



(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)



1-0025567

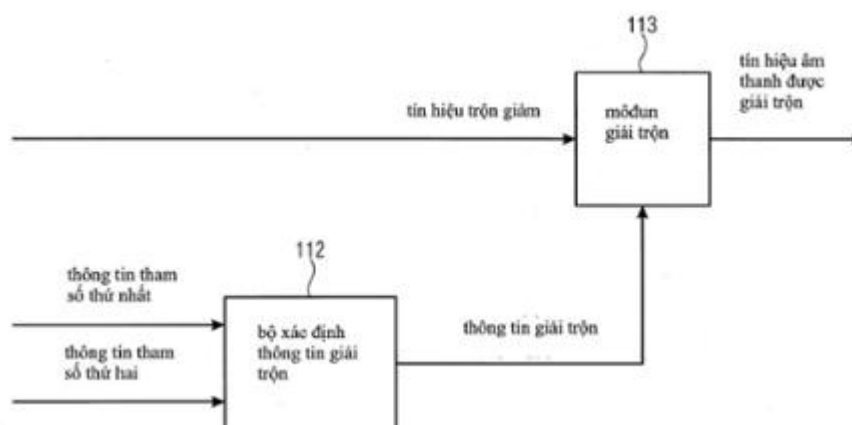
(51)⁷ G10L 19/008; G10L 19/02

(13) B

- (21) 1-2015-01536 (22) 02/10/2013
(86) PCT/EP2013/070533 02/10/2013 (87) WO/2014/053537 10/04/2014
(30) 61/710,128 05/10/2012 US; 13167485.5 13/05/2013 EP
(45) 25/09/2020 390 (43) 27/07/2015 328A
(73) Fraunhofer-Gesellschaft zur Foerderung der angewandten Forschung e.V. (DE)
Hansastrasse 27c, 80686 Muenchen, Germany
(72) DISCH, Sascha (DE); FUCHS, Harald (DE); PAULUS, Jouni (FI); TERENTIV,
Leon (DE); HELLMUTH, Oliver (DE); HERRE, Juergen (DE).
(74) CÔNG TY LUẬT TRÁCH NHIỆM HỮU HẠN AMBYS HÀ NỘI (AMBYS
HANOI)

(54) BỘ MÃ HÓA, BỘ GIẢI MÃ, HỆ THỐNG VÀ PHƯƠNG PHÁP MÃ HÓA ĐỐI
TƯỢNG ÂM THANH KHÔNG GIAN ĐA PHÂN GIẢI TƯỢNG THÍCH NGƯỢC

(57) Sáng chế đề xuất bộ giải mã để tạo ra tín hiệu âm thanh giải trộn bao gồm nhiều kênh âm thanh được giải trộn. Ngoài ra, bộ mã hóa và tín hiệu âm thanh được mã hóa được đề xuất. Bộ giải mã bao gồm bộ xác định thông tin giải trộn để xác định thông tin giải trộn bằng cách tiếp nhận thông tin phụ tham số thứ nhất trên ít nhất một tín hiệu đối tượng âm thanh và thông tin phụ tham số thứ hai trên ít nhất một tín hiệu đối tượng âm thanh, trong đó độ phân giải tần số của thông tin phụ tham số thứ hai cao hơn độ phân giải tần số của thông tin phụ tham số thứ nhất. Hơn nữa, bộ giải mã bao gồm môđun giải trộn để áp dụng thông tin giải trộn trên tín hiệu trộn giảm, biểu thị sự trộn giảm của ít nhất một tín hiệu đối tượng âm thanh, để thu tín hiệu âm thanh được giải trộn bao gồm nhiều kênh âm thanh được giải trộn. Bộ xác định thông tin giải trộn được tạo cấu hình để xác định thông tin giải trộn bằng cách thay đổi thông tin tham số thứ nhất và thông tin tham số thứ hai để thu thông tin tham số được thay đổi, sao cho thông tin tham số được thay đổi có độ phân giải tần số cao hơn độ phân giải tần số thứ nhất.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến mã hóa tín hiệu âm thanh, giải mã tín hiệu âm thanh và xử lý tín hiệu âm thanh, và cụ thể, đề cập đến bộ mã hóa, bộ giải mã và các phương pháp mã hóa đối tượng âm thanh không gian đa phân giải tương thích ngược (spatial-audio-object-coding - SAOC).

Tình trạng kỹ thuật của sáng chế

Trong các hệ thống âm thanh số hiện đại, xu hướng chính là cho phép các thay đổi liên quan đến đối tượng âm thanh của nội dung được truyền ở phía bộ tiếp nhận. Những thay đổi này bao gồm các thay đổi độ khuếch đại của các phần đã được chọn của tín hiệu âm thanh và/hoặc tái định vị trong không gian của các đối tượng âm thanh được dành riêng cho trường hợp phát lại đa kênh qua các loa được phân bố trong không gian. Điều này có thể đạt được bằng cách phân phối riêng biệt các phần khác nhau của nội dung âm thanh đến các loa khác nhau.

Nói cách khác, trong kỹ thuật xử lý âm thanh, truyền âm thanh, và lưu trữ âm thanh, nhu cầu về việc cho phép tương tác người dùng để phát lại nội dung âm thanh định hướng đối tượng cũng như nhu cầu sử dụng các khả năng mở rộng phát lại đa kênh để kết xuất riêng biệt các nội dung âm thanh hoặc các phần của các nội dung âm thanh để nâng cao được cảm giác nghe đã tăng lên. Bằng cách này, việc sử dụng nội dung âm thanh đa kênh mang đến những cải thiện đáng kể cho người dùng. Ví dụ, có thể có được cảm giác nghe ba chiều, mang đến sự thỏa mãn cải thiện về các ứng dụng giải trí cho người dùng. Tuy nhiên, nội dung âm thanh đa kênh cũng hữu dụng trong các môi trường chuyên nghiệp, ví dụ, trong các ứng dụng hội thoại bằng điện thoại, bởi vì khả năng nghe rõ người nói có thể được cải thiện bằng cách sử dụng sự phát lại âm thanh đa kênh. Có thể có các ứng dụng khác như để cho người nghe đoạn nhạc điều chỉnh riêng biệt mức phát lại và/hoặc vị trí không gian của các phần khác nhau (còn được gọi là “đối tượng âm thanh”) hoặc bản nhạc, như phần xướng âm hoặc các nhạc cụ khác nhau. Người dùng có thể thực hiện việc điều chỉnh này cho các lý do sở

thích cá nhân, để chuyển biến dễ dàng hơn một hoặc nhiều phần từ một đoạn nhạc, cho các mục đích giáo dục, karaoke, buổi diễn tập, v.v..

Việc truyền rời rạc theo hướng thẳng tất cả nội dung âm thanh đa đối tượng hoặc đa kênh dạng số, ví dụ, dưới dạng dữ liệu điều biến mã xung (pulse code modulation – PCM) hoặc các định dạng âm thanh được nén, yêu cầu tốc độ truyền dữ liệu rất cao. Tuy nhiên, cũng có những mong muốn truyền và lưu trữ dữ liệu âm thanh theo hướng hiệu quả về tốc độ truyền dữ liệu. Do đó, người ta mong muốn chấp nhận sự cân bằng hợp lý giữa chất lượng âm thanh và các yêu cầu tốc độ truyền dữ liệu để tránh được việc tải nguồn quá mức gây ra bởi các ứng dụng đa đối tượng/đa kênh.

Thực vậy, trong lĩnh vực mã hóa âm thanh, các kỹ thuật tham số dành cho truyền/lưu trữ hiệu quả tốc độ truyền dữ liệu của các tín hiệu âm thanh đa đối tượng/đa kênh được đưa ra bởi, ví dụ, nhóm các chuyên gia hình ảnh động (Moving Picture Experts Group - MPEG) và các nhóm khác nữa. Một ví dụ là MPEG Surround (MPS) là phương pháp định hướng kênh [MPS, BCC], hoặc mã hóa đối tượng âm thanh không gian MPEG (MPEG Spatial Audio Object Coding - SAOC) như phương pháp định hướng đối tượng [JSC, SAOC, SAOC1, SAOC2]. Phương pháp định hướng đối tượng khác được gọi là “phân tách nguồn được thông báo” [ISS1, ISS2, ISS3, ISS4, ISS5, ISS6]. Các kỹ thuật này nhằm khôi phục cảnh âm thanh đầu ra mong muốn hoặc đối tượng nguồn âm thanh mong muốn trên cơ sở của việc trộn giảm các kênh/các đối tượng và thông tin phụ bổ sung mô tả cảnh âm thanh được lưu trữ/được truyền và/hoặc các đối tượng nguồn âm thanh trong cảnh âm thanh.

Phép ước lượng và ứng dụng thông tin phụ liên quan đến kênh/đối tượng trong các hệ thống này được thực hiện theo cách thức chọn lọc thời gian-tần số. Do đó, các hệ thống này sử dụng các phép biến đổi thời gian-tần số như biến đổi Fourier rời rạc (Discrete Fourier Transform - DFT), biến đổi Fourier thời gian ngắn (Short Time Fourier Transform - STFT) hoặc các giàn bộ lọc như các giàn bộ lọc gương cầu phương (Quadrature Mirror Filter - QMF), v.v.. Nguyên lý cơ bản của các hệ thống này được mô tả trên Fig.4, sử dụng ví dụ MPEG SAOC.

Trong trường hợp của STFT, kích thước theo thời gian được biểu diễn bởi số khối thời gian và kích thước phổ được lấy bởi số hệ số phổ (“bin”). Trong trường hợp

của QMF, kích thước theo thời gian được biểu diễn bởi số khe thời gian và kích thước phổ được lấy bởi số dải con. Nếu độ phân giải phổ của QMF được cải thiện bởi ứng dụng sau đó của giai đoạn bộ lọc thứ hai, toàn bộ giàn bộ lọc được gọi là QMF lai và các dải con có độ phân giải tốt được gọi là các dải con lai.

Như đã chỉ ra bên trên, trong SAOC, xử lý chung được tiến hành theo cách lựa chọn thời gian-tần số và có thể được mô tả như sau trong mỗi dải tần số:

- N tín hiệu đối tượng âm thanh đầu vào $s_1 \dots s_N$ được trộn giảm còn P kênh $x_1 \dots x_P$ như phần xử lý của bộ mã hóa sử dụng ma trận trộn giảm bao gồm các phần tử $d_{1,1} \dots d_{N,P}$. Ngoài ra, bộ mã hóa trích xuất thông tin phụ mô tả các đặc tính của các đối tượng âm thanh đầu vào (môđun bộ ước lượng thông tin phụ (side-information-estimator - SEI)). Đối với MPEG SAOC, các mối quan hệ của các công suất đối tượng w.r.t với nhau là dạng cơ bản nhất của thông tin phụ này.

- (Các) tín hiệu trộn giảm và thông tin phụ được truyền/được lưu trữ. Để đạt được điều này, (các) tín hiệu âm thanh trộn giảm có thể được nén, ví dụ, sử dụng các bộ mã hóa âm thanh cảm giác đã biết như MPEG-1/2 lớp II hoặc III (aka .mp3), MPEG-2/4, mã hóa âm thanh tiên tiến (Advanced Audio Coding - AAC) v.v..

- Phía đầu tiếp nhận, bộ giải mã cố gắng dựa trên khái niệm khôi phục lại các đối tượng âm thanh gốc (“chia tách đối tượng”) từ các tín hiệu trộn giảm (đã được giải mã) sử dụng thông tin phụ được đã được truyền. Các tín hiệu đối tượng được lấy gần đúng $\hat{s}_1 \dots \hat{s}^N$ này tiếp đó được trộn trong cảnh mục tiêu được kết xuất bởi M kênh đầu ra âm thanh $\hat{y}_1 \dots \hat{y}_M$ sử dụng ma trận kết xuất được mô tả bởi các hệ số $r_{1,1} \dots r_{N,M}$ trên Fig.4. Cảnh mục tiêu mong muốn, trong trường hợp đặc biệt, có thể là phép kết xuất duy nhất một tín hiệu nguồn ngoài hỗn hợp (kịch bản phân tách nguồn), ngoài ra còn cảnh âm thanh tùy hứng khác bất kỳ chứa các đối tượng được truyền. Ví dụ, đầu ra có thể là cảnh mục tiêu kênh đơn, âm lập thể kênh đôi hoặc đa kênh 5.1.

Các hệ thống dựa trên thời gian-tần số có thể sử dụng biến đổi thời gian-tần số (t/f) với độ phân giải theo thời gian và tần số tĩnh. Lựa chọn lưới phân giải t/f cố định nhất định đặc trưng bao gồm cân bằng giữa độ phân giải thời gian và tần số.

Hiệu quả của độ phân giải t/f cố định có thể được chứng minh bằng ví dụ của các tín hiệu đối tượng đặc trưng trong hỗn hợp tín hiệu âm thanh. Ví dụ, phổ của âm tiếng nói biểu hiện cấu trúc liên quan có tính điều hòa với tần số cơ bản và một số âm bội. Năng lượng của các tín hiệu này được tập trung tại khoảng tần số nhất định. Với các tín hiệu đã nêu, độ phân giải tần số cao của phép biểu diễn t/f được sử dụng có lợi để phân tách các vùng phổ âm dải hẹp từ hỗn hợp tín hiệu. Ngược lại, các tín hiệu nhất thời, như tiếng trống, thường có cấu trúc theo thời gian riêng biệt: năng lượng đáng kể chỉ xuất hiện trong các giai đoạn ngắn của thời gian và lan truyền qua khoảng rộng của tần số. Với các tín hiệu này, độ phân giải theo thời gian ở mức cao của phép biểu diễn t/f được sử dụng là ưu điểm để phân tách phần tín hiệu nhất thời từ hỗn hợp tín hiệu.

Độ phân giải theo tần số thu được từ phép biểu diễn SAOC tiêu chuẩn bị giới hạn số lượng dải tham số, có giá trị cực đại là 28 trong SAOC tiêu chuẩn. Chúng thu được từ giàn QMF lai bao gồm phân tích QMF dải 64 với tầng lọc lai bổ sung trên các dải thấp nhất còn chia chúng thành tới 4 dải con phức hợp. Các dải tần số thu được được nhóm thành các dải tham số mô phỏng theo độ phân giải theo dải tới hạn của hệ thống thính giác của con người. Việc nhóm cho phép giảm tỉ lệ dữ liệu thông tin phụ được yêu cầu thành kích cỡ mà có thể sử dụng hiệu quả trong các ứng dụng thực tế.

Các sơ đồ mã hóa đối tượng âm thanh hiện tại chỉ đưa ra tính biến thiên bị giới hạn trong việc chọn lọc thời gian-tần số của xử lý SAOC. Ví dụ, MPEG, SAOC [SAOC] [SAOC1] [SAOC2] bị giới hạn bởi độ phân giải thời gian-tần số mà có thể thu được bởi việc sử dụng cái được gọi là Giàn bộ lọc Gương cầu Phương Lai (Lai-QMF) và tiếp theo nhóm nó thành các dải tham số. Do đó, việc hoàn trả đối tượng trong SAOC tiêu chuẩn thường trải qua từ độ phân giải tần số thô của QMF lai dẫn đến xuyên âm được điều biến có thể nghe được từ các đối tượng âm thanh khác (ví dụ các giả âm nhại tiếng nói hoặc các giả âm âm thính giác trong âm nhạc).

Một sơ đồ mã hóa đối tượng âm thanh khác tạo ra việc mã hóa đối tượng cơ sở và mã hóa đối tượng nâng cao với độ phân giải tần số cao hơn được bộc lộ trong công bố đơn đăng ký sáng chế số US 2010/0087938 A1 (ngày 08 tháng 04 năm 2010).

Hệ thống hiện có đưa ra chất lượng phân tách hợp lý nhận được tốc độ dữ liệu thấp hợp lý. Vấn đề chính là độ phân giải tần số không đủ để phân tách hoàn toàn các

âm tiếng nói. Điều đó được biểu hiện như "quảng" các đối tượng khác bao quanh các thành phần âm của đối tượng. Về mặt nhận thức điều đó có thể được quan sát như hỗn âm hoặc âm giả kiểu bộ ghi mã tiếng nói. Các tác dụng bất lợi của quảng này có thể được giảm bằng cách tăng độ phân giải tần số tham số. Đã được ghi nhận rằng độ phân giải bằng hoặc lớn hơn dải 512 (tại tốc độ lấy mẫu 44,1kHz) là đủ để đưa ra sự phân tách được cải thiện đáng kể về cảm giác trong các tín hiệu thử nghiệm. Vấn đề với độ phân giải tham số cao như vậy là lượng thông tin phụ cần tăng đáng kể, thành lượng không thực tế. Ngoài ra, khả năng tương thích với các hệ thống SAOC tiêu chuẩn hiện có sẽ bị mất.

Do đó sẽ được đánh giá cao nếu các khái niệm có thể được đề xuất mà hướng dẫn cách khắc phục những hạn chế của tình trạng kỹ thuật.

Tài liệu tham khảo

[BCC] C. Faller and F. Baumgarte, "Binaural Cue Coding - Part II: Schemes and applications," IEEE Trans. on Speech and Audio Proc., vol. 11, no. 6, Nov. 2003

[JSC] C. Faller, "Parametric Joint-Coding of Audio Sources", 120th AES Convention, Paris, 2006

[SAOC1] J. Herre, S. Disch, J. Hilpert, O. Hellmuth: "From SAC To SAOC - Recent Developments in Parametric Coding of Spatial Audio", 22nd Regional UK AES Conference, Cambridge, UK, April 2007

[SAOC2] J. Engdegård, B. Resch, C. Falch, O. Hellmuth, J. Hilpert, A. Hölzer, L. Terentiev, J. Breebaart, J. Koppens, E. Schuijers and W. Oomen: " Spatial Audio Object Coding (SAOC) – The Upcoming MPEG Standard on Parametric Object Based Audio Coding", 124th AES Convention, Amsterdam 2008

[SAOC] ISO/IEC, "MPEG audio technologies – Part 2: Spatial Audio Object Coding (SAOC)," ISO/IEC JTC1/SC29/WG11 (MPEG) International Standard 23003-2.

[AAC] M .Bosi, K. Brandenburg, S. Quackenbush, L. Fielder, K. Akagiri, H. Fuchs, M. Dietz, "ISO/IEC MPEG-2 Advanced Audio Coding", J. Audio Eng. Soc, vol 45, no 10, pp. 789-814, 1997.

[ISS1] M. Parvaix and L. Girin: “Informed Source Separation of underdetermined instantaneous Stereo Mixtures using Source Index Embedding”, IEEE ICASSP, 2010

[ISS2] M. Parvaix, L. Girin, J.-M. Brossier: “A watermarking-based method for informed source separation of audio signals with a single sensor”, IEEE Transactions on Audio, Speech and Language Processing, 2010

[ISS3] A. Liutkus and J. Pinel and R. Badeau and L. Girin and G. Richard: “Informed source separation through spectrogram coding and data embedding”, Signal Processing Journal, 2011

[ISS4] A. Ozerov, A. Liutkus, R. Badeau, G. Richard: “Informed source separation: source coding meets source separation”, IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, 2011

[ISS5] Shuhua Zhang and Laurent Girin: “An Informed Source Separation System for Speech Signals”, INTERSPEECH, 2011

[ISS6] L. Girin and J. Pinel: “Informed Audio Source Separation from Compressed Linear Stereo Mixtures”, AES 42nd International Conference: Semantic Audio, 2011

[ISS7] A. Nesbit, E. Vincent, and M. D. Plumbley: “Benchmarking flexible adaptive time-frequency transforms for underdetermined audio source separation”, IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 37-40, 2009.

Bản chất kỹ thuật của sáng chế

Mục đích của sáng chế là đề xuất các khái niệm được cải thiện đã nêu cho việc mã hóa đối tượng âm thanh. Mục đích này được giải quyết bằng bộ giải mã theo điểm 1, bằng bộ mã hóa theo điểm 9, bằng hệ thống theo điểm 14, bằng phương pháp giải mã theo điểm 15, bằng phương pháp mã hóa theo điểm 16 và bằng vật ghi có thể đọc được bằng máy tính chứa chương trình máy tính theo điểm 17 và 18.

Ngược lại với SAOC thuộc tình trạng kỹ thuật, các phương án của sáng chế đề xuất sự tham số hóa phổ, đó là

- các dòng bit tham số SAOC có nguồn từ bộ mã hóa SAOC tiêu chuẩn có thể vẫn được giải mã bởi bộ giải mã được nâng cao với chất lượng cảm giác có thể so sánh được với một dòng bit thu được với bộ giải mã tiêu chuẩn,

- các dòng bit tham số SAOC được nâng cao có thể được giải mã với bộ giải mã SAOC tiêu chuẩn với chất lượng có thể so sánh với một dòng bit có thể thu được với các dòng bit SAOC tiêu chuẩn.

- các dòng bit tham số SAOC được nâng cao có thể được giải mã với chất lượng tối ưu với bộ giải mã được nâng cao,

- bộ giải mã SAOC được nâng cao có thể điều chỉnh một cách linh hoạt mức nâng cao, ví dụ, phụ thuộc vào các nguồn điện toán có sẵn,

- các dòng bit tham số SAOC tiêu chuẩn và được nâng cao có thể được trộn, ví dụ, trong cảnh bộ phận điều khiển đa điểm (multi-point control unit - MCU), thành một dòng bit chung mà có thể được giải mã bằng bộ giải mã tiêu chuẩn hoặc được nâng cao với chất lượng được cung cấp bởi bộ giải mã, và

- phép tham số hóa bổ sung được thu gọn.

