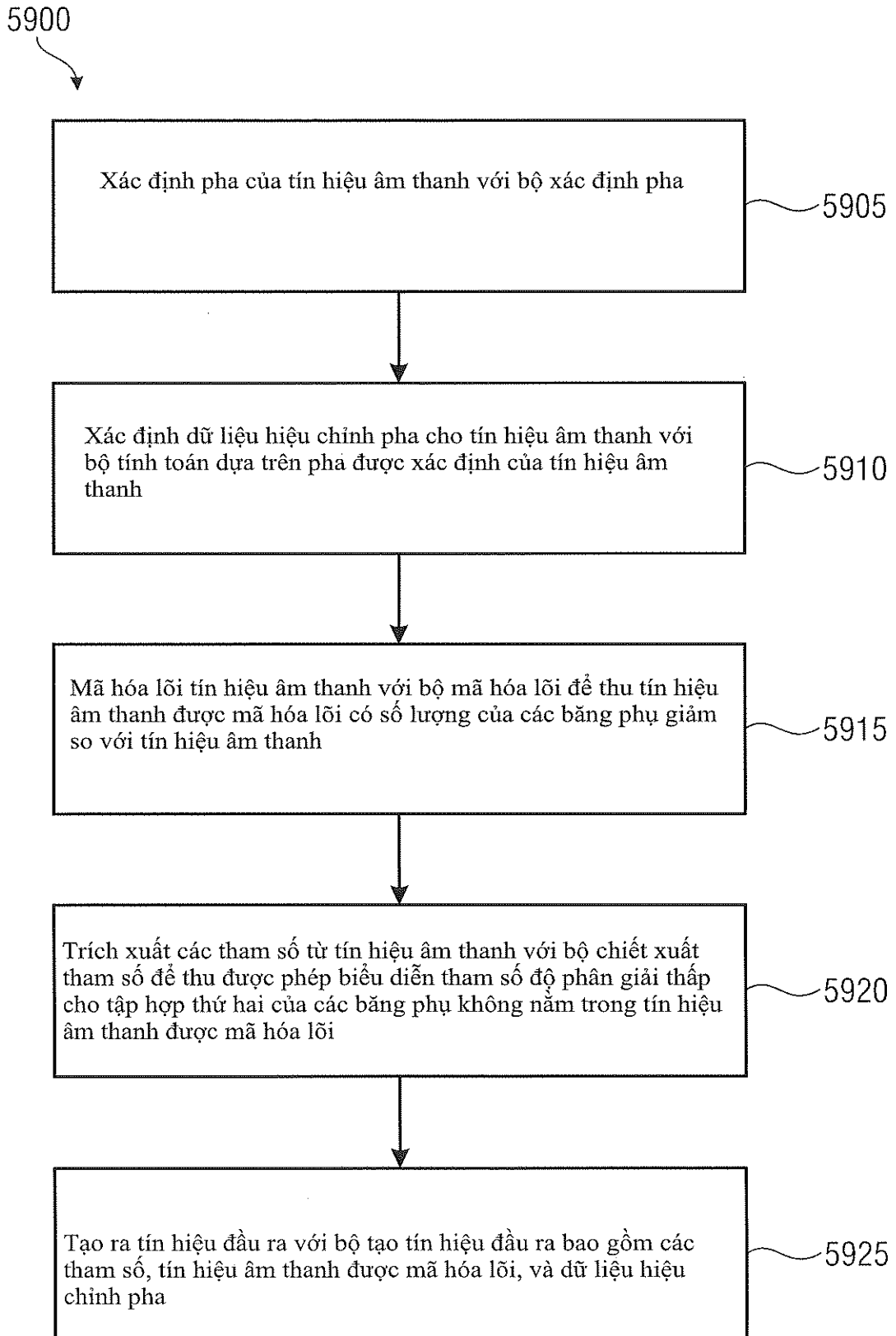


FIG 59