Với các đặc tính được nêu ra ở trên, sáng chế ưu tiên có sự tham số hóa mà có thể hiểu được bởi bộ giải mã SAOC tiêu chuẩn nhưng cũng cho phép phân phối hiệu quả thông tin về độ phân giải tần số cao hơn. Độ phân giải của phép biểu diễn tần số-thời gian cơ bản xác định hiệu suất cực đại của các việc nâng cao. Sáng chế định rõ phương pháp để phân phối thông tin tần số cao được nâng cao theo cách mà được thu gọn và cho phép giải mã tương thích ngược.

Chất lượng cảm giác SAOC được nâng cao có thể thu được, ví dụ, bằng việc làm thích ứng một cách linh hoạt độ phân giải thời gian/tần số của giàn bộ lọc hoặc biến đổi mà được sử dụng để ước lượng hoặc sử dụng để tổng hợp đối tượng âm thanh bổ sung các đặc tính cụ thể của đối tượng âm thanh đầu vào. Ví dụ, nếu đối tượng âm thanh là gần như - ổn định trong suốt khoảng thời gian nhất định, phép ước lượng tham số và tổng hợp được thực hiện hữu ích trên độ phân giải thời gian thô và độ phân giải tần số mịn. Nếu đối tượng âm thanh chứa các tín hiệu nhất thời hoặc bất định trong suốt khoảng thời gian nhất định, sự ước lượng tham số và tổng hợp được thực hiện

hữu ích sử dụng độ phân giải thời gian mịn và độ phân giải tần số thô. Bằng cách đó, sự thích ứng linh hoạt của giàn bộ lọc hoặc sự biến đổi cho phép để

- độ chọn lọc tần số cao trong phân tách phổ của các tín hiệu gần như - ổn định để tránh sự giao tiếp chéo các đối tượng liên kết, và

- độ chính xác theo thời gian ở mức cao cho lúc bắt đầu đối tượng hoặc các sự kiện nhất thời để giảm thiểu tiếng vang trước và sau.

Đồng thời, chất lượng SAOC truyền thống có thể thu được bởi việc ánh xạ dữ liệu SAOC tiêu chuẩn lên trên lưới thời gian-tần số được cung cấp bởi phép biến đổi thích ứng tín hiệu tương thích ngược theo sáng chế mà phụ thuộc vào thông tin phụ mô tả các đặc điểm tín hiệu đối tượng.

Để có thể giải mã cả dữ liệu SAOC tiêu chuẩn và dữ liệu SAOC được nâng cao, sử dụng một biến đổi chung, cho phép tương thích ngược trực tiếp cho các ứng dụng mà trộn bao quanh dữ liệu SAOC tiêu chuẩn và dữ liệu SAOC nâng cao mới. Sáng chế cũng cho phép nâng cao có chọn lọc thời gian-tần số vượt quá chất lượng tiêu chuẩn.

Các phương án được đề xuất không bị giới hạn bởi bất kỳ biến đổi thời gian-tần số cụ thể, nhưng có thể được ứng dụng với bất kỳ biến đổi cung cấp độ phân giải tần số cao thích đáng. Sáng chế mô tả ứng dụng cho Biến đổi Fourier rời rạc (Discrete Fourier Transform - DFT) dựa trên giàn bộ lọc với độ phân giải thời gian-tần số chuyển đổi. Trong cách tiếp cận này, các tín hiệu miền thời gian được chia nhỏ thêm thành các khối ngắn hơn, mà cũng có thể chồng lấp. Tín hiệu trong mỗi khối ngắn hơn được gán trọng số bởi hàm cửa sổ (thông thường có các giá trị lớn ở giữa và cả hai đầu được giảm dần xuống không). Cuối cùng, tín hiệu được gán trọng số được biến đổi thành miền tần số bởi sự biến đổi được lựa chọn, ở đây, bởi ứng dụng của DFT.

Bộ giải mã để tạo ra tín hiệu âm thanh giải trộn bao gồm nhiều kênh âm thanh giải trộn được đề xuất. Bộ giải mã bao gồm bộ xác định thông tin giải trộn để xác định thông tin giải trộn bằng cách tiếp nhận thông tin phụ tham số thứ nhất trên ít nhất một tín hiệu đối tượng âm thanh và thông tin phụ tham số thứ hai trên ít nhất một tín hiệu đối tượng âm thanh, trong đó độ phân giải tần số của thông tin phụ tham số thứ hai cao hơn độ phân giải tần số của thông tin phụ tham số thứ nhất. Hơn nữa, bộ giải mã bao

gồm môđun giải trộn để ứng dụng thông tin giải trộn trên tín hiệu trộn giảm, biểu thị sự trộn giảm của ít nhất một tín hiệu đối tượng âm thanh, để thu tín hiệu âm thanh giải trộn bao gồm nhiều kênh âm thanh giải trộn. Bộ xác định thông tin giải trộn được tạo cấu hình để xác định thông tin giải trộn bằng cách thay đổi thông tin tham số thứ nhất và thông tin tham số thứ hai để thu thông tin tham số được thay đổi, sao cho thông tin tham số được thay đổi có độ phân giải tần số cao hơn độ phân giải tần số thứ nhất.

Ngoài ra, bộ mã hóa để mã hóa một hoặc nhiều tín hiệu đối tượng âm thanh đầu vào được đề xuất. Bộ mã hóa bao gồm bộ phận trộn giảm để trộn giảm một hoặc nhiều tín hiệu đối tượng âm thanh đầu vào để thu một hoặc nhiều tín hiệu trộn giảm. Hơn nữa, bộ mã hóa bao gồm bộ tạo thông tin phụ tham số để tạo ra thông tin phụ tham số thứ nhất trên ít nhất một tín hiệu đối tượng âm thanh và thông tin phụ tham số thứ hai trên ít nhất một tín hiệu đối tượng âm thanh, sao cho độ phân giải tần số của thông tin phụ tham số thứ hai cao hơn độ phân giải tần số của thông tin phụ tham số thứ nhất.

Hơn nữa, tín hiệu âm thanh được mã hóa được đề xuất. Tín hiệu âm thanh được mã hóa bao gồm phần trộn giảm, biểu thị sự trộn giảm của một hoặc nhiều tín hiệu đối tượng âm thanh đầu vào, và phần thông tin phụ tham số bao gồm thông tin phụ tham số thứ nhất trên ít nhất một tín hiệu đối tượng âm thanh và thông tin phụ tham số thứ hai trên ít nhất một tín hiệu đối tượng âm thanh. Độ phân giải tần số của thông tin phụ tham số thứ hai cao hơn độ phân giải tham số của thông tin phụ tham số thứ nhất.

Ngoài ra, hệ thống được đề xuất. Hệ thống bao gồm bộ mã hóa như được mô tả ở trên và bộ giải mã như được mô tả ở trên. Bộ mã hóa được tạo cấu hình để mã hóa một hoặc nhiều tín hiệu đối tượng âm thanh đầu vào bằng cách thu một hoặc nhiều tín hiệu trộn giảm biểu thị sự trộn giảm của một hoặc nhiều tín hiệu đối tượng âm thanh đầu vào, bằng cách thu thông tin phụ tham số thứ nhất trên ít nhất một tín hiệu đối tượng âm thanh, và bằng cách thu thông tin phụ tham số thứ hai trên ít nhất một tín hiệu đối tượng âm thanh, trong đó độ phân giải tần số của thông tin phụ tham số thứ hai cao hơn độ phân giải tần số của thông tin phụ tham số thứ nhất. Bộ giải mã được tạo cấu hình để tạo ra tín hiệu âm thanh giải trộn dựa trên một hoặc nhiều tín hiệu trộn giảm, và dựa trên thông tin phụ tham số thứ nhất và thông tin phụ tham số thứ hai.

Bộ mã hóa được tạo cấu hình để mã hóa một hoặc nhiều tín hiệu đối tượng âm thanh đầu vào bằng cách thu một hoặc nhiều tín hiệu trộn giảm biểu thị sự trộn giảm của một hoặc nhiều tín hiệu đối tượng âm thanh đầu vào, bằng cách thu được thông tin phụ tham số thứ nhất trên ít nhất một tín hiệu đối tượng âm thanh, và bằng cách thu được thông tin phụ tham số thứ hai trên ít nhất một tín hiệu đối tượng âm thanh, trong đó độ phân giải tần số của thông tin phụ tham số thứ hai cao hơn độ phân giải tần số của thông tin phụ tham số thứ nhất. Bộ giải mã được tạo cấu hình để tạo ra tín hiệu đầu ra âm thanh dựa trên một hoặc nhiều tín hiệu trộn giảm, và dựa trên thông tin phụ tham số thứ nhất và thông tin phụ tham số thứ hai.

Hơn nữa, phương pháp để tạo ra tín hiệu âm thanh giải trộn bao gồm nhiều kênh âm thanh giải trộn được đề xuất. Phương pháp bao gồm:

- Xác định thông tin giải trộn bằng cách tiếp nhận thông tin phụ tham số thứ nhất trên ít nhất một tín hiệu đối tượng âm thanh và thông tin phụ tham số thứ hai trên ít nhất một tín hiệu đối tượng âm thanh, trong đó độ phân giải tần số của thông tin phụ tham số thứ hai cao hơn độ phân giải tần số của thông tin phụ tham số thứ nhất. Và:

- Áp dụng thông tin giải trộn trên tín hiệu trộn giảm, biểu thị sự trộn giảm của ít nhất một tín hiệu đối tượng âm thanh, để thu tín hiệu âm thanh giải trộn bao gồm nhiều kênh âm thanh giải trộn.

Xác định thông tin giải trộn bao gồm thay đổi thông tin tham số thứ nhất và thông tin tham số thứ hai để thu được thông tin tham số được thay đổi, sao cho thông tin tham số được thay đổi có độ phân giải tần số cao hơn độ phân giải tần số thứ nhất.

Ngoài ra, phương pháp mã hóa một hoặc nhiều tín hiệu đối tượng âm thanh đầu vào được đề xuất. Phương pháp gồm có:

- Trộn giảm một hoặc nhiều tín hiệu đối tượng âm thanh đầu vào để thu một hoặc nhiều tín hiệu trộn giảm. Và:

- Tạo ra thông tin phụ tham số thứ nhất trên ít nhất một tín hiệu đối tượng âm thanh và thông tin phụ tham số thứ hai trên ít nhất một tín hiệu đối tượng âm thanh, sao cho độ phân giải tần số của thông tin phụ tham số thứ hai cao hơn độ phân giải tần số của thông tin phụ tham số thứ nhất.

Ngoài ra, chương trình máy tính để thực hiện một trong các phương pháp được mô tả ở trên khi chạy trên máy tính hoặc bộ xử lý tín hiệu được đề xuất.

Các phương án ưu tiên sẽ được đề xuất trong các điểm yêu cầu bảo hộ phụ thuộc.

Mô tả vắn tắt các hình vẽ

Sau đây, các phương án của sáng chế được mô tả chi tiết hơn với sự tham chiếu đến các hình vẽ, trong đó

Fig.1a minh họa bộ giải mã theo phương án của sáng chế,

Fig.1b minh họa bộ giải mã theo phương án khác của sáng chế,

Fig.2a minh họa bộ mã hóa theo phương án của sáng chế,

Fig.2b minh họa bộ mã hóa theo phương án khác của sáng chế,

Fig.2c minh họa tín hiệu âm thanh được mã hóa theo phương án của sáng chế,

Fig.3 minh họa hệ thống theo phương án của sáng chế,

Fig.4 thể hiện sơ đồ khối dạng giản đồ tổng quan khái niệm của hệ thống SAOC.

Fig.5 thể hiện sơ đồ minh họa dạng giản đồ phép biểu diễn thời gian-phổ của tín hiệu âm thanh kênh đơn,

Fig.6 thể hiện sơ đồ khối dạng giản đồ phép tính toán có chọn lọc thời gian-tần số của thông tin phụ trong bộ mã hóa SAOC.

Fig.7 minh họa phép biểu diễn tương thích ngược theo phương án của sáng chế,

Fig.8 minh họa đường cong chênh lệch giữa giá trị tham số đúng và giá trị trung bình độ phân giải ở mức thấp theo phương án của sáng chế,

Fig.9 là hình vẽ minh họa ở mức cao của bộ mã hóa được nâng cao cung cấp dòng bit tương thích ngược với các nâng cao theo phương án của sáng chế,

Fig.10 minh họa sơ đồ khối của bộ mã hóa theo phương án cụ thể thực hiện đường tham số của bộ mã hóa,

Fig.11 là hình vẽ sơ đồ khối ở mức cao của bộ giải mã được nâng cao theo phương án của sáng chế mà có thể giải mã cả hai dòng bit tiêu chuẩn và dòng bit được nâng cao,

Fig.12 minh họa sơ đồ khối minh họa phương án của bộ phận giải mã PSI được nâng cao,

Fig.13 là hình vẽ sơ đồ khối của việc giải mã dòng bit SAOC tiêu chuẩn với bộ giải mã SAOC được nâng cao theo phương án của sáng chế,

Fig.14 là hình vẽ các khối chức năng chính của bộ giải mã theo phương án của sáng chế,

Fig.15 minh họa tín hiệu nhiều và âm, cụ thể, minh họa phổ công suất độ phân giải cao và các phép khôi phục thô tương ứng,

Fig.16 minh họa phép biến đổi cho cả hai tín hiệu ví dụ, cụ thể, minh họa các thừa số hiệu chỉnh cho các tín hiệu ví dụ,

Fig.17 minh họa các hệ số hiệu chỉnh ban đầu và dự báo tuyến tính giảm bậc dựa vào phép tính gần đúng cho cả hai tín hiệu ví dụ, và

Fig.18 minh họa kết quả của việc ứng dụng các hệ số hiệu chỉnh mô hình trên các phép khôi phục thô.

Mô tả chi tiết sáng chế

Trước khi mô tả các phương án của sáng chế, các thông tin bổ sung thuộc tình trạng kỹ thuật của các hệ thống SAOC được cung cấp.

Fig.4 thể hiện cách sắp xếp thông thường của bộ mã hóa SAOC 10 và bộ giải mã SAOC 12. Bộ mã hóa SAOC 10 nhận N đối tượng đầu vào, tức là các tín hiệu âm thanh từ s_1 đến s_N . Cụ thể, bộ mã hóa 10 bao gồm bộ trộn giảm 16 nhận các tín hiệu âm thanh từ s_1 đến s_N và trộn giảm các tín hiệu âm thanh này thành tín hiệu trộn giảm 18. Hoặc, sự trộn có thể được cung cấp bên ngoài ("sự trộn giảm mỹ thuật") và hệ thống ước lượng thông tin phụ bổ sung để làm khớp sự trộn giảm được đề xuất với sự trộn giảm được tính toán. Trên Fig.4, tín hiệu trộn giảm được thể hiện là tín hiệu kênh P . Do đó, bất kỳ cấu hình tín hiệu trộn giảm đơn ($P=1$), lập thể ($P=2$) hoặc đa kênh ($P>2$) có thể hiểu được.

Trong trường hợp trộn lập thể, các kênh của tín hiệu trộn giảm 18 được chỉ ra là $L0$ và $R0$, trong trường hợp trộn đơn, các kênh của tín hiệu trộn giảm được chỉ ra đơn

giản là $L0$. Để cho phép bộ giải mã SAOC 12 khôi phục các đối tượng riêng biệt s_I đến s_N , bộ ước lượng thông tin phụ 17 cung cấp cho bộ giải mã SAOC 12 thông tin phụ bao gồm các tham số SAOC. Ví dụ, trong trường hợp trộn giảm lập thể, các tham số SAOC gồm có các chênh lệch mức đối tượng (object level differences - OLD), các tương quan liên đối tượng (inter-object correlations - IOC) (các tham số tương quan chéo liên đối tượng), các giá trị độ khuếch đại trộn giảm (downmix gain values - DMG) và các chênh lệch mức kênh trộn giảm (downmix channel level differences - DCLD). Thông tin phụ 20, bao gồm các tham số SAOC cùng với tín hiệu trộn giảm 18, tạo ra dòng dữ liệu đầu ra SAOC được nhận bởi bộ giải mã SAOC 12.

Bộ giải mã SAOC 12 gồm có bộ trộn tăng để nhận tín hiệu trộn giảm 18 cũng như thông tin phụ 20 để khôi phục và kết xuất các tín hiệu âm thanh $\hat{s}_I$ và $\hat{s}_N$ lên trên tập hợp các kênh được chọn $\hat{y}_I$ đến $\hat{y}_M$, của người dùng bất kỳ với việc kết xuất được mô tả sơ bộ bởi thông tin kết xuất 26 đưa vào trong bộ giải mã SAOC 12.

Các tín hiệu âm thanh s_I đến s_N có thể được nhập vào cho bộ mã hóa 10 trong miền mã hóa bất kỳ, như, trong miền thời gian hoặc phổ. Trong trường hợp các tín hiệu âm thanh s_I đến s_N được đưa vào trong bộ mã hóa 10 trong miền thời gian, như PCM được mã hóa, bộ mã hóa 10 có thể sử dụng giàn bộ lọc, như giàn QMF lai, để truyền các tín hiệu trong miền phổ, trong đó các tín hiệu âm thanh được biểu diễn trong vài dải con được kết hợp với các phần phổ khác nhau, tại độ phân giải giàn bộ lọc cụ thể. Nếu các tín hiệu âm thanh s_I đến s_N sẵn sàng trong phép biểu diễn được mong đợi bởi bộ mã hóa 10, các tín hiệu âm thanh không phải thực hiện sự phân giải phổ.

Fig.5 thể hiện tín hiệu âm thanh trong miền phổ vừa được nêu. Như có thể được thấy, tín hiệu âm thanh được thể hiện là nhiều tín hiệu dải con. Mỗi tín hiệu dải con 30_1 đến 30_K bao gồm chuỗi theo thời gian các giá trị dải con được chỉ ra bởi các hộp nhỏ 32. Như có thể được thấy, các giá trị dải con 32 của các tín hiệu dải con 30_1 đến 30_K được đồng bộ với nhau theo thời gian do đó, đối với mỗi khe thời gian của giàn bộ lọc liên tiếp 34, mỗi dải con 30_1 đến 30_K gồm có chính xác một giá trị dải con 32. Như được minh họa bởi trục tần số 36, các tín hiệu dải con 30_1 đến 30_K được kết hợp với

các khoảng tần số khác nhau, và như được minh họa bởi trục thời gian 38, các khe thời gian của giàn bộ lọc 34 được sắp đặt liên tiếp theo thời gian.

Như được chỉ ra bên trên, bộ trích xuất thông tin phụ 17 của Fig.4 tính toán các tham số SAOC từ các tín hiệu âm thanh đầu vào s_1 đến s_N . Theo tiêu chuẩn SAOC được cung cấp hiện tại, bộ mã hóa 10 thực hiện việc tính toán này trong độ phân giải thời gian/tần số mà độ phân giải này có thể bị giảm đi tương đối với độ phân giải thời gian/tần số ban đầu như được xác định bởi các khe thời gian của giàn bộ lọc 34 và phép phân tích dải con, bởi một lượng nhất định, với lượng nhất định này được báo hiệu cho phía bộ giải mã trong thông tin phụ 20. Các nhóm khe thời gian của giàn bộ lọc liên tiếp 34 có thể tạo ra khung SAOC 41. Ngoài ra, số lượng dải tham số trong khung SAOC 41 được truyền trong thông tin phụ 20. Do đó, miền thời gian/tần số được chia thành các ô thời gian/tần số được minh họa trên Fig.5 bởi các đường nét đứt 42. Trên Fig.5 các dải tham số được phân bố theo cùng cách thức trong các khung SAOC được mô tả khác nhau 41 sao cho sự sắp xếp đều đặn các ô thời gian/tần số thu được. Tuy nhiên, nói chung, các dải tham số có thể thay đổi từ một khung SAOC 41 đến khung tiếp theo, phụ thuộc vào các nhu cầu khác nhau về sự phân giải phổ trong các khung SAOC 41 tương ứng. Hơn nữa, chiều dài của các khung SAOC 41 cũng có thể thay đổi. Kết quả là, sự sắp xếp các ô thời gian/tần số có thể là không đều. Tuy nhiên, các ô thời gian/tần số trong khung SAOC cụ thể 41 thường có cùng khoảng thời gian và được sắp theo hướng thời gian, cụ thể tất cả các ô t/f trong khung SAOC 41 đã nêu bắt đầu tại điểm bắt đầu khung SAOC 41 đã cho và kết thúc tại điểm kết thúc khung SAOC 41 này.

Bộ trích xuất thông tin phụ 17 được mô tả trên Fig.4 tính toán các tham số SAOC theo công thức dưới đây. Cụ thể, bộ trích xuất thông tin phụ 17 tính toán các chênh lệch mức đối tượng cho mỗi đối tượng i là

$$OLD_i^{l,m} = \frac{\sum_{n \in I} \sum_{k \in m} x_i^{n,k} x_i^{n,k*}}{\max_j (\sum_{n \in I} \sum_{k \in m} x_j^{n,k} x_j^{n,k*})}$$

trong đó các tổng và chỉ số n và k , lần lượt, xuyên suốt tất cả các chỉ số thời gian 34, và tất cả các chỉ số phổ 30 thuộc về ô thời gian/tần số 42 nhất định, được tham chiếu bởi các chỉ số l cho khung SAOC (hoặc khe thời gian xử lý) và m cho dải tham số, và

$x_i^{n,k*}$ là liên hợp phức của $x_i^{n,k}$. Do đó, các năng lượng của tất cả các giá trị dải con x_i của tín hiệu hoặc đối tượng âm thanh i được tính tổng lại và được chuẩn hóa đến giá trị năng lượng cao nhất của ô đó trong số tất cả các đối tượng hoặc các tín hiệu âm thanh.

Hơn nữa, bộ trích xuất thông tin phụ SAOC 17 có thể tính toán phép đo tương tự của các ô thời gian/tần số tương ứng của các cặp đối tượng đầu vào khác nhau s_I đến s_N . Mặc dù bộ trích xuất thông tin phụ SAOC 17 có thể tính toán phép đo tương tự giữa tất cả các cặp đối tượng đầu vào s_I đến s_N , bộ trích xuất thông tin phụ SAOC 17 cũng có thể nhấn mạnh việc tín hiệu hóa các phép đo tương tự hoặc giới hạn việc tính toán của các phép đo tương tự cho các đối tượng âm thanh s_I đến s_N mà các đối tượng này tạo thành các kênh bên trái hoặc phải của kênh lập thể thông thường. Trong trường hợp bất kỳ, phép đo tương tự được gọi là tham số hiệu chỉnh chéo liên đối tượng $IOC_{i,j}^{l,m}$. Phép toán như sau

$$IOC_{i,j}^{l,m} = IOC_{j,i}^{l,m} = Re \left\{ \frac{\sum_{n \in I} \sum_{k \in m} x_i^{n,k} x_j^{n,k*}}{\sqrt{\sum_{n \in I} \sum_{k \in m} x_i^{n,k} x_i^{n,k*} \sum_{n \in I} \sum_{k \in m} x_j^{n,k} x_j^{n,k*}}} \right\}$$

với các chỉ số lặp lại n và k xuyên suốt tất cả các giá trị dải con thuộc ô thời gian/tần số nhất định 42, i và j chỉ ra cặp đối tượng âm thanh nhất định s_I đến s_N , và $Re\{ \}$ chỉ ra phép toán chỉ giữ lại phần thực (cụ thể, loại bỏ phần ảo) của đối số giá trị phức.

Bộ trộn giảm 16 của Fig.4 trộn giảm các đối tượng từ s_I đến s_N bằng cách sử dụng các độ khuếch đại được ứng dụng cho từng đối tượng s_I đến s_N . Tức là, thừa số độ khuếch đại d_i được ứng dụng cho đối tượng i và tiếp đó tất cả các đối tượng được gán trọng số s_I đến s_N được cộng lại để thu được tín hiệu trộn đơn, mà được minh họa trên Fig.4 nếu $P=1$. Trong ví dụ khác, trường hợp của tín hiệu trộn giảm hai kênh, được mô tả trên Fig.4 nếu $P=2$, thừa số độ khuếch đại $d_{1,i}$ được ứng dụng cho đối tượng i và tiếp đó tất cả các đối tượng được khuếch đại nhờ độ khuếch đại này được cộng lại để thu được kênh trộn giảm trái $L0$, và các thừa số độ khuếch đại $d_{2,i}$ được ứng dụng cho đối tượng i và tiếp đó các đối tượng được khuếch đại nhờ độ khuếch đại được cộng lại để thu được kênh trộn giảm phải $R0$. Xử lý tương tự với bên trên được ứng dụng trong trường hợp trộn giảm đa kênh ($P>2$).

Nguyên tắc trộn giảm này được báo hiệu cho phía bộ giải mã bởi các hệ số khếch đại trộn giảm DMG và trong trường hợp của tín hiệu trộn giảm lập thể, bởi các chênh lệch mức kênh trộn giảm (downmix channel level differences - $DCLD_i$).

Các hệ số khếch đại trộn giảm được tính toán theo:

$$DMG_i = 20 \log_{10}(d_i + \varepsilon) \quad , \text{ (trộn giảm đơn),}$$

$$DMG_i = 10 \log_{10}(d_{1,i}^2 + d_{2,i}^2 + \varepsilon) \quad , \text{ (trộn giảm lập thể),}$$

trong đó ε là số nhỏ chẳng hạn như 10^{-9} .

Đối với các DCLD công thức sau đây được áp dụng:

$$DCLD_i = 20 \log_{10} \left(\frac{d_{1,i}}{d_{2,i} + \varepsilon} \right).$$

Trong chế độ thông thường, bộ trộn giảm 16 tạo ra tín hiệu trộn giảm theo:

$$(LO) = (d_i) \begin{pmatrix} s_1 \\ \vdots \\ s_N \end{pmatrix}$$

đối với trộn giảm đơn, hoặc

$$\begin{pmatrix} LO \\ RO \end{pmatrix} = \begin{pmatrix} d_{1,i} \\ d_{2,i} \end{pmatrix} \begin{pmatrix} s_1 \\ \vdots \\ s_N \end{pmatrix}$$

đối với trộn giảm lập thể, tương ứng.

Do đó, theo các công thức nêu trên, các tham số OLD và IOC là hàm của các tín hiệu âm thanh và các tham số DMG và DCLD là hàm của các hệ số trộn giảm d . Nhân đây, lưu ý rằng d có thể thay đổi theo thời gian và tần số.

Do đó, trong chế độ thông thường, bộ trộn giảm 16 trộn tất cả các đối tượng s_1 đến s_N mà không có sự ưu tiên nào, tức là, xử lý tất cả các đối tượng s_1 đến s_N ngang nhau.

Ở phía bộ giải mã, bộ trộn tăng thực hiện đảo ngược quy trình trộn giảm và thực hiện: “thông tin kết xuất” 26 được biểu hiện bởi ma trận R (trong tài liệu đôi khi còn được gọi là A) trong một bước tính toán, tức là, trong trường hợp trộn giảm hai kênh

$$\begin{pmatrix} \hat{y}_1 \\ \vdots \\ \hat{y}_M \end{pmatrix} = \mathbf{RED}^* (\mathbf{DED}^*)^{-1} \begin{pmatrix} L\mathbf{0} \\ R\mathbf{0} \end{pmatrix},$$

trong đó ma trận $\mathbf{E}$ là hàm của các tham số OLD và IOC, và ma trận $\mathbf{D}$ chứa các hệ số trộn giảm là

$$\mathbf{D} = \begin{pmatrix} d_{1,1} & \cdots & d_{1,N} \\ \vdots & \ddots & \vdots \\ d_{P,1} & \cdots & d_{P,N} \end{pmatrix},$$

và trong đó $\mathbf{D}^*$ biểu thị chuyển vị phức hợp của $\mathbf{D}$. Ma trận $\mathbf{E}$ là ma trận hiệp phương sai được ước lượng của các đối tượng âm thanh s_1 đến s_N . Trong các phương án SAOC hiện tại, phép toán ma trận hiệp phương sai đã được ước lượng $\mathbf{E}$ thường được thực hiện trong phép phân giải phổ/thời gian của các tham số SAOC, cụ thể, đối với mỗi (l,m) , sao cho ma trận hiệp phương sai được ước lượng có thể được viết là $\mathbf{E}^{l,m}$. Ma trận hiệp phương sai được ước lượng $\mathbf{E}^{l,m}$ có kích cỡ $N \times N$ với các hệ số của nó được định rõ là

$$e_{i,j}^{l,m} = \sqrt{OLD_i^{l,m} OLD_j^{l,m} IOC_{i,j}^{l,m}}.$$

Do đó, ma trận $\mathbf{E}^{l,m}$ với

$$\mathbf{E}^{l,m} = \begin{pmatrix} e_{1,1}^{l,m} & \cdots & e_{1,N}^{l,m} \\ \vdots & \ddots & \vdots \\ e_{N,1}^{l,m} & \cdots & e_{N,N}^{l,m} \end{pmatrix}$$

có các chênh lệch mức đối tượng dọc theo đường chéo của ma trận, cụ thể, $e_{i,j}^{l,m} = OLD_i^{l,m}$ đối với $i=j$, vì $OLD_i^{l,m} = OLD_j^{l,m}$ và $IOC_{i,j}^{l,m} = 1$ đối với $i=j$. Phía ngoài đường chéo của ma trận, ma trận hiệp phương sai được ước lượng $\mathbf{E}$ có các hệ số ma trận biểu diễn số trung bình nhân của các chênh lệch mức đối tượng của các đối tượng i và j , lần lượt, được gán trọng số với phép đo tương quan chéo liên đối tượng $IOC_{i,j}^{l,m}$.

Fig.6 thể hiện một nguyên lý có thể sử dụng để thực hiện trên một ví dụ của bộ ước lượng thông tin phụ (Side-information estimator - SIE) như là một phần của bộ mã hóa SAOC 10. Bộ mã hóa SAOC 10 gồm có bộ trộn 16 và bộ ước lượng thông tin phụ

(SEI) 17. SEI theo khái niệm gồm có hai môđun: Một môđun 45 để tính toán phép biểu diễn t/f dựa trên thời gian ngắn (ví dụ, STFT hoặc QMF) của mỗi tín hiệu. Phép biểu diễn t/f thời gian ngắn được tính toán được cấp vào trong môđun thứ hai 46, môđun ước lượng thông tin phụ chọn lọc t/f (t/f-SEI). Môđun t/f-SEI 46 tính toán thông tin phụ cho mỗi ô t/f. Trong các phương án SAOC hiện tại, phép biến đổi thời gian/tần số được cố định và giống nhau đối với tất cả các đối tượng âm thanh s_1 đến s_N . Hơn nữa, các tham số SAOC được xác định trên các khung SAOC mà các khung này là giống nhau đối với tất cả các đối tượng âm thanh và có cùng độ phân giải thời gian/tần số cho tất cả các đối tượng âm thanh s_1 đến s_N , do đó bỏ đi các yêu cầu đối tượng cụ thể cho độ phân giải thời gian mịn trong một số trường hợp hoặc độ phân giải phổ mịn trong các trường hợp khác.

Dưới đây, các phương án của sáng chế được mô tả.

Fig.1a minh họa bộ giải mã để tạo ra tín hiệu âm thanh giải trộn bao gồm nhiều kênh âm thanh giải trộn theo phương án của sáng chế.

Bộ giải mã bao gồm bộ xác định thông tin giải trộn 112 để xác định thông tin giải trộn bằng cách tiếp nhận thông tin phụ tham số thứ nhất trên ít nhất một tín hiệu đối tượng âm thanh và thông tin phụ tham số thứ hai trên ít nhất một tín hiệu đối tượng âm thanh, trong đó độ phân giải tần số của thông tin phụ tham số thứ hai cao hơn độ phân giải tần số của thông tin phụ tham số thứ nhất.

Hơn nữa, bộ giải mã bao gồm môđun giải trộn 113 để áp dụng thông tin giải trộn trên tín hiệu trộn giảm, biểu thị phép trộn giảm của ít nhất một tín hiệu đối tượng âm thanh, để thu tín hiệu âm thanh giải trộn bao gồm nhiều kênh âm thanh giải trộn.

Bộ xác định thông tin giải trộn 112 được tạo cấu hình để xác định thông tin giải trộn bằng cách thay đổi thông tin tham số thứ nhất và thông tin tham số thứ hai để thu thông tin tham số được thay đổi, sao cho thông tin tham số được thay đổi có độ phân giải tần số cao hơn độ phân giải tần số thứ nhất.

Fig.1b minh họa bộ giải mã để tạo ra tín hiệu âm thanh giải trộn bao gồm nhiều kênh âm thanh giải trộn theo phương án của sáng chế. Bộ giải mã theo Fig.1b còn bao gồm bộ phận biến đổi thứ nhất 111 để biến đổi đầu vào trộn giảm, được biểu diễn

trong miền thời gian, để thu tín hiệu trộn giảm, được biểu diễn trong miền thời gian-tần số. Hơn nữa, bộ giải mã của Fig.1b bao gồm bộ phận biến đổi thứ hai 114 để biến đổi tín hiệu âm thanh giải trộn từ miền thời gian-tần số thành miền thời gian.

Fig.2a minh họa bộ mã hóa để mã hóa một hoặc nhiều tín hiệu đối tượng âm thanh đầu vào theo phương án của sáng chế.

Bộ mã hóa bao gồm bộ phận trộn giảm 91 để trộn giảm một hoặc nhiều tín hiệu đối tượng âm thanh đầu vào để thu một hoặc nhiều tín hiệu trộn giảm.

Hơn nữa, bộ mã hóa bao gồm bộ tạo thông tin phụ tham số 93 tạo ra thông tin phụ tham số thứ nhất trên ít nhất một tín hiệu đối tượng âm thanh và thông tin phụ tham số thứ hai trên ít nhất một tín hiệu đối tượng âm thanh, sao cho độ phân giải tần số của thông tin phụ tham số thứ hai cao hơn độ phân giải tần số của thông tin phụ tham số thứ nhất.

Fig.2b minh họa bộ mã hóa để mã hóa một hoặc nhiều tín hiệu đối tượng âm thanh đầu vào theo phương án khác của sáng chế. Bộ mã hóa của Fig.2b bao gồm bộ phận biến đổi 92 để biến đổi một hoặc nhiều tín hiệu đối tượng âm thanh đầu vào từ miền thời gian thành miền thời gian-tần số để thu một hoặc nhiều tín hiệu đối tượng âm thanh được biến đổi. Trong phương án của Fig.2b, bộ tạo thông tin phụ tham số 93 được tạo cấu hình để tạo ra thông tin phụ tham số thứ nhất và thông tin phụ tham số thứ hai dựa trên một hoặc nhiều tín hiệu đối tượng âm thanh được biến đổi.

Fig.2c minh họa tín hiệu âm thanh được mã hóa theo phương án của sáng chế. Tín hiệu âm thanh được mã hóa bao gồm phần trộn giảm 51, biểu thị sự trộn giảm của một hoặc nhiều tín hiệu đối tượng âm thanh đầu vào, và phần thông tin phụ tham số 52 bao gồm thông tin phụ tham số thứ nhất trên ít nhất một tín hiệu đối tượng âm thanh và thông tin phụ tham số thứ hai trên ít nhất một tín hiệu đối tượng âm thanh. Độ phân giải tần số của thông tin phụ tham số thứ hai cao hơn độ phân giải tần số của thông tin phụ tham số thứ nhất.

Fig.3 minh họa hệ thống theo phương án của sáng chế. Hệ thống bao gồm bộ mã hóa 61 như được mô tả ở trên và bộ giải mã 62 như được mô tả ở trên.

Bộ mã hóa 61 được tạo cấu hình để mã hóa một hoặc nhiều tín hiệu đối tượng âm thanh đầu vào bằng cách thu một hoặc nhiều tín hiệu trộn giảm biểu thị sự trộn giảm của một hoặc nhiều tín hiệu đối tượng âm thanh đầu vào, bằng cách thu thông tin phụ tham số thứ nhất trên ít nhất một tín hiệu đối tượng âm thanh, và bằng cách thu thông tin phụ tham số thứ hai trên ít nhất một tín hiệu đối tượng âm thanh, trong đó độ phân giải tần số của thông tin phụ tham số thứ hai cao hơn độ phân giải tần số của thông tin phụ tham số thứ nhất.

Bộ giải mã 62 được tạo cấu hình để tạo ra tín hiệu âm thanh giải trộn dựa trên một hoặc nhiều tín hiệu trộn giảm, và dựa trên thông tin phụ tham số thứ nhất và thông tin phụ tham số thứ hai.

Dưới đây, SAOC được nâng cao sử dụng sự cải thiện độ phân giải tương thích ngược được mô tả.

Fig.7 minh họa sự biểu diễn tương thích ngược theo phương án của sáng chế. Các tính chất tín hiệu được biểu diễn, ví dụ đường bao phổ công suất 71, thay đổi trên tần số. Trục tần số được phân chia thành các dải tham số, và tập hợp đơn của các bộ mô tả tín hiệu được chuyển cho mỗi dải con. Sử dụng chúng thay cho việc phân phối sự mô tả cho mỗi ngăn tần số riêng cho phép lưu trong lượng thông tin phụ được yêu cầu mà không mất đáng kể chất lượng cảm giác. Trong SAOC tiêu chuẩn, bộ mô tả đơn cho mỗi dải là giá trị trung bình 72, 73, 74 của các bộ mô tả theo ngăn. Có thể được hiểu rằng, có thể đưa sự mất thông tin mà độ lớn của nó phụ thuộc vào các đặc điểm tín hiệu. Trên Fig.7, các dải $k - 1$ và k có lỗi khá lớn, trong khi dải $k + 1$ lỗi rất nhỏ.

Fig.8 minh họa đường cong chênh lệch 81 giữa giá trị tham số đúng và giá trị trung bình độ phân giải thấp theo phương án, ví dụ, thông tin cấu trúc mịn mất trong phép tham số hóa SAOC tiêu chuẩn. Sáng chế mô tả phương pháp để tham số hóa và truyền đường cong chênh lệch 81 giữa các giá trị trung bình 72, 73, 74 (ví dụ, bộ mô tả SAOC tiêu chuẩn) và các giá trị độ phân giải mịn, đúng theo cách hiệu quả cho phép tính gần đúng cấu trúc độ phân giải mịn trong bộ giải mã.

Cần lưu ý rằng việc bổ sung thông tin nâng cao cho đối tượng đơn trong hỗn hợp không chỉ cải thiện chất lượng của các đối tượng này, mà cả chất lượng của tất cả các

đối tượng chia sẻ vị trí không gian gần đúng và có một số chồng lấp phổ được cải thiện.

Sau đây, mã hóa SAOC được nâng cao tương thích ngược với bộ mã hóa được nâng cao được mô tả, cụ thể, bộ mã hóa SAOC được nâng cao với việc đưa ra dòng bit chứa phần thông tin phụ tương thích ngược và các dòng bit nâng cao bổ sung. Thông tin được bổ sung có thể được chèn vào trong dòng bit SAOC tiêu chuẩn mà theo cách cũ, bộ giải mã phù hợp tiêu chuẩn chỉ đơn giản là bỏ qua dữ liệu bổ sung trong khi bộ giải mã được nâng cao tạo việc sử dụng chúng. Các bộ giải mã SAOC tiêu chuẩn hiện có có thể giải mã phần tương thích ngược của thông tin phụ tham số (parametric side information - PSI) và đưa ra các phép khôi phục các đối tượng, trong khi thông tin bổ sung được sử dụng bởi bộ giải mã SAOC được nâng cao cải thiện chất lượng cảm giác của các khôi phục trong hầu hết các trường hợp. Ngoài ra, nếu bộ giải mã SAOC được nâng cao đang chạy trên các nguồn bị giới hạn, việc nâng cao có thể được bỏ qua và chất lượng khôi phục cơ bản vẫn thu được. Cần lưu ý rằng các sự khôi phục từ bộ giải mã SAOC tiêu chuẩn và bộ giải mã SAOC được nâng cao chỉ sử dụng SAOC tiêu chuẩn tương thích khác biệt PSI, nhưng được đánh giá là rất giống nhau về mặt cảm giác (sự chênh lệch là bản chất tương tự như trong giải mã các dòng bit SAOC tiêu chuẩn với bộ giải mã SAOC được nâng cao).

Fig.9 là hình vẽ minh họa ở mức cao bộ mã hóa được nâng cao cung cấp dòng bit tương thích ngược với sự nâng cao theo phương án của sáng chế.

Bộ mã hóa bao gồm bộ phận trộn giảm 91 trộn giảm nhiều tín hiệu đối tượng âm thanh để thu một hoặc nhiều tín hiệu trộn giảm. Ví dụ, các tín hiệu đối tượng âm thanh (ví dụ, các đối tượng (âm thanh) riêng lẻ) được sử dụng bởi bộ phận trộn giảm 91 để tạo ra tín hiệu trộn giảm. Điều đó có thể xảy ra trong miền thời gian, miền tần số, hoặc thậm chí trộn giảm được cung cấp bên ngoài có thể được sử dụng.

Trong đường PSI, các tín hiệu đối tượng (âm thanh) được biến đổi bởi bộ phận biến đổi 92 từ miền thời gian thành miền tần số, miền thời gian-tần số hoặc miền phổ (ví dụ, bởi bộ phận biến đổi 92 bao gồm một hoặc nhiều bộ phận con biến đổi t/f 921, 922).

Ngoài ra, bộ mã hóa bao gồm bộ tạo thông tin phụ tham số 93 để tạo ra thông tin phụ tham số. Trong phương án của Fig.9, bộ tạo thông tin phụ tham số 93 có thể, ví dụ, bao gồm bộ phận trích xuất PSI 94 và bộ tách PSI 95. Theo phương án đã nêu, trong miền tần số, PSI được trích xuất bởi bộ phận trích xuất PSI 94. Bộ tách PSI 95 sau đó được điều khiển để tách PSI ra thành hai phần: phần độ phân giải tần số tiêu chuẩn mà được giải mã với bộ giải mã SAOC tương thích tiêu chuẩn bất kỳ, và phần độ phân giải tần số được nâng cao. Phần độ phân giải tần số được nâng cao có thể được "ẩn" trong các yếu tố phân tử dòng bit, sao cho phần độ phân giải tần số được nâng cao là sẽ được bỏ qua bởi các bộ giải mã tiêu chuẩn nhưng được sử dụng bởi các bộ giải mã được nâng cao.

Fig.10 minh họa sơ đồ khối của bộ mã hóa theo phương án cụ thể thực hiện đường tham số của bộ mã hóa, được mô tả ở trên. Các khối chức năng màu đen đậm (102, 105, 106, 107, 108, 109) biểu thị các thành phần chính của xử lý theo sáng chế. Cụ thể, Fig.10 minh họa sơ đồ khối của việc mã hóa hai giai đoạn đưa ra dòng bit tương thích ngược với việc nâng cao cho bộ giải mã có nhiều khả năng. Bộ mã hóa được tạo cấu hình để đưa ra PSI mà có thể được giải mã với cả hai bản bộ giải mã. Bộ phận biến đổi của Fig.9 được thực hiện bởi bộ phận phát hiện nhất thời 101, bởi bộ phận tạo chuỗi cửa sổ 102, và bởi bộ phận phân tích t/f 103 trên Fig.10. Các bộ phận khác 104, 105, 106, 107, 108, 109 trên Fig.10 thực hiện bộ tạo thông tin phụ tham số 93 (ví dụ, các bộ phận 104, 105, 106, 107, 108, 109 có thể thực hiện chức năng của tổ hợp bộ phận trích xuất PSI và bộ tách PSI 95).

Thứ nhất, tín hiệu được chia nhỏ thành các khung phân tích, mà sau đó được biến đổi thành miền tần số. Nhiều khung phân tích được nhóm lại thành khung tham số chiều dài được cố định, ví dụ, trong SAOC tiêu chuẩn, chiều dài các khung phân tích 16 và 32 là phổ biến. Được giả định rằng các đặc điểm tín hiệu còn lại gần như - ổn định trong khung tham số và do đó có thể chỉ được đặc trưng bởi một tập hợp tham số. Nếu các đặc điểm tín hiệu thay đổi trong khung tham số, mô hình hóa lỗi phải tiến hành, và điều đó là có lợi cho việc chia nhỏ khung tham số dài hơn thành các phần mà giả định gần như - ổn định được thỏa mãn một lần nữa. Với mục đích này, sự phát hiện tín hiệu nhất thời là cần thiết.

Trong phương án, bộ phận biến đổi 92 được tạo cấu hình để biến đổi một hoặc nhiều tín hiệu đối tượng âm thanh đầu vào từ miền thời gian thành miền thời gian-tần số phụ thuộc vào chiều dài cửa sổ của khối biến đổi tín hiệu bao gồm các giá trị tín hiệu của ít nhất một trong số một hoặc nhiều tín hiệu đối tượng âm thanh đầu vào. Bộ phận biến đổi 92 bao gồm bộ phận phát hiện tín hiệu nhất thời 101 để xác định kết quả phát hiện tín hiệu nhất thời biểu thị có hoặc không tín hiệu nhất thời trong một hoặc nhiều trong số ít nhất một tín hiệu đối tượng âm thanh, trong đó tín hiệu nhất thời biểu thị tín hiệu thay đổi trong một hoặc nhiều trong số ít nhất một tín hiệu đối tượng âm thanh. Ngoài ra, bộ phận biến đổi 92 còn bao gồm bộ phận tạo chuỗi cửa sổ 102 để xác định chiều dài cửa sổ phụ thuộc vào kết quả phát hiện tín hiệu nhất thời.

Ví dụ, các tín hiệu nhất thời có thể được phát hiện bởi bộ phận phát hiện nhất thời 101 từ tất cả các đối tượng đầu vào riêng biệt, và khi có sự kiện nhất thời trong chỉ một đối tượng mà vị trí được công nhận như vị trí nhất thời toàn bộ. Thông tin của các vị trí nhất thời được sử dụng để xây dựng chuỗi cửa sổ thích hợp. Việc xây dựng có thể được dựa trên, ví dụ, theo logic sau:

- Thiết lập độ dài cửa sổ mặc định, tức là, chiều dài của khối biến đổi tín hiệu mặc định, ví dụ, 2048 mẫu.

- Thiết lập chiều dài khung tham số, ví dụ, 4096 mẫu, tương ứng với 4 cửa sổ mặc định với mức chồng lấp 50%. Các khung tham số nhóm nhiều cửa sổ với nhau và tập hợp đơn gồm các bộ mô tả tín hiệu được sử dụng cho khối nguyên thay vì có các bộ mô tả cho mỗi cửa sổ riêng. Điều đó cho phép giảm lượng PSI.

- Nếu không có tín hiệu nhất thời được phát hiện, sử dụng các cửa sổ mặc định và chiều dài khung tham số đầy đủ.

- Nếu tín hiệu nhất thời được phát hiện, thích ứng tạo cửa sổ để cung cấp độ phân giải theo thời gian tốt nhất tại vị trí của tín hiệu nhất thời.

Bộ phận tạo chuỗi cửa sổ 102 xây dựng chuỗi cửa sổ. Đồng thời, cũng tạo ra các khung con tham số từ một hoặc nhiều cửa sổ phân tích. Mỗi tập hợp con được phân tích như thực thể và chỉ một tập hợp gồm các tham số PSI được truyền cho mỗi khối con. Để cung cấp SAOC tiêu chuẩn tương thích PSI, chiều dài khối tham số được định

rõ được sử dụng như chiều dài khối tham số chính, và các tín hiệu nhất thời có thể được định vị trong khối định rõ các tập hợp con tham số.

Chuỗi cửa sổ được xây dựng được đưa ra để phân tích thời gian-tần số của các tín hiệu âm thanh đầu vào được xây dựng bởi bộ phận phân tích t/f 103, và được truyền trong phần nâng cao SAOC được nâng cao của PSI.

PSI bao gồm các tập hợp chênh lệch mức đối tượng (object level differences - OLD), tương quan liên đối tượng (inter-object correlations - IOC) và thông tin của ma trận trộn giảm $\mathbf{D}$ được sử dụng để tạo tín hiệu trộn giảm từ các đối tượng riêng trong bộ mã hóa. Mỗi tập hợp tham số được liên kết với biên tham số mà định rõ khoảng theo thời gian mà tham số được liên kết với.

Dữ liệu phổ của mỗi cửa sổ phân tích được sử dụng bởi bộ phận ước lượng PSI 104 để ước lượng PSI cho phần SAOC tiêu chuẩn. Điều đó được thực hiện bằng cách nhóm các ngăn phổ thành dải tham số của SAOC tiêu chuẩn và ước lượng các IOC, OLD và năng lượng đối tượng tuyệt đối (absolute objects energies-NRG) trong các dải. Sau đây kí hiệu của SAOC tiêu chuẩn, tích số chuẩn hóa của hai phổ đối tượng $\mathbf{S}_i(f, n)$ và $\mathbf{S}_j(f, n)$ trong ô tham số được định rõ

$$nr g_{i,j}(b) = \frac{\sum_{n=0}^{N-1} \sum_{f=0}^{F_n-1} K(b, f, n) \mathbf{S}_i(f, n) \mathbf{S}_j^*(f, n)}{\sum_{n=0}^{N-1} \sum_{f=0}^{F_n-1} K(b, f, n)},$$

trong đó ma trận $\mathbf{K}(b, f, n): \mathbb{R}^{B \times F_n \times N} \rightarrow \mathbb{R}^{B \times F_n \times N}$ định rõ ánh xạ từ các ngăn biểu diễn F_n t/f trong khung n thành các dải tham số B bởi

$$\mathbf{K}(b, f, n) = \begin{cases} 1, & \text{nếu } f \in b \\ 0, & \text{trái lại} \end{cases}$$

Độ phân giải phổ có thể thay đổi giữa các khung trong khối tham số đơn, vì ma trận ánh xạ chuyển đổi dữ liệu thành cơ sở phân giải thông thường. Năng lượng đối tượng cực đại trong ô tham số hóa được định rõ để là năng lượng đối tượng cực đại $NRG(b) = \max_i(nr g_{i,i}(b))$. Có giá trị này, các OLD sau đó được định rõ là năng lượng đối tượng chuẩn hóa

$$OLD_i(b) = \frac{nr g_{i,i}(b)}{NRG(b)}.$$

Và cuối cùng IOC có thể thu được từ công suất chéo theo

$$IOC_{i,j}(b) = \text{Re} \left\{ \frac{nrg_{i,j}(b)}{\sqrt{nrg_{i,i}(b)nrg_{j,j}(b)}} \right\}.$$

Điều này kết luận sự ước lượng của SAOC tiêu chuẩn tương thích các phần của dòng bit.

Bộ phận khôi phục phổ công suất thô 105 được tạo cấu hình để sử dụng các OLD và NRG để khôi phục ước lượng thô của đường bao phổ trong khối phân tích tham số. Đường bao được xây dựng trong độ phân giải tần số cao nhất được sử dụng trong khối đó.

Phổ ban đầu của mỗi cửa sổ phân tích được sử dụng bởi bộ phận ước lượng phổ công suất 106 để tính toán phổ công suất trong cửa sổ đó.

Phổ công suất thu được được biến đổi thành phép biểu diễn độ phân giải tần số cao thông thường bởi bộ phận thích ứng độ phân giải tần số 107. Điều đó có thể thực hiện, ví dụ, bởi việc nội suy các giá trị phổ công suất. Sau đó biên dạng phổ công suất trung bình được tính toán bằng cách lấy trung bình phổ trong khối tham số. Điều này gần như tương ứng với sự ước lượng OLD bỏ qua sự gộp dải tham số. Biên dạng phổ thu được được xem xét như độ phân giải mịn OLD.

Bộ mã hóa còn bao gồm bộ phận ước lượng delta 108 để ước lượng nhiều thừa số hiệu chỉnh bằng cách chia mỗi OLD trong số các OLD của một trong số ít nhất một tín hiệu đối tượng âm thanh bằng giá trị của khôi phục phổ công suất của một tín hiệu đối tượng âm thanh đã nêu trong số ít nhất một tín hiệu đối tượng âm thanh để thu được thông tin phụ tham số thứ hai, trong đó các OLD đã nêu có độ phân giải tần số cao hơn sự khôi phục phổ công suất đã nêu.

Theo phương án, bộ phận ước lượng delta 108 được tạo cấu hình để ước lượng nhiều thừa số hiệu chỉnh dựa trên nhiều giá trị tham số phụ thuộc vào ít nhất một tín hiệu đối tượng âm thanh để thu được thông tin phụ tham số thứ hai. Ví dụ, bộ phận ước lượng delta 108 có thể được tạo cấu hình để ước lượng thừa số hiệu chỉnh, "delta", ví dụ, bằng cách chia độ phân giải mịn OLD bởi sự khôi phục phổ công suất thô. Kết

quả là, điều này cung cấp cho mỗi ngân tần số (ví dụ, nhân lên) thừa số hiệu chỉnh mà có thể được sử dụng để tính gần đúng độ phân giải mịn OLD đã đưa ra cho phổ thô.

Cuối cùng, bộ phận mô hình hóa delta 109 được tạo cấu hình để tạo mô hình cho thừa số hiệu chỉnh được ước lượng theo cách hiệu quả cho việc truyền. Một khả năng để mô hình hóa bằng cách sử dụng hệ số dự đoán tuyến tính (Linear Prediction Coefficients - LPC) được mô tả sau đây.

Một cách hiệu quả, các thay đổi SAOC được nâng cao bao gồm việc bổ sung thông tin chuỗi cửa sổ và tham số để truyền "delta" đến dòng bit.

Sau đây, bộ giải mã được nâng cao được mô tả.

Fig.11 minh họa sơ đồ khối ở mức cao của bộ giải mã được nâng cao theo phương án của sáng chế mà có thể giải mã cả dòng bit tiêu chuẩn và dòng bit được nâng cao. Cụ thể, Fig.11 minh họa sơ đồ khối hoạt động của bộ giải mã được nâng cao có thể giải mã cả các dòng bit tiêu chuẩn cũng như dòng bit chứa các nâng cao độ phân giải tần số.

Tín hiệu trộn giảm đầu vào được biến đổi thành miền tần số bởi bộ phận biến đổi t/f 111.

Ma trận giải trộn được ước lượng được áp dụng trên tín hiệu trộn giảm được biến đổi bởi bộ phận giải trộn 110 để tạo ra đầu ra giải trộn.

Ngoài ra, đường khử tương quan được chứa để cho phép điều khiển trong không gian tốt hơn của các đối tượng trong việc giải trộn. Bộ phận khử tương quan 119 xây dựng sự khử tương quan trên tín hiệu trộn giảm được biến đổi và kết quả của sự khử tương quan được nạp vào trong bộ phận giải trộn 110. Bộ phận giải trộn 110 sử dụng kết quả khử tương quan để tạo ra đầu ra giải trộn.

Đầu ra giải trộn sau đó được biến đổi sau thành miền thời gian bằng bộ phận biến đổi f/t 114.

Đường xử lý tham số có thể nhận độ phân giải tiêu chuẩn PSI như đầu vào, trong trường hợp PSI được giải mã, mà được tạo ra bởi bộ phận giải mã PSI tiêu chuẩn 115, được làm thích ứng bởi bộ phận chuyển đổi độ phân giải tần số 116 thành độ phân giải tần số được sử dụng trong các biến đổi t/f .

Đầu vào thay thế kết hợp phần độ phân giải tần số của PSI với phần độ phân giải tần số được nâng cao và các phép tính toán gồm thông tin độ phân giải tần số được nâng cao. Cụ thể hơn, bộ phận giải mã PSI được nâng cao 117 tạo ra PSI được giải mã biểu thị độ phân giải tần số được nâng cao.

Bộ tạo ma trận giải trộn 118 tạo ra ma trận giải trộn dựa trên PSI được giải mã nhận được từ bộ phận chuyển đổi độ phân giải tần số 116 hoặc từ bộ phận giải mã PSI được nâng cao 117. Bộ tạo ma trận giải trộn 118 có thể cũng tạo ra ma trận giải trộn dựa trên thông tin kết xuất, ví dụ, dựa trên ma trận kết xuất. Bộ phận giải trộn 110 được tạo cấu hình để tạo ra đầu ra giải trộn bằng cách ứng dụng ma trận giải trộn, được tạo ra bởi bộ tạo ma trận giải trộn 118 trên tín hiệu trộn giảm được biến đổi.

Fig.12 minh họa sơ đồ khối minh họa phương án của bộ phận giải mã PSI được nâng cao 117 theo Fig.11.

Thông tin tham số thứ nhất bao gồm nhiều giá trị tham số thứ nhất, trong đó thông tin tham số thứ hai bao gồm nhiều giá trị tham số thứ hai. Bộ xác định thông tin giải trộn 112 bao gồm bộ con chuyển đổi độ phân giải tần số 122 và bộ kết hợp 124. Bộ phận chuyển đổi độ phân giải tần số 112 được tạo cấu hình để tạo ra các giá trị tham số bổ sung, ví dụ, bằng cách tái tạo các giá trị tham số thứ nhất, trong đó các giá trị tham số thứ nhất và các giá trị tham số bổ sung cùng nhau tạo nhiều giá trị tham số được xử lý thứ nhất. Bộ kết hợp 124 được tạo cấu hình để kết hợp các giá trị tham số được xử lý thứ nhất và các giá trị tham số thứ hai để thu các giá trị tham số được thay đổi như thông tin tham số được thay đổi.

Theo phương án, phần độ phân giải tần số tiêu chuẩn được giải mã bởi bộ phận con giải mã 121 và được chuyển đổi bởi bộ phận con chuyển đổi độ phân giải tần số 122 thành độ phân giải tần số được sử dụng bởi phần nâng cao. Phần nâng cao được giải mã, được tạo ra bởi bộ phận con giải mã PSI được nâng cao 123 được kết hợp bởi bộ kết hợp 124 với phần độ phân giải tiêu chuẩn được chuyển đổi.

Sau đây, hai mô hình giải mã có thể thực hiện được mô tả cụ thể.

Thứ nhất, việc giải mã các dòng bit SAOC tiêu chuẩn với bộ giải mã được nâng cao được mô tả.

Bộ giải mã SAOC được nâng cao được thiết kế để có thể giải mã các dòng bit từ các bộ mã hóa SAOC tiêu chuẩn với chất lượng tốt. Việc giải mã chỉ bị giới hạn sự khôi phục tham số, và các dòng dư có thể có được bỏ qua.

Fig.13 vẽ sơ đồ khối của việc giải mã dòng bit SAOC tiêu chuẩn với bộ giải mã SAOC được nâng cao minh họa xử lý giải mã theo phương án của sáng chế. Các khối chức năng màu đen đậm (131, 132, 133, 135) biểu thị các phần chính của xử lý theo sáng chế.

Bộ tính toán ma trận giải trộn 131, bộ nội suy theo thời gian 132, và bộ phận làm thích ứng độ phân giải tần số-cửa sổ 131 thực hiện chức năng của bộ phận giải mã PSI tiêu chuẩn 115, của bộ phận chuyển đổi độ phân giải tần số 116 và của bộ tạo ma trận giải trộn 118 của Fig.11. Bộ tạo chuỗi cửa sổ 134 và môđun phân tích t/f 135 thực hiện bộ phận biến đổi t/f 111 của Fig.11.

Thông thường, các ngăn tần số của phép biểu diễn thời gian/tần số được nhóm lại thành các dải tham số. Việc giãn cách của các dải tương tự của dải tới hạn trong hệ thống thính giác con người. Hơn nữa, nhiều khung biểu diễn t/f có thể được nhóm thành khung tham số. Cả hai toán tử này cung cấp sự giảm bớt lượng thông tin phụ yêu cầu với giá trị mô hình hóa không chính xác.

Như mô tả trong tiêu chuẩn SAOC, các OLD và IOC được sử dụng để tính toán ma trận giải trộn $\mathbf{G} = \mathbf{E}\mathbf{D}^*\mathbf{J}$, tại các yếu tố phân tử của $\mathbf{E}$ được định rõ như $E(i, j) = IOC_{i,j} \sqrt{OLD_i OLD_j}$ tính gần đúng ma trận tương quan chéo đối tượng, i và j là các chỉ số đối tượng, $\mathbf{J} \approx (\mathbf{D}\mathbf{E}\mathbf{D}^*)^{-1}$. Bộ tính toán ma trận giải trộn 131 có thể điều khiển để tính toán ma trận giải trộn.

Ma trận giải trộn sau đó được nội suy tuyến tính bởi bộ nội suy theo thời gian 132 từ ma trận giải trộn của khung dự báo trên khung tham số đến biên tham số trên đó các giá trị ước lượng đạt được, như mỗi SAOC tiêu chuẩn. Kết quả là thành các ma trận giải trộn cho mỗi cửa sổ phân tích thời gian/tần số và dải tham số.

Độ phân giải tần số dải tham số của các ma trận giải trộn được mở rộng thành độ phân giải của biểu diễn thời gian/tần số trong đó phân tích cửa sổ bằng bộ phận làm thích ứng độ phân giải cửa sổ-tần số 133. Khi ma trận giải trộn nội suy cho dải tham số

b trong khung thời gian được định rõ như $\mathbf{G}(b)$, các hệ số giải trộn tương tự được sử dụng cho tất cả các ngăn tần số bên trong dải tham số.

Bộ tạo chuỗi cửa sổ 134 được tạo cấu hình để sử dụng thông tin khoảng tập hợp tham số từ PSI để xác định chuỗi cửa sổ gần đúng để phân tích tín hiệu âm thanh trộn giảm đầu vào. Yêu cầu chính là khi có biên tập hợp tham số trong PSI, điểm giao nhau giữa các cửa sổ phân tích liên tiếp phải khớp với nó. Việc tạo cửa sổ cũng xác định độ phân giải tần số của dữ liệu trong mỗi cửa sổ (sử dụng trong mở rộng dữ liệu giải trộn, vì được mô tả ở trên).

Dữ liệu được tạo cửa sổ sau đó được biến đổi bằng môđun phân tích t/f 135 thành sự biểu diễn miền tần số sử dụng biến đổi thời gian-tần số thích hợp, ví dụ, biến đổi fourier rời rạc (discrete fourier transform - DFT), biến đổi cosin rời rạc được thay đổi phức hợp (complex modified discrete cosine transform - CMDCT) hoặc biến đổi fourier rời rạc được chồng lè (oddly stacked discrete fourier transform - ODFT).

Cuối cùng, bộ phận giải trộn 136 áp dụng các ma trận giải trộn mỗi ngăn tần số mỗi khung trên sự biểu diễn phổ của tín hiệu trộn giảm $\mathbf{X}$ để thu kết xuất tham số $\mathbf{Y}$. Kênh đầu ra j là kết hợp tuyến tính của các kênh trộn giảm $\mathbf{Y}_j = \sum_i \mathbf{G}_{j,i} \mathbf{X}_i$.

Chất lượng mà có thể thu được với việc xử lý này là dành cho hầu hết các mục đích đồng nhất về cảm giác từ kết quả thu được với bộ giải mã SAOC tiêu chuẩn.

Cần lưu ý rằng phần trên đã mô tả việc khôi phục các đối tượng riêng lẻ, nhưng trong SAOC tiêu chuẩn sự kết xuất được chứa trong ma trận giải trộn, cụ thể, nó được chứa trong phép nội suy tham số. Như phép toán tuyến tính, thứ tự của các phép toán không quan trọng, nhưng sự chênh lệch là đáng lưu ý.

Sau đây, việc giải mã các dòng bit SAOC nâng cao với bộ giải mã được nâng cao được mô tả.

Chức năng chính của bộ giải mã SAOC được nâng cao đã được mô tả ở trên trong giải mã dòng bit SAOC tiêu chuẩn. Phần này sẽ mô tả chi tiết về cách nâng cao SAOC được nâng cao được đưa ra trong PSI có thể được sử dụng để thu được chất lượng cảm giác tốt hơn.

Fig.14 minh họa các khối chức năng của bộ giải mã theo phương án minh họa việc mã hóa của các sự nâng cao độ phân giải tần số. Các khối chức năng màu đen đậm (141, 142, 143) biểu thị các phần chính của xử lý theo sáng chế. Bộ phận giá trị mở rộng qua dải 141, bộ phận khôi phục hàm delta 142, bộ phận áp dụng delta 143, bộ tính toán ma trận giải trộn 131, bộ nội suy theo thời gian 132, và bộ phận thích ứng độ phân giải tần số của số 133 thực hiện chức năng của bộ phận giải mã PSI được nâng cao 117 và của bộ tạo ma trận giải trộn 118 theo Fig.11.

Bộ giải mã theo Fig.14 bao gồm bộ xác định thông tin giải trộn 112. Ngoài ra, bộ xác định thông tin giải trộn 112 bao gồm bộ phận khôi phục hàm delta 142 và bộ phận áp dụng delta 143. Thông tin tham số thứ nhất bao gồm nhiều giá trị tham số phụ thuộc vào ít nhất một tín hiệu đối tượng âm thanh, ví dụ, các giá trị chênh lệch mức đối tượng. Thông tin tham số thứ hai bao gồm sự tham số hóa thừa số hiệu chỉnh. Bộ phận khôi phục hàm delta 142 được tạo cấu hình để đảo ngược sự tham số hóa thừa số hiệu chỉnh để thu được hàm delta. Bộ phận áp dụng delta 143 được tạo cấu hình để áp dụng hàm delta trên các giá trị tham số, ví dụ, trên các giá trị chênh lệch mức đối tượng, để xác định thông tin giải trộn. Trong phương án, sự tham số hóa thừa số hiệu chỉnh bao gồm nhiều hệ số dự báo tuyến tính và bộ phận khôi phục hàm delta 142 được tạo cấu hình để đảo ngược sự tham số hóa thừa số hiệu chỉnh bằng cách tạo ra nhiều thừa số hiệu chỉnh phụ thuộc vào nhiều hệ số dự báo tuyến tính, và được tạo cấu hình để tạo ra hàm delta dựa vào nhiều thừa số hiệu chỉnh.

Ví dụ, thứ nhất, bộ phận giá trị mở rộng qua dải 141, thích ứng giá trị OLD và IOC cho mỗi dải tham số đến độ phân giải tần số được sử dụng trong nâng cao, ví dụ, đến gần 1024. Điều đó được thực hiện bằng cách tái tạo giá trị qua các ngăn tần số mà tương ứng với dải tham số. Điều đó thu được OLDs mới $OLD_i^{enh}(f) = K(f, b)OLD_i(b)$ và IOC mới $IOC_{i,j}^{enh}(f) = K(f, b)IOC_{i,j}(b)$. $K(f, b)$ là ma trận kernel định rõ sự chuyển các ngăn tần số f thành các dải tham số b .

Song song với đó, bộ phận khôi phục hàm delta 142 đảo ngược sự tham số hóa thừa số hiệu chỉnh để thu được hàm delta $\mathcal{C}_i^{res}(f)$ cùng kích cỡ như OLD và IOC được mở rộng.

Sau đó, bộ phận áp dụng delta 143 áp dụng delta trên các giá trị OLD được mở rộng, và các giá trị OLD độ phân giải mịn thu được được thu bởi $OLD_i^{fine}(f) = C_i^{rec}(f)OLD_i^{enh}(f)$.

Trong phương án cụ thể, việc tính toán các ma trận giải trộn, có thể, ví dụ, được thực hiện bởi bộ tính toán ma trận giải trộn 131 như với việc giải mã dòng bit SAOC tiêu chuẩn:

$$G(f) = E(f)D^*(f)J(f), \quad \text{với}$$

$$E_{i,j}(f) = 10C_{i,j}^{enh}(f)\sqrt{OLD_i^{fine}(f)OLD_j^{fine}(f)}, \text{ và } J(f) \approx (D(f)E(f)D^*(f))^{-1}. \text{ Nếu}$$

muốn, ma trận kết xuất có thể được nhân lên thành ma trận giải trộn $G(f)$. Phép nội suy theo thời gian bằng bộ nội suy theo thời gian 132 sau mỗi SAOC tiêu chuẩn.

Vì độ phân giải tần số trong mỗi cửa sổ có thể khác với (thấp hơn) độ phân giải tần số cao danh định, bộ phận thích ứng độ phân giải tần số cửa sổ 133 cần để thích ứng các ma trận giải trộn khớp độ phân giải của dữ liệu phổ từ âm thanh để cho phép ứng dụng nó. Điều đó có thể được thực hiện, ví dụ, bằng việc lấy mẫu lại các hệ số trên trục tần số đến độ phân giải chính xác. Hoặc nếu các độ phân dải là bội số nguyên, thì lấy trung bình đơn giản từ dữ liệu độ phân giải cao các chỉ số mà tương ứng với một ngăn tần số trong độ phân giải thấp hơn $G^{low}(b) = \mathbf{1}/\|b\| \sum_{f \in b} G(f)$.

Thông tin chuỗi tạo cửa sổ từ dòng bit có thể được sử dụng để thu được phân tích thời gian-tần số bù đầy đủ cho một chuỗi được sử dụng trong bộ mã hóa, hoặc chuỗi tạo cửa sổ có thể được xây dựng dựa trên các biên tham số, như được thực hiện trong giải mã dòng bit SAOC tiêu chuẩn. Với điều này bộ tạo chuỗi cửa sổ 134 có thể được sử dụng.

Phân tích thời gian-tần số của âm thanh trộn giảm sau đó được xây dựng bằng môđun phân tích t/f 135 sử dụng các cửa sổ đã có.

Cuối cùng, các ma trận giải trộn (có thể) được thích ứng theo phổ và nội suy theo thời gian được ứng dụng bởi bộ phận giải trộn 136 trên sự biểu diễn thời gian-tần số của âm thanh đầu vào, và kênh đầu ra j có thể thu được như kết hợp tuyến tính của các kênh đầu vào $Y_j(f) = \sum_i G_{j,i}^{low}(f)X_i(f)$.

Sau đây, các khía cạnh cụ thể của các phương án có thể được mô tả.

Theo phương án, bộ phận mô hình hóa delta 109 của Fig.10 được tạo cấu hình để xác định các hệ số dự báo tuyến tính từ các thừa số hiệu chỉnh (delta) bằng cách xây dựng phép dự báo tuyến tính.

Sau đây, xử lý ước lượng thừa số hiệu chỉnh, delta, và thay thế mô hình hóa cso thể có sử dụng các hệ số dự báo tuyến tính (linear prediction coefficients - LPC) theo phương án đã nêu được mô tả.

Trước hết, phép ước lượng delta theo phương án được mô tả.

Đầu vào vào phép ước lượng bao gồm biên dạng phổ công suất độ phân giải mịn được ước lượng qua khối tham số và từ phép khôi phục thô biên dạng phổ công suất dựa trên các tham số NRG và OLD. Các biên dạng phổ công suất mịn được tính toán theo cách sau. $S_i(f, n)$ là phổ hỗn hợp của đối tượng thứ i với f là chỉ số ngăn tần số và $0 \leq n \leq N - 1$ là chỉ số cửa sổ theo thời gian trong khối mô hình hóa chiều dài N . Phổ công suất độ phân giải mịn sau đó là

$$P_i(f) = \frac{1}{N} \sum_{n=0}^{N-1} S_i(f, n) S_i^*(f, n).$$

Sự khôi phục thô được tính toán từ các OLD và NRGs (khử lượng tử hóa) bởi

$$Z_i(f) = K(f, b) OLD_i(b) NRG_i(b),$$

tại đó $K(f, b)$ là ma trận kernel định rõ sự chuyển các ngăn tần số f thành các dải tham số b .

Hai tín hiệu với các đặc điểm phổ khác biệt sẽ được sử dụng như các ví dụ trong phần này: một tín hiệu thứ nhất là nhiễu âm (hồng) với phổ phẳng thực tế (bỏ qua độ nghiêng phổ) và tín hiệu thứ hai là âm từ nhạc cụ đàn chuông mà có âm cao, cụ thể, đỉnh, phổ.

Fig.15 minh họa phổ công suất của tín hiệu âm và nhiễu âm. Phổ công suất độ phân giải cao của chúng ("orig") và sự khôi phục thô tương ứng dựa trên các OLD và NRG ("recon"). Cụ thể, Fig.15 minh họa phổ công suất thô và mịn của cả hai tín hiệu. Cụ thể hơn, phổ công suất của tín hiệu âm ban đầu 15 và tín hiệu nhiễu âm ban đầu 152, và phổ công suất được khôi phục của tín hiệu âm 153 và tín hiệu nhiễu âm 154 được thể hiện. Cần lưu ý rằng, trong các hình vẽ sau đây, các tín hiệu 153 và 154 đúng

hơn là các thừa số tỷ lệ (tham số phổ công suất được khôi phục) và không phải là các tín hiệu được khôi phục đầy đủ được phác họa.

Có thể được ghi chú nhanh, độ chênh lệch trung bình giữa giá trị thô và mịn đúng hơn là nhỏ trong trường hợp của tín hiệu nhiễu âm, trong khi độ chênh lệch là rất lớn trong tín hiệu âm. Những độ chênh lệch này gây ra suy giảm cảm giác trong sự khôi phục tham số của tất cả các đối tượng.

Thừa số điều chỉnh thu được bằng cách chia đường cong độ phân giải mịn bằng đường cong độ phân giải thô.

$$C_i(f) = P_i(f)/Z_i(f).$$

Điều đó cho phép khôi phục thừa số bội mà có thể được ứng dụng trên sự khôi phục thô để thu được đường cong độ phân giải mịn:

$$P_i^{rec}(f) = Z_i(f)C_i(f).$$

Fig.16 minh họa cả hai tín hiệu ví dụ, cụ thể, minh họa các thừa số hiệu chỉnh cho các tín hiệu ví dụ. Cụ thể, các thừa số hiệu chỉnh cho tín hiệu âm 151 và tín hiệu nhiễu âm 152 được thể hiện.

Dưới đây, mô hình hóa delta được mô tả.

Đường cong hiệu chỉnh C được chuyển thành một hoặc nhiều khối mô hình hóa qua trực tần số. Thay thế tự nhiên là sử dụng các định nghĩa dải tham số tương tự như được sử dụng cho PSI SAOC tiêu chuẩn. Việc mô hình hóa sau đó được thực hiện cho mỗi khối riêng với các bước sau:

1. Thừa số hiệu chỉnh phổ C được biến đổi thành chuỗi tự tương quan miền thời gian với Biến đổi Fourier Rời rạc Nghịch đảo (Inverse Discrete Fourier Transform - IDFT).

Khi chiều dài của khối mô hình hóa lẻ, phổ giả sẽ được biến đổi được định rõ như sau

$$R(l) = \begin{cases} C(l), & \text{khi } 0 \leq l \leq N-1 \\ C(2N-2-l), & \text{khi } N \leq l \leq 2N-3 \end{cases}.$$

Khi khối mô hình hóa chẵn, phổ giả được định rõ như sau

$$R(l) = \begin{cases} C(l), & \text{khi } 0 \leq l \leq N-1 \\ C(2N-1-l), & \text{khi } N \leq l \leq 2N-2 \end{cases}$$

Kết quả biến đổi sau đó $r(t) = IDFT(R(l))$.

2. Kết quả được chắt thành nửa thứ nhất: $\begin{cases} 0 \leq t \leq N-2, & \text{khi } N \text{ lẻ} \\ 0 \leq t \leq N-1, & \text{khi } N \text{ chẵn} \end{cases}$

3. Các phép truy hồi Levinson-Durbin được ứng dụng trên chuỗi tự tương quan $r(t)$ để có được các hệ số phản xạ k và mô hình hóa các phương sai dư e để tăng bậc mô hình.

4. Tùy chọn: Dựa trên mô hình hóa phương sai dư e , bỏ qua toàn bộ mô hình (vì độ khếch đại không thu được) hoặc chọn bậc thích hợp.

5. Các tham số mô hình được lượng tử hóa để truyền.

Có thể tạo quyết định nếu delta cần được truyền cho mỗi ô t-f (dải tham số tiêu chuẩn định rõ khoảng tần số và khối tham số khoảng thời gian) một cách độc lập. Quyết định này có thể được tạo dựa trên, ví dụ:

- Kiểm tra năng lượng dư của mô hình hóa delta. Nếu năng lượng dư của mô hình hóa không vượt quá ngưỡng nhất định, thông tin nâng cao không được truyền.

- Đo "độ nhọn"/không phẳng của mô tả tham số được mô hình hóa có độ phân giải mịn, mô hình hóa delta, hoặc đường bao phổ công suất của tín hiệu đối tượng âm thanh. Phụ thuộc vào giá trị đo được các tham số mô hình hóa delta, mà mô tả độ phân giải phổ mịn, và được truyền hoặc không, hoặc được tính toán tại tất cả phụ thuộc vào độ không phẳng của đường bao phổ công suất của tín hiệu đối tượng âm thanh). Các phép đo thích hợp ví dụ là thừa số đỉnh phổ, phép đo độ phẳng phổ, hoặc tỉ lệ cực tiểu-cực đại.

- Chất lượng cảm giác của sự khôi phục thu được. Bộ mã hóa tính toán sự khôi phục kết xuất nâng cao và không nâng cao, và xác định độ khếch đại chất lượng cho mỗi nâng cao. Sau đó điểm cân bằng thích hợp giữa độ phức tạp mô hình hóa và độ khếch đại chất lượng được định vị, và các nâng cao được biểu thị được truyền. Ví dụ, sự biến dạng được gán trọng số về cảm giác cho tỷ lệ tín hiệu hoặc các phép đo cảm giác được nâng cao có thể được sử dụng cho quyết định. Quyết định có thể được tạo

cho mỗi dải tham số (thô) riêng biệt (tức là, tối ưu hóa chất lượng cục bộ), nhưng cũng dưới điều kiện xem xét các dải liên kế để tính toán cho những biến dạng tín hiệu gây ra bởi việc sử dụng thời gian và tần số khác nhau của các hệ số thời gian-tần số (cụ thể là, tối ưu hóa chất lượng toàn cục).

Sau đây, việc khôi phục delta và ứng dụng được mô tả.

Việc khôi phục của đường cong hiệu chỉnh theo các bước sau:

1. Các hệ số phản chiếu nhận được $\mathbf{k}$ (vec tơ có chiều dài $L - 1$) được khử tương quan và biến đổi thành các hệ số bộ lọc IIR $\mathbf{a}$ chiều dài L , trong cú pháp mã giả (mà hàm $X = \text{diag}(x)$ xuất ra ma trận X với các yếu tố phân tử chéo của X là x và tất cả các yếu tố phân tử không chéo của X là không):

```
A = diag(k)
for ii=1 to L
    for l=1 to ii-1
        A(l,ii) = A(l,ii-1) + k(ii)*A(ii-1,ii-1)
    end
end
a = [1; A(1 to end,end)]
```

2. Đáp ứng tần số $\mathbf{h}(n)$ của bộ lọc thu được $\mathbf{a}$ được tính toán bằng công thức $\mathbf{h}(n) = 1 / \sum_{l=0}^{L-1} \mathbf{a}(l) \exp(-in2\pi/N)$, mà i kí hiệu bộ phận ảo $i = \sqrt{-1}$.

3. Khôi phục hàm tương quan thu được từ đó bằng công thức $\mathbf{C}^{raw}(n) = \mathbf{h}(n)\mathbf{h}^*(n)$.

4. Đáp ứng được chuẩn hóa để có trung bình đơn vị, mà năng lượng tổng của khối được mô hình hóa không thay đổi $\mathbf{C}^{rec}(n) = \mathbf{C}^{raw}(n) / \sum_{n=0}^{N-1} \mathbf{C}^{raw}(n)$.

5. Các thừa số hiệu chỉnh được ứng dụng trên các OLD, mà đã được mở rộng đến độ phân giải mịn $OLD_i^{fine}(f) = \mathbf{C}_i^{rec}(f)\mathbf{K}(f,b)OLD_i(b)$. Lưu ý, năng lượng tuyệt đối có thể được bỏ qua vì chúng được triệt tiêu trong các phép toán sau này.

Fig.17 minh họa các thừa số hiệu chỉnh ban đầu và các phép tính gần đúng dựa trên LPC có bậc đã giảm (sau khi mô hình hóa) cho cả hai tín hiệu ví dụ. Cụ thể hơn, các thừa số hiệu chỉnh ban đầu của tín hiệu âm 151, tín hiệu nhiễu âm ban đầu 152, và

thừa số hiệu chỉnh được khôi phục ước lượng tín hiệu âm 153 và tín hiệu nhiễu âm 154 được thể hiện.

Fig.18 minh họa kết quả của việc ứng dụng các thừa số hiệu chỉnh được mô hình hóa trên sự khôi phục thô được minh họa trên Fig.15. Cụ thể hơn, phổ công suất của tín hiệu âm ban đầu 151 và tín hiệu nhiễu âm ban đầu 152, và phổ công suất được khôi phục của tín hiệu âm 153 và tín hiệu nhiễu âm 154 được thể hiện. Các đường cong này có thể được sử dụng thay cho các OLD trong các tính toán sau đây, cụ thể, phổ công suất độ phân giải mịn được khôi phục sau khi áp dụng các thừa số hiệu chỉnh được mô hình hóa. Ở đây, thông tin năng lượng tuyệt đối được chứa để tạo sự so sánh tốt hơn có thể thấy được, nhưng nguyên tắc tương tự cũng hoạt động mà không có chúng.

Phương pháp và thiết bị theo sáng chế làm giảm các hạn chế đã nêu ở trên thuộc tình trạng kỹ thuật của xử lý SAOC sử dụng giàn bộ lọc hoặc biến đổi thời gian-tần số với độ phân giải tần số cao và cung cấp quá trình tham số hóa hiệu quả của thông tin bổ sung. Ngoài ra, có thể truyền thông tin bổ sung này theo cách mà các bộ giải mã SAOC tiêu chuẩn có thể giải mã phần tương thích ngược của thông tin với chất lượng có thể đạt được có thể so sánh chất lượng của thông tin thu được sử dụng bộ mã hóa SAOC thích ứng tiêu chuẩn, và vẫn cho phép các bộ giải mã được nâng cao sử dụng thông tin bổ sung cho chất lượng cảm giác tốt hơn. Điều quan trọng nhất, thông tin bổ sung có thể được biểu diễn trong nhiều cách rất gọn để truyền hoặc lưu trữ hiệu quả.

Phương pháp theo sáng chế có thể áp dụng trên sơ đồ SAOC bất kỳ. Phương pháp theo sáng chế có thể được kết hợp với các định dạng âm thanh hiện tại bất kỳ cũng như tương lai. Phương pháp theo sáng chế cho phép nâng cao chất lượng âm thanh cảm giác trong các ứng dụng SAOC bằng việc biểu diễn mức hai của thông tin phụ dạng phổ.

Ý tưởng tương tự có thể được sử dụng liên hợp với MPEG lập thể khi thay thế khái niệm của các OLD với các độ chênh lệch mức kênh (channel-level differences - CLDs).

Bộ mã hóa âm thanh hoặc phương pháp mã hóa âm thanh hoặc chương trình máy tính liên quan như được mô tả được đề xuất. Ngoài ra, bộ giải mã âm thanh hoặc phương pháp giải mã âm thanh hoặc chương trình máy tính liên quan như được mô tả được đề xuất. Ngoài ra, tín hiệu âm thanh được mã hóa hoặc vật ghi lưu trữ có lưu trữ tín hiệu âm thanh được mã hóa được mô tả ở trên được đề xuất.

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

Tín hiệu đã được tách ra có thể được lưu trữ trên môi trường lưu trữ số hoặc có thể được truyền trên môi trường truyền như môi trường truyền dẫn không dây hoặc môi trường truyền dẫn 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, CD, ROM, PROM, EPROM, EEPROM hoặc bộ nhớ FLASH, có các tín hiệu điều khiển có thể đọc được theo cách điện tử được lưu trữ trên đó, mà kết hợp (hoặc có thể kết hợp) với hệ thống máy tính có thể lập trình được sao cho phương pháp tương ứng được thực hiện.

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

Thông thường, các phương án của sáng chế có thể được thực hiện như một sản phẩm chương trình máy tính với mã chương trình, mã chương trình hoạt động để thực hiện một trong các phương pháp khi sản phẩm chương trình máy tính chạy trên máy tính. Mã chương trình có thể, ví dụ, được lưu trữ trên vật mang có thể đọc được bằng máy.

Các phương án khác gồm có 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, một 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.

Phương án khác của các phương pháp theo sáng chế là, vật mang dữ liệu (hoặc vật ghi lưu trữ số, hoặc vật ghi có thể đọc được bằng máy tính) bao gồm chương trình máy tính được ghi lại trên đó để thực hiện một trong các phương pháp được mô tả trong sáng chế.

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

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

Phương án khác bao gồm máy tính chứa chương trình máy tính được cài đặt trên máy tính để thực hiện một trong các phương pháp được mô tả trong sáng chế.

Theo một số phương án, thiết bị logic lập trình được (ví dụ, mảng cổng lập trình được dạng trường) có thể được sử dụng để thực hiện một số hoặc tất cả các chức năng của các phương pháp được mô tả trong sáng chế. Theo 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 các phương pháp được mô tả trong sáng chế. 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 họa cho các nguyên lý của sáng chế. Cần hiểu rằng các biến thể và biến đổi của các phương án và các chi tiết được mô tả trong sáng chế sẽ là rõ ràng đối với người có hiểu biết trong lĩnh vực kỹ thuật tương ứng. Do đó, mục đích của sáng chế chỉ bị giới hạn bởi phạm vi của các

điểm yêu cầu bảo hộ và không bị giới hạn bởi các chi tiết cụ thể được thể hiện bởi bản mô tả và phân giải thích các phương án trong sáng chế.

YÊU CẦU BẢO HỘ

1. Bộ giải mã để tạo ra tín hiệu âm thanh giải trộn bao gồm nhiều kênh âm thanh giải trộn, trong đó bộ giải mã bao gồm:

bộ xác định thông tin giải trộn (112) để xác định thông tin giải trộn bằng cách tiếp nhận thông tin phụ tham số thứ nhất trên ít nhất một tín hiệu đối tượng âm thanh và thông tin phụ tham số thứ hai trên ít nhất một tín hiệu đối tượng âm thanh, trong đó độ phân giải tần số của thông tin phụ tham số thứ hai cao hơn độ phân giải tần số của thông tin phụ tham số thứ nhất, và

môđun giải trộn (113) để áp dụng thông tin giải trộn trên tín hiệu trộn giảm, biểu thị sự trộn giảm của ít nhất một tín hiệu đối tượng âm thanh, để thu tín hiệu âm thanh được giải trộn bao gồm nhiều kênh âm thanh được giải trộn,

trong đó bộ xác định thông tin giải trộn (112) được tạo cấu hình để xác định thông tin giải trộn bằng cách thay đổi thông tin tham số thứ nhất và thông tin tham số thứ hai để thu thông tin tham số được thay đổi, sao cho thông tin tham số được thay đổi có độ phân giải tần số cao hơn độ phân giải tần số thứ nhất.

2. Bộ giải mã theo điểm 1,

trong đó bộ giải mã còn bao gồm bộ phận biến đổi thứ nhất (111) để biến đổi đầu vào trộn giảm, sẽ được biểu diễn trong miền thời gian, để thu tín hiệu trộn giảm, được biểu diễn trong miền thời gian-tần số, và

trong đó bộ giải mã bao gồm bộ phận biến đổi thứ hai (114) để biến đổi tín hiệu âm thanh được giải trộn từ miền thời gian-tần số thành miền thời gian.

3. Bộ giải mã theo điểm 1 hoặc điểm 2, trong đó bộ xác định thông tin giải trộn (112) được tạo cấu hình để xác định thông tin giải trộn bằng cách kết hợp thông tin tham số thứ nhất và thông tin tham số thứ hai để thu thông tin tham số được thay đổi, sao cho thông tin tham số được thay đổi có độ phân giải tần số bằng độ phân giải tần số thứ hai.

4. Bộ giải mã theo một điểm bất kỳ trong số các điểm nêu trên,

trong đó, thông tin tham số thứ nhất bao gồm nhiều giá trị tham số thứ nhất, trong đó thông tin tham số thứ hai bao gồm nhiều giá trị tham số thứ hai,

trong đó bộ xác định thông tin giải trộn (112) bao gồm bộ con chuyển đổi độ phân giải tần số (122) và bộ kết hợp (124),

trong đó bộ phận chuyển đổi độ phân giải tần số (112) được tạo cấu hình để tạo ra các giá trị tham số bổ sung, trong đó các giá trị tham số thứ nhất và các giá trị tham số bổ sung cùng nhau tạo nhiều giá trị tham số được xử lý thứ nhất, và

trong đó bộ kết hợp (124) được tạo cấu hình để kết hợp các giá trị tham số được xử lý thứ nhất và các giá trị tham số thứ hai để thu nhiều giá trị tham số được thay đổi như thông tin tham số được thay đổi.

5. Bộ giải mã theo một điểm bất kỳ trong số các điểm nêu trên,

trong đó bộ xác định thông tin giải trộn (112) bao gồm bộ phận khôi phục hàm delta (142) và bộ phận áp dụng delta (143),

trong đó thông tin tham số thứ nhất bao gồm nhiều giá trị tham số phụ thuộc vào ít nhất một tín hiệu đối tượng âm thanh, và trong đó thông tin tham số thứ hai bao gồm phép tham số hóa thừa số hiệu chỉnh,

trong đó bộ phận khôi phục hàm delta (142) được tạo cấu hình để đảo ngược phép tham số hóa thừa số hiệu chỉnh để thu hàm delta, và

trong đó bộ phận áp dụng delta (143) được tạo cấu hình để áp dụng hàm delta trên các giá trị tham số để xác định thông tin giải trộn.

6. Bộ giải mã theo điểm 5,

trong đó phép tham số hóa thừa số hiệu chỉnh bao gồm nhiều hệ số dự báo tuyến tính,

trong đó bộ phận khôi phục hàm delta (142) được tạo cấu hình để đảo ngược phép tham số hóa thừa số hiệu chỉnh bằng cách tạo ra nhiều thừa số hiệu chỉnh phụ thuộc vào nhiều hệ số dự báo tuyến tính, và

trong đó bộ phận khôi phục hàm delta (142) được tạo cấu hình để tạo ra hàm delta dựa trên nhiều thừa số hiệu chỉnh.

7. Bộ giải mã theo một điểm bất kỳ trong số các điểm nêu trên,

trong đó bộ giải mã còn bao gồm bộ tạo ma trận giải trộn (118) để tạo ra ma trận giải trộn phụ thuộc vào thông tin phụ tham số thứ nhất, phụ thuộc vào thông tin phụ tham số thứ hai, và phụ thuộc vào thông tin kết xuất, và

trong đó môđun giải trộn (113) được tạo cấu hình để áp dụng ma trận giải trộn trên sự trộn giảm được biến đổi để thu tín hiệu âm thanh được giải trộn.

8. Bộ giải mã theo một điểm bất kỳ trong số các điểm nêu trên,

trong đó môđun giải trộn (113) bao gồm bộ phận khử tương quan (119) và bộ phận giải trộn (110),

trong đó bộ phận khử tương quan (119) được tạo cấu hình để điều khiển khử tương quan trên sự trộn giảm được biến đổi để thu kết quả khử tương quan,

và trong đó bộ phận giải trộn (110) được tạo cấu hình để sử dụng kết quả khử tương quan để thu tín hiệu âm thanh được giải trộn.

9. Bộ mã hóa để mã hóa một hoặc nhiều tín hiệu đối tượng âm thanh đầu vào, bộ mã hóa bao gồm:

bộ phận trộn giảm (91) để trộn giảm một hoặc nhiều tín hiệu đối tượng âm thanh đầu vào để thu một hoặc nhiều tín hiệu trộn giảm, và

bộ tạo thông tin phụ tham số (93) để tạo ra thông tin phụ tham số thứ nhất trên ít nhất một tín hiệu đối tượng âm thanh và thông tin phụ tham số thứ hai trên ít nhất một tín hiệu đối tượng âm thanh, sao cho độ phân giải tần số của thông tin phụ tham số thứ hai cao hơn độ phân giải tần số của thông tin phụ tham số thứ nhất,

trong đó bộ tạo thông tin phụ tham số (93) được tạo cấu hình để tạo thông tin phụ tham số thứ nhất và thông tin phụ tham số thứ hai để một hoặc nhiều tín hiệu đối tượng âm thanh đầu vào có thể giải mã được từ một hoặc nhiều tín hiệu trộn giảm sử dụng thông tin phụ tham số thứ nhất cùng với thông tin phụ tham số thứ hai.

10. Bộ mã hóa theo điểm 9,

trong đó bộ mã hóa còn bao gồm bộ phận biến đổi (92) để biến đổi một hoặc nhiều tín hiệu đối tượng âm thanh đầu vào từ miền thời gian thành miền thời gian-tần số để thu một hoặc nhiều tín hiệu đối tượng âm thanh được biến đổi, và

trong đó bộ tạo thông tin phụ tham số (93) được tạo cấu hình để tạo ra thông tin phụ tham số thứ nhất và thông tin phụ tham số thứ hai dựa trên một hoặc nhiều tín hiệu đối tượng âm thanh được biến đổi.

11. Bộ mã hóa theo điểm 10,

trong đó bộ phận biến đổi (92) được tạo cấu hình để biến đổi một hoặc nhiều tín hiệu đối tượng âm thanh đầu vào từ miền thời gian thành miền thời gian-tần số phụ thuộc vào chiều dài cửa sổ của khối biến đổi tín hiệu bao gồm các giá trị tín hiệu của ít nhất một trong số một hoặc nhiều tín hiệu đối tượng âm thanh đầu vào,

trong đó bộ phận biến đổi (92) bao gồm bộ phận phát hiện tín hiệu nhất thời (101) để xác định kết quả phát hiện tín hiệu nhất thời biểu thị có hoặc không có tín hiệu nhất thời có mặt trong một hoặc nhiều của ít nhất một tín hiệu đối tượng âm thanh, trong đó tín hiệu nhất thời biểu thị sự thay đổi tín hiệu trong một hoặc nhiều tín hiệu trong số ít nhất một tín hiệu đối tượng âm thanh, và

trong đó bộ phận biến đổi (92) còn bao gồm bộ phận tạo chuỗi cửa sổ (102) để xác định chiều dài cửa sổ phụ thuộc vào kết quả phát hiện tín hiệu nhất thời.

12. Bộ mã hóa theo một điểm trong số các điểm từ 9 đến 11, trong đó bộ mã hóa còn bao gồm bộ phận ước lượng delta (108) để ước lượng nhiều thừa số hiệu chỉnh dựa trên nhiều giá trị tham số phụ thuộc vào ít nhất một tín hiệu đối tượng âm thanh để thu thông tin phụ tham số thứ hai.

13. Bộ mã hóa theo điểm 12, trong đó bộ mã hóa còn bao gồm bộ phận mô hình hóa delta (109) để xác định các hệ số dự báo tuyến tính từ nhiều thừa số hiệu chỉnh bằng cách điều khiển dự báo tuyến tính.

14. Hệ thống mã hóa đối tượng âm thanh không gian đa phân giải tương thích ngược bao gồm:

bộ mã hóa (61) theo một điểm bất kỳ trong số các điểm từ 9 đến 13 để mã hóa một hoặc nhiều tín hiệu đối tượng âm thanh đầu vào bằng cách thu một hoặc nhiều tín

hiệu trộn giảm biểu thị sự trộn giảm của một hoặc nhiều tín hiệu đối tượng âm thanh đầu vào, bằng cách thu thông tin phụ tham số thứ nhất trên ít nhất một tín hiệu đối tượng âm thanh, và bằng cách thu thông tin phụ tham số thứ hai trên ít nhất một tín hiệu đối tượng âm thanh, trong đó độ phân giải tần số của thông tin phụ tham số thứ hai cao hơn độ phân giải tần số của thông tin phụ tham số thứ nhất, và

bộ giải mã (62) theo một điểm bất kỳ trong số các điểm từ 1 đến 8 để tạo ra tín hiệu âm thanh giải trộn dựa trên một hoặc nhiều tín hiệu trộn giảm, và dựa trên thông tin phụ tham số thứ nhất và thông tin phụ tham số thứ hai.

15. Phương pháp tạo ra tín hiệu âm thanh giải trộn bao gồm nhiều kênh âm thanh được giải trộn, trong đó phương pháp bao gồm các bước:

xác định thông tin giải trộn bằng cách nhận thông tin phụ tham số thứ nhất trên ít nhất một tín hiệu đối tượng âm thanh và thông tin phụ tham số thứ hai trên ít nhất một tín hiệu đối tượng âm thanh, trong đó độ phân giải tần số của thông tin phụ tham số thứ hai cao hơn độ phân giải tần số của thông tin phụ tham số thứ nhất, và

áp dụng thông tin giải trộn trên tín hiệu trộn giảm, biểu thị sự trộn giảm của ít nhất một tín hiệu đối tượng âm thanh, để thu tín hiệu âm thanh được giải trộn bao gồm nhiều kênh âm thanh được giải trộn,

trong đó bước xác định thông tin giải trộn bao gồm việc thay đổi thông tin tham số thứ nhất và thông tin tham số thứ hai để thu thông tin tham số được thay đổi, sao cho thông tin tham số được thay đổi có độ phân giải tần số cao hơn độ phân giải tần số thứ nhất.

16. Phương pháp mã hóa một hoặc nhiều tín hiệu đối tượng âm thanh đầu vào, phương pháp bao gồm các bước:

trộn giảm một hoặc nhiều tín hiệu đối tượng âm thanh đầu vào để thu một hoặc nhiều tín hiệu trộn giảm, và

tạo ra thông tin phụ tham số thứ nhất trên ít nhất một tín hiệu đối tượng âm thanh và thông tin phụ tham số thứ hai trên ít nhất một tín hiệu đối tượng âm thanh, sao cho độ phân giải tần số của thông tin phụ tham số thứ hai cao hơn độ phân giải tần số của thông tin phụ tham số thứ nhất,

trong đó thông tin phụ tham số thứ nhất và thông tin phụ tham số thứ hai được tạo ra để một hoặc nhiều tín hiệu đối tượng âm thanh đầu vào có thể giải mã được từ một hoặc nhiều tín hiệu trộn giảm sử dụng thông tin phụ tham số thứ nhất cùng với thông tin phụ tham số thứ hai.

17. Vật ghi có thể đọc được bằng máy tính chứa chương trình máy tính để thực hiện phương pháp theo điểm 15 khi được chạy trên máy tính hoặc bộ xử lý tín hiệu.

18. Vật ghi có thể đọc được bằng máy tính chứa chương trình máy tính để thực hiện phương pháp theo điểm 16 khi được chạy trên máy tính hoặc bộ xử lý tín hiệu.

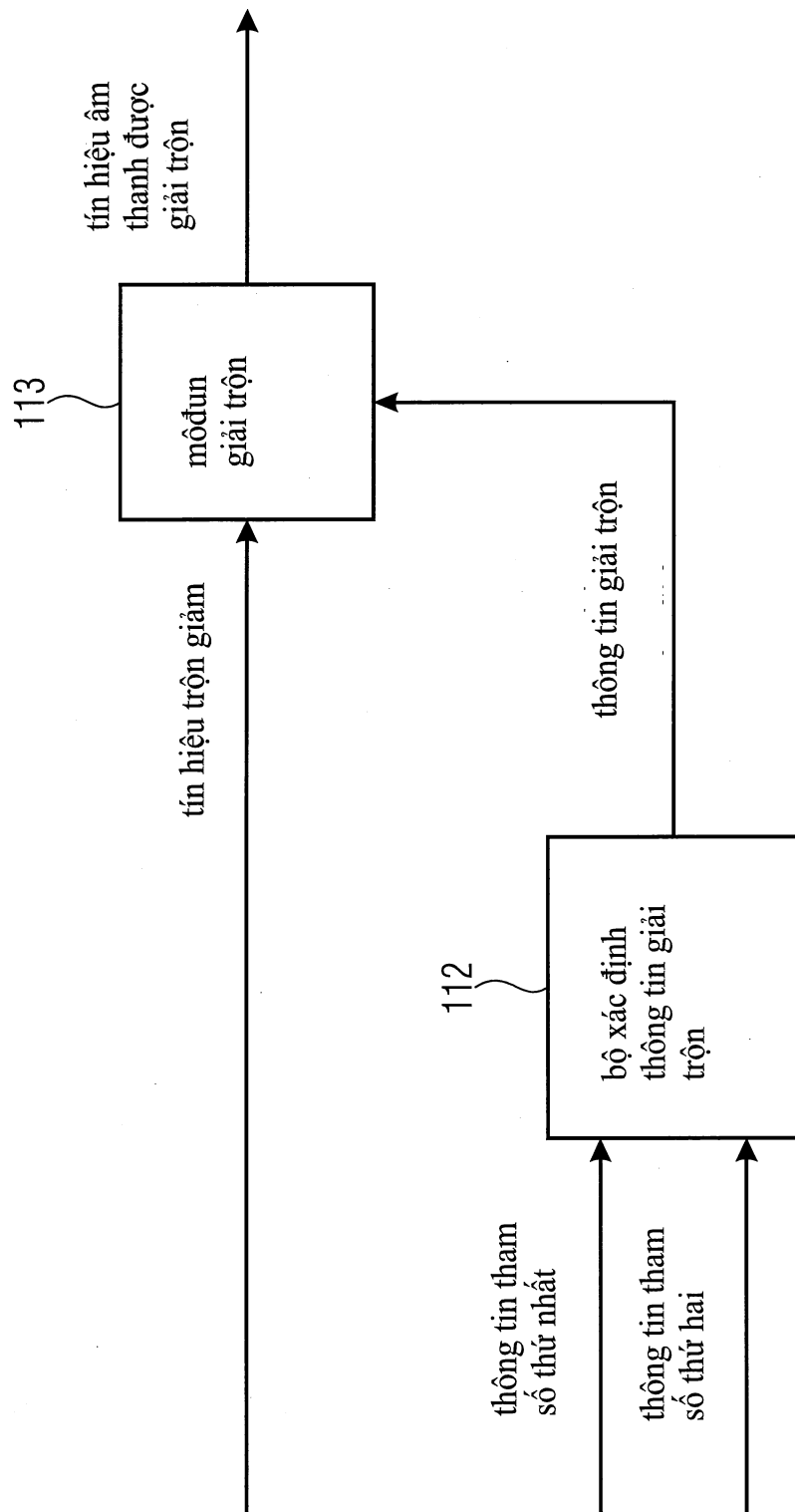


Fig.1A

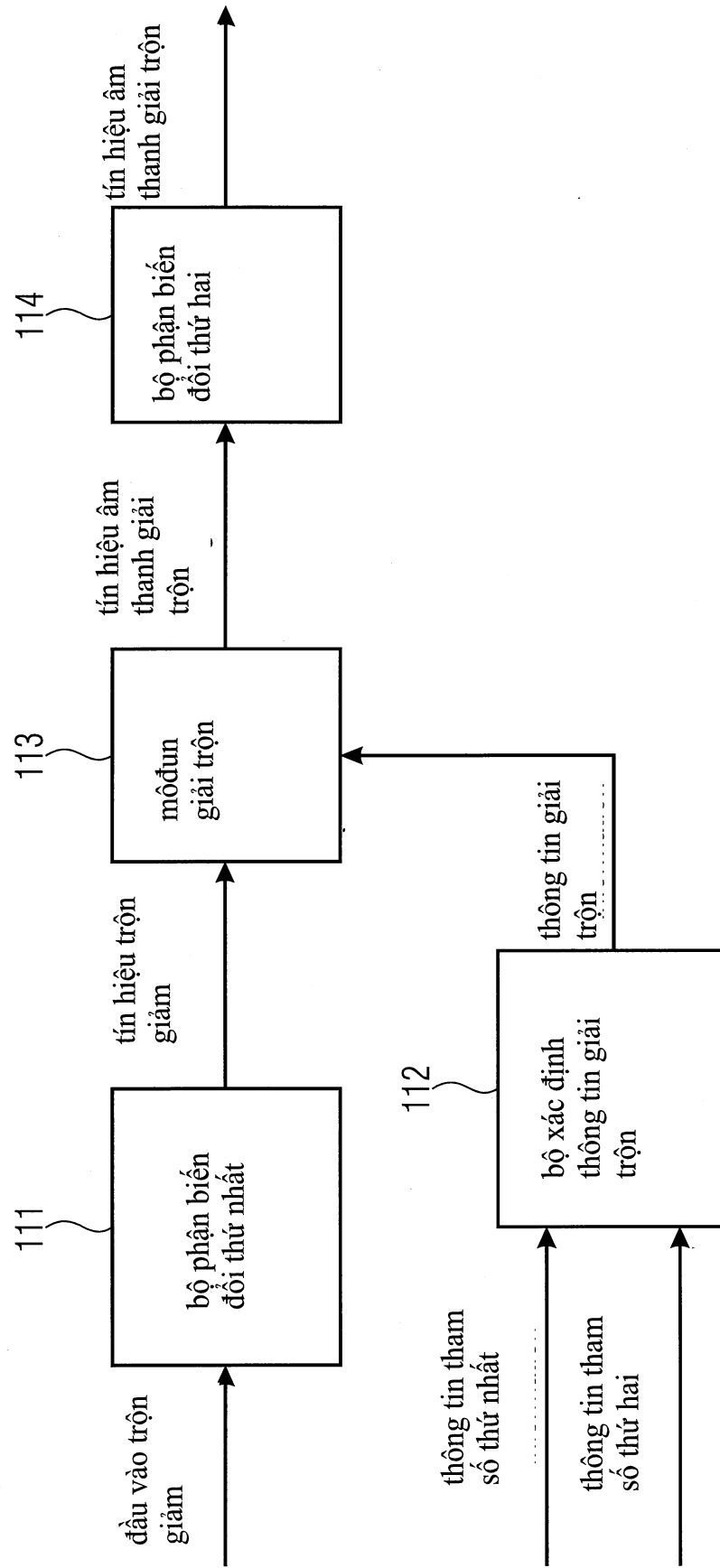
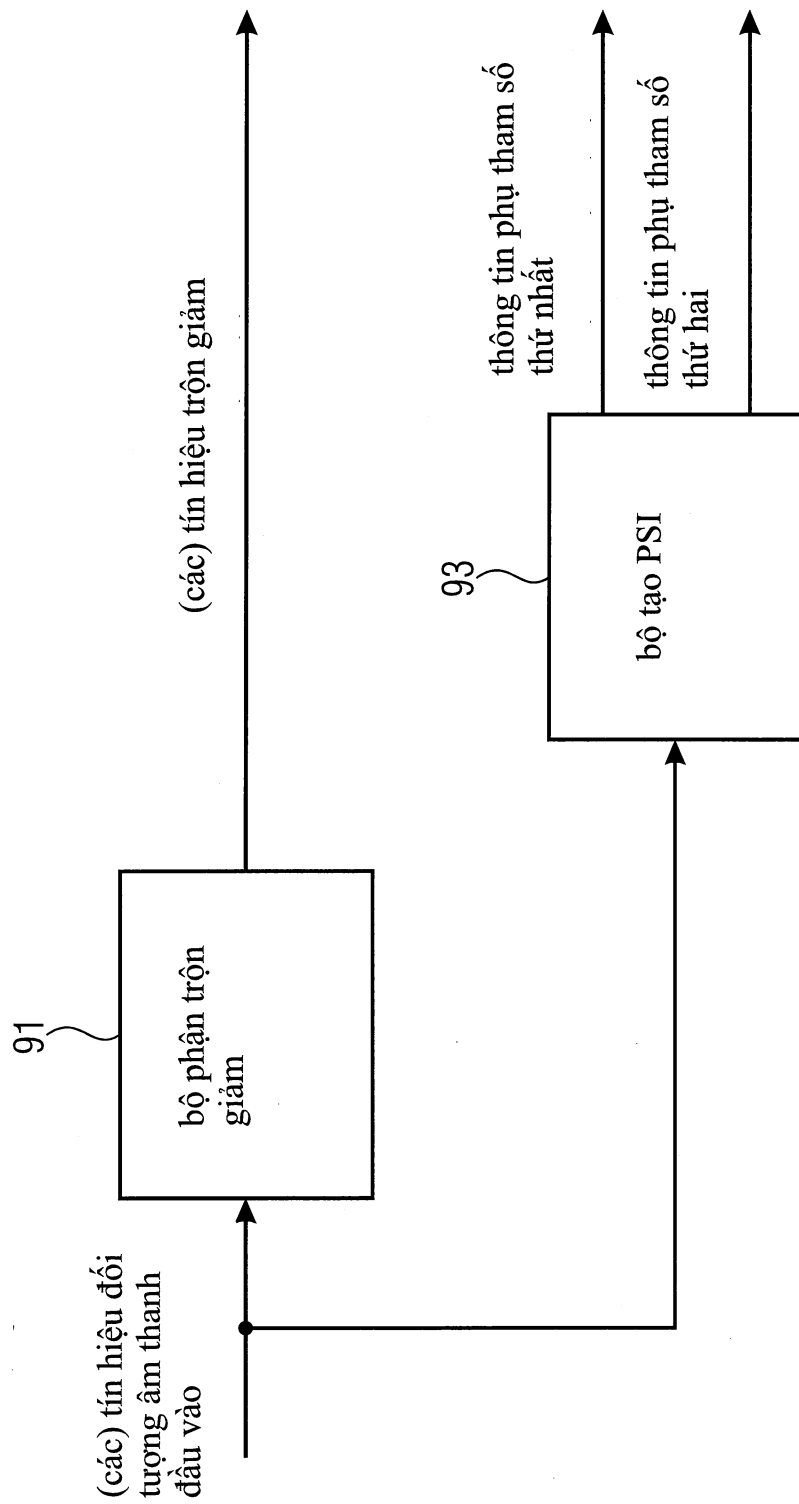


Fig.1B

**Fig.2A**

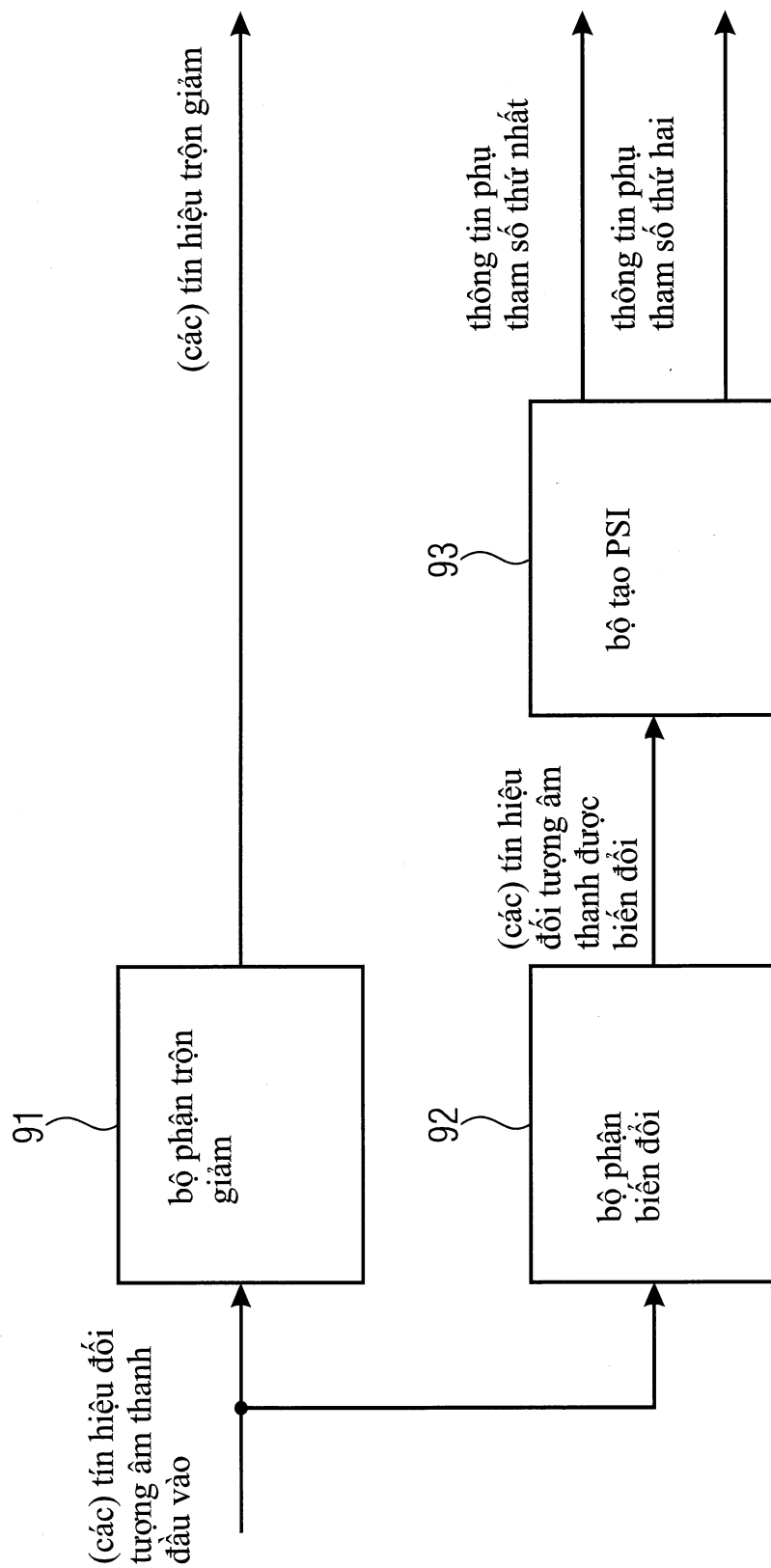


Fig.2B

	phần trộn giảm	phần thông tin phụ tham số	
		PSI thứ nhất	PSI thứ hai

Fig.2C

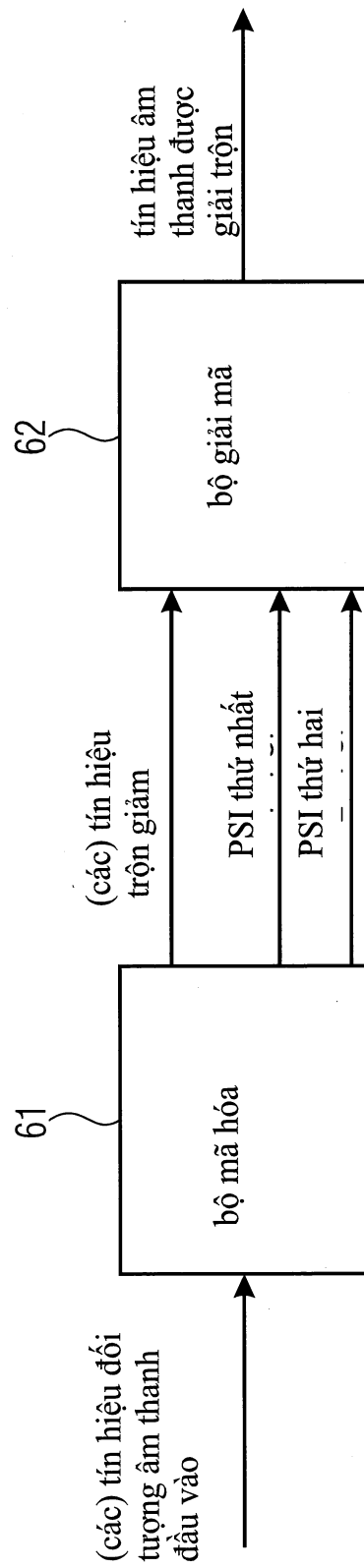


Fig.3

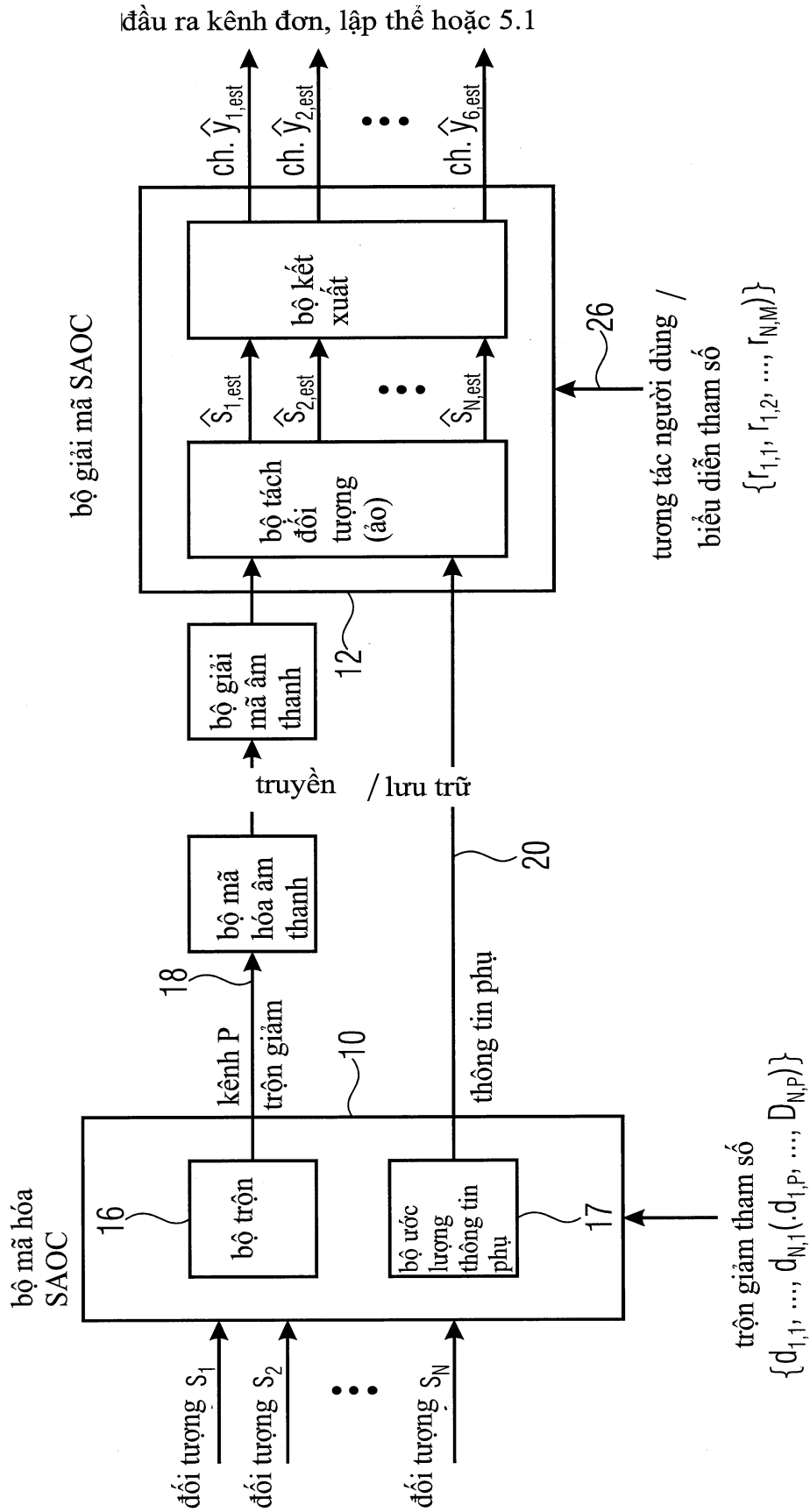


Fig.4

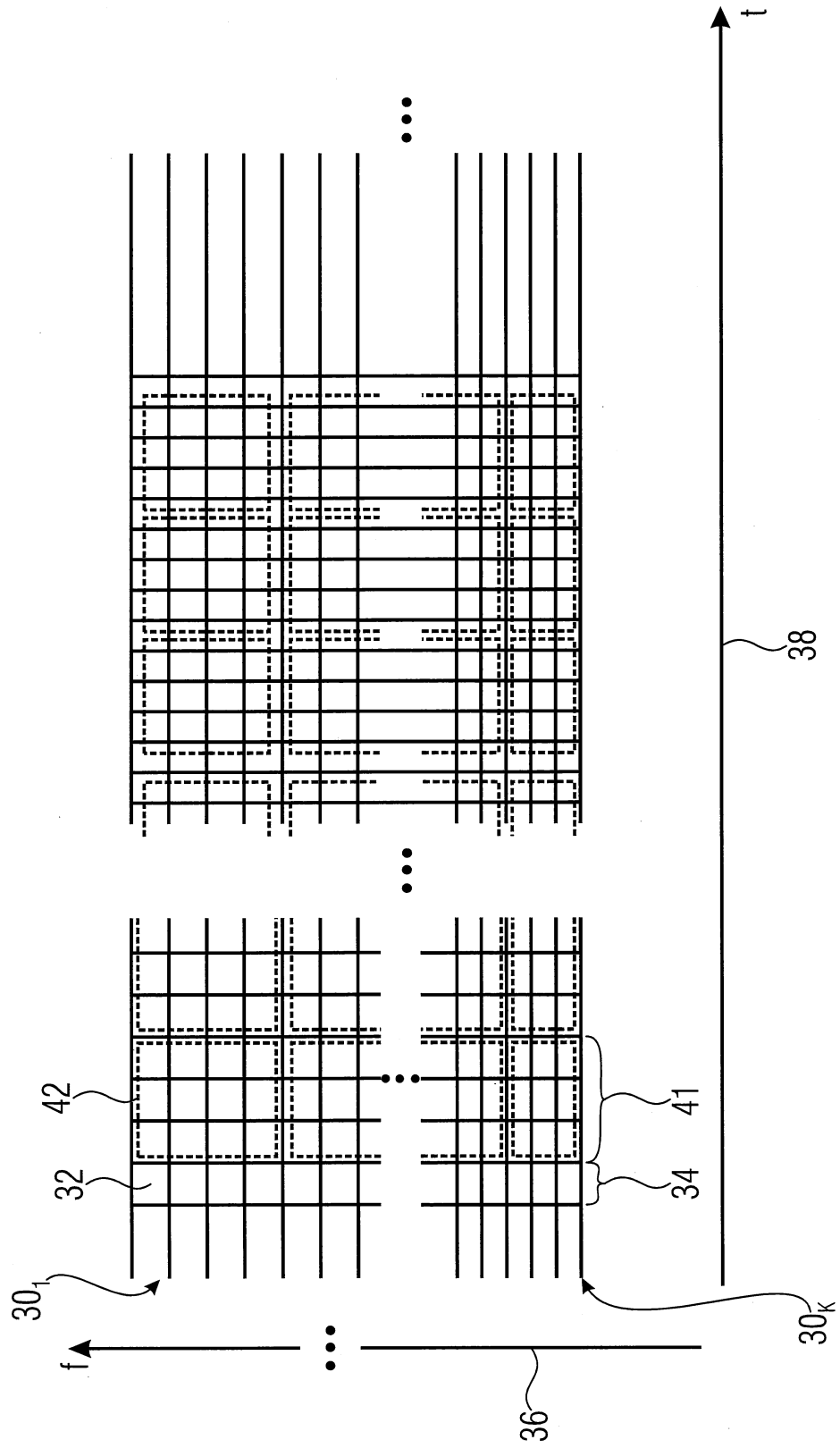


Fig.5

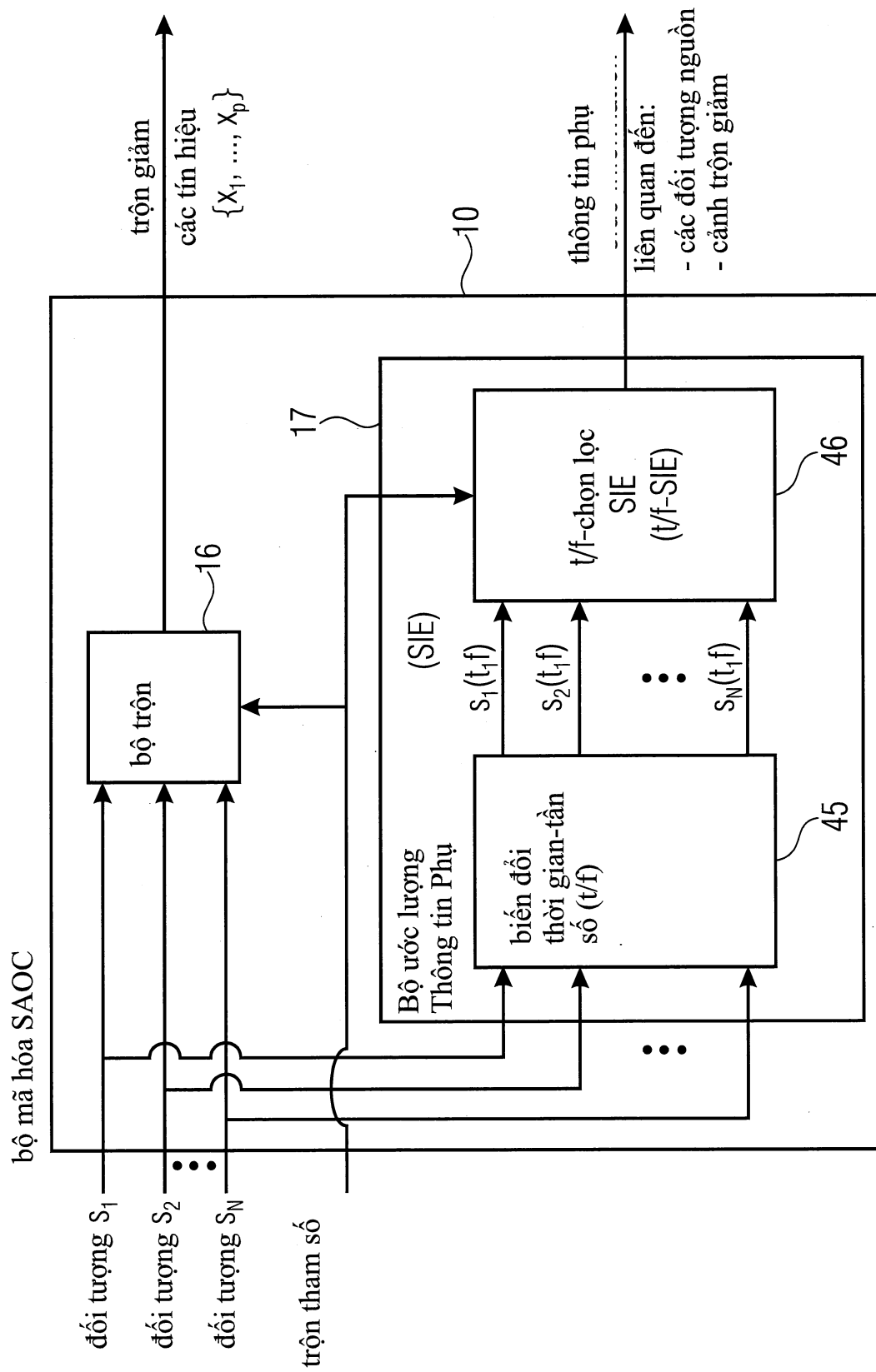


Fig.6

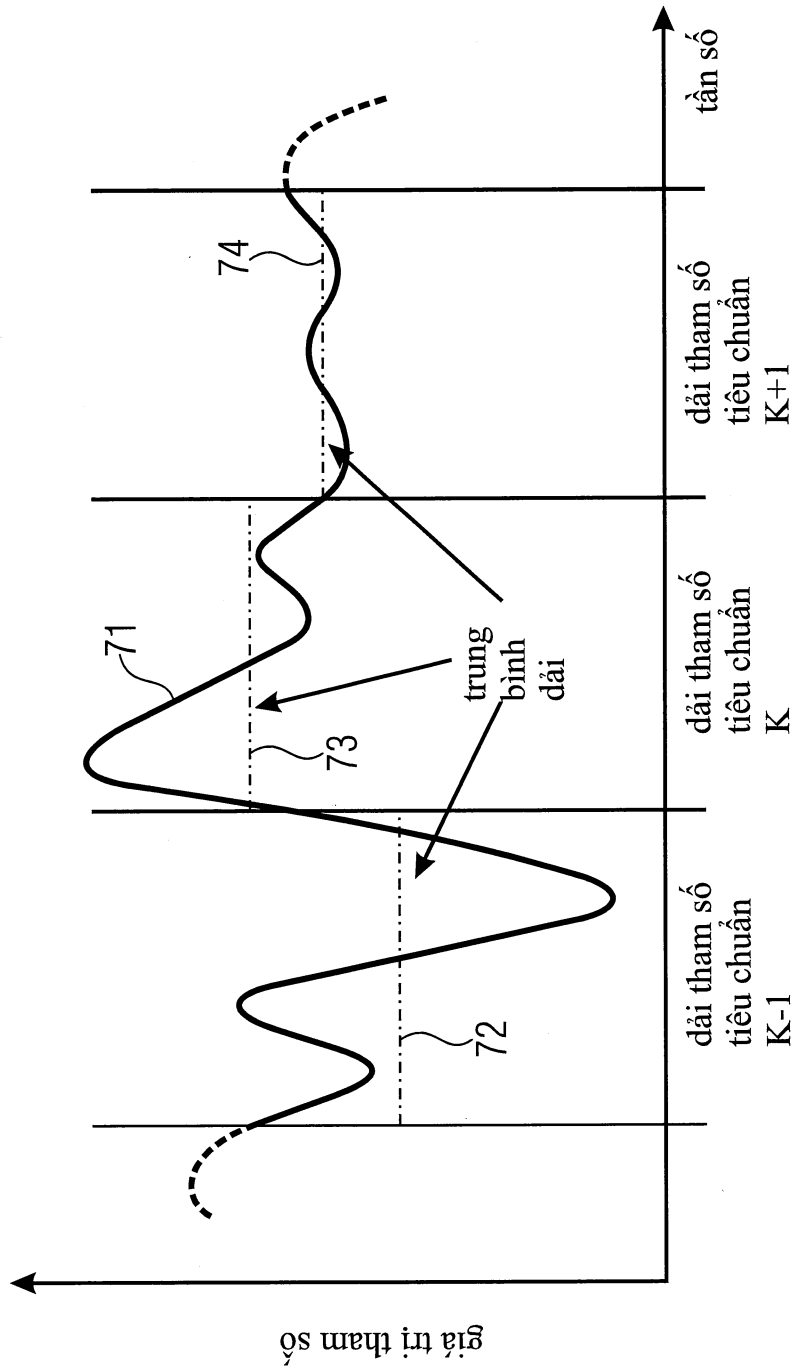


Fig.7

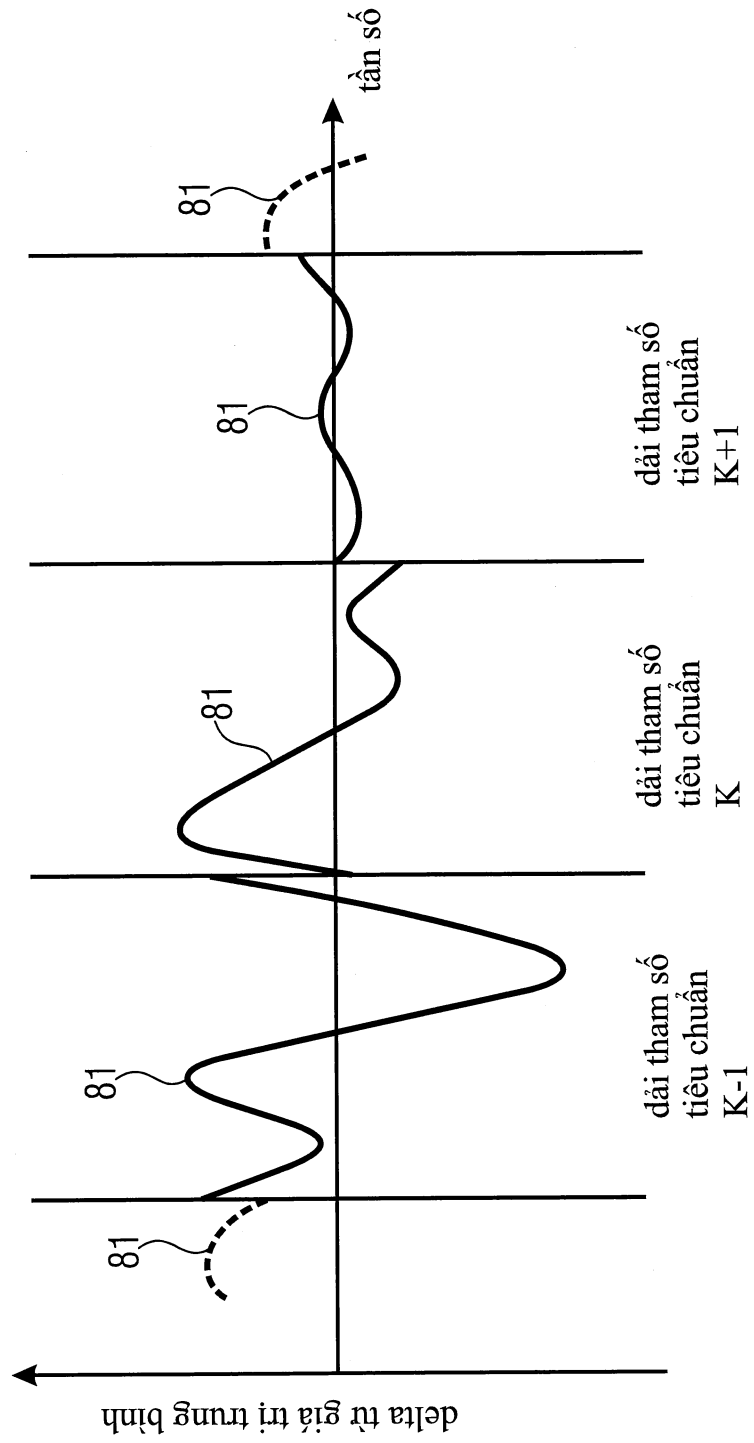


Fig.8

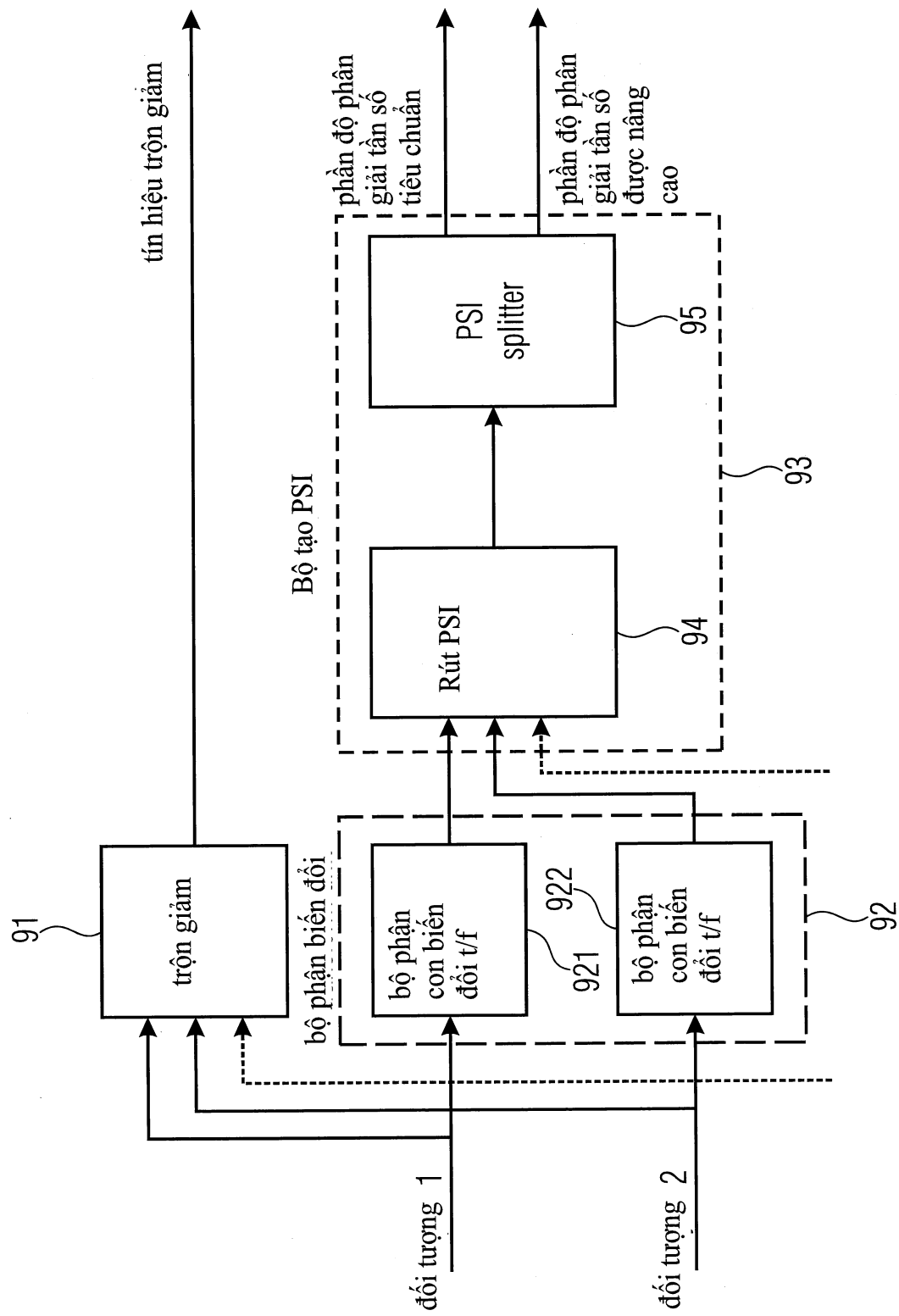


Fig.9

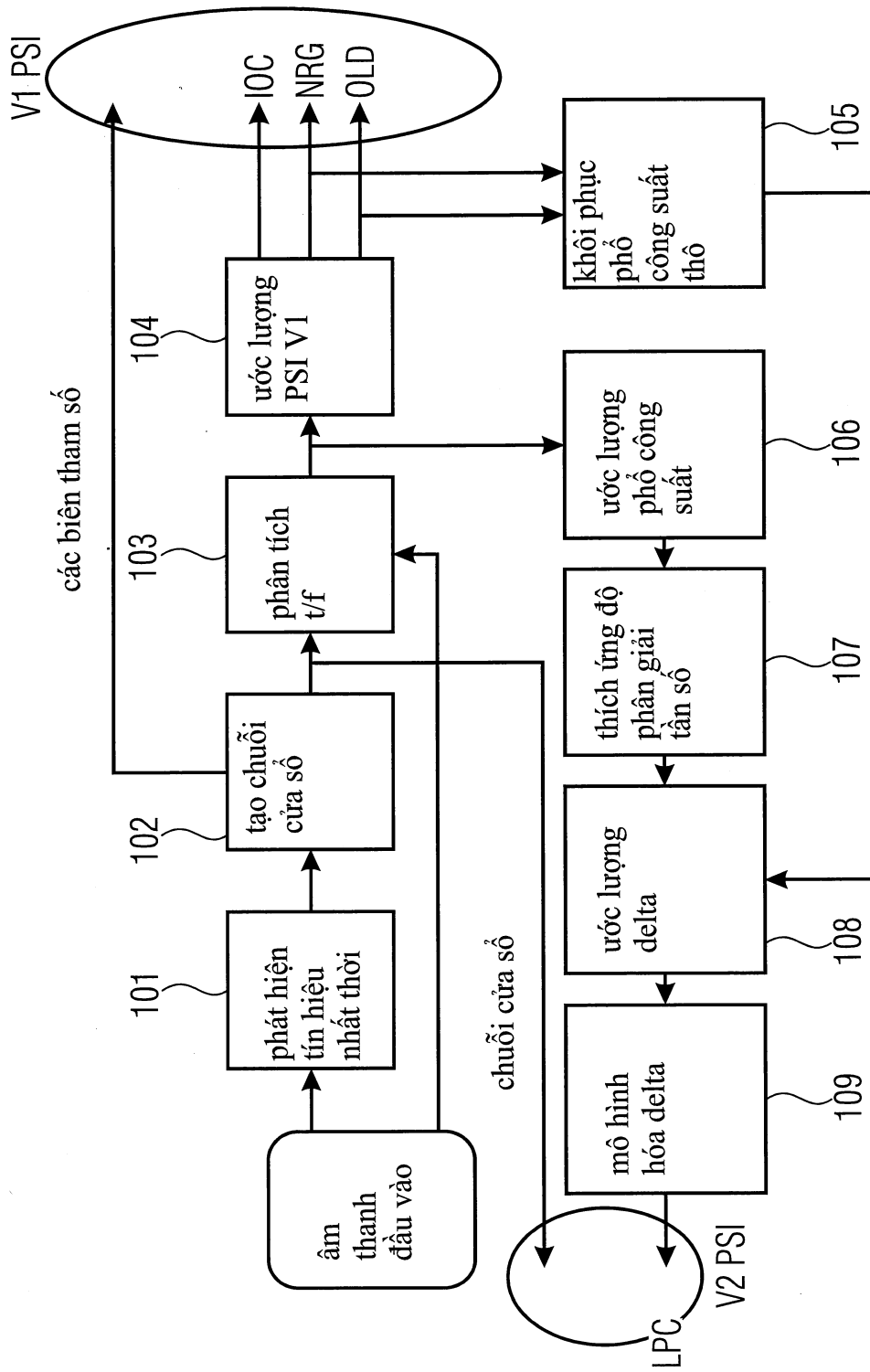


Fig.10

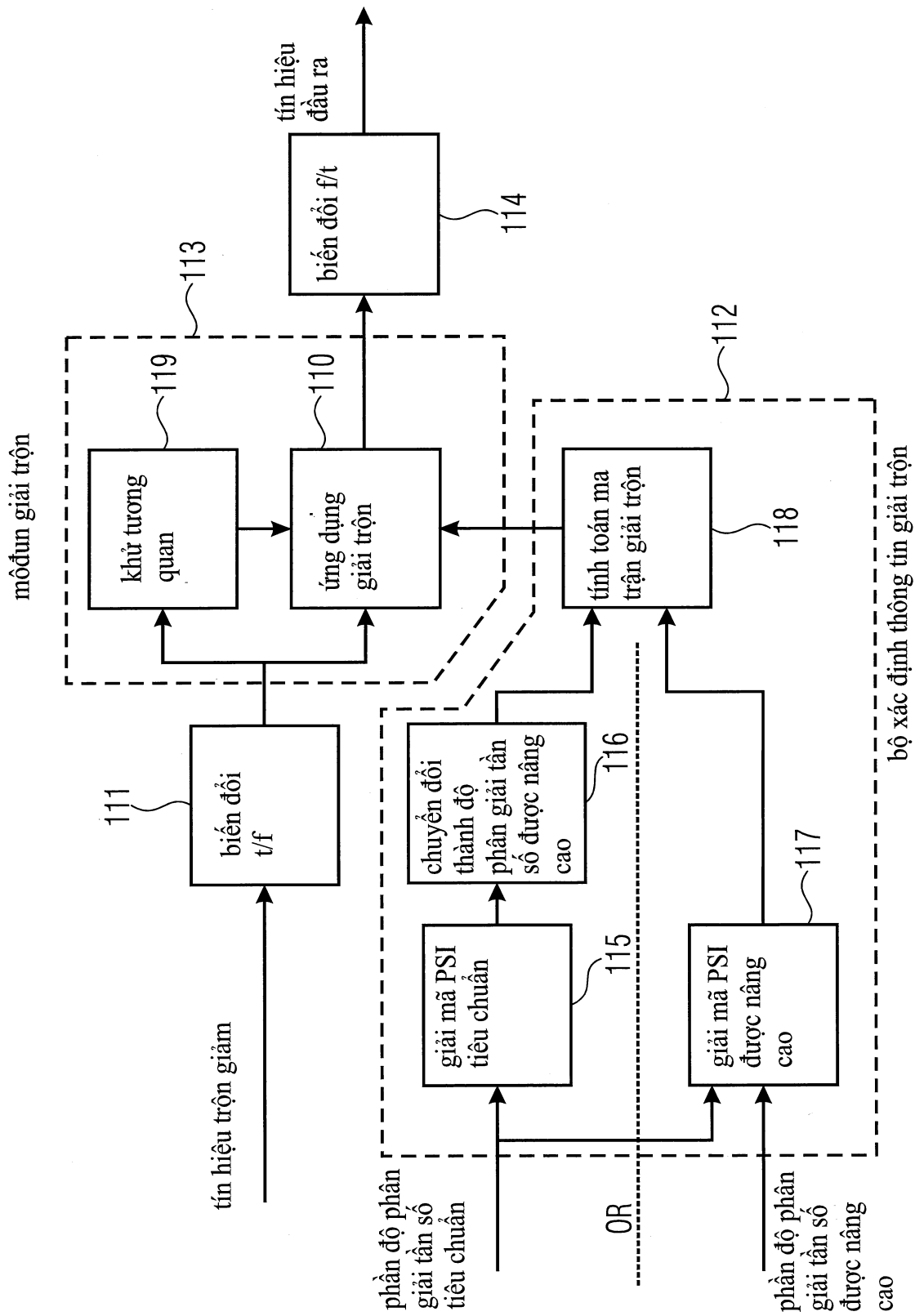


Fig.11

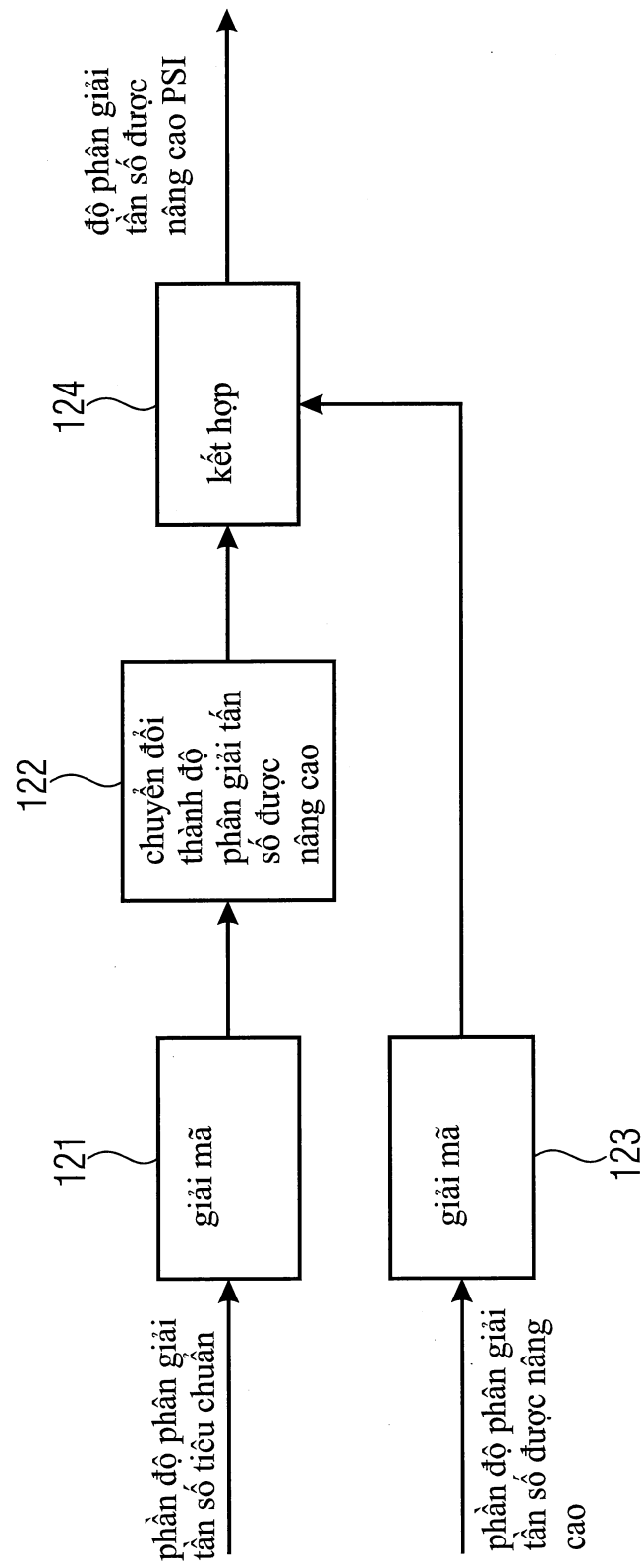


Fig.12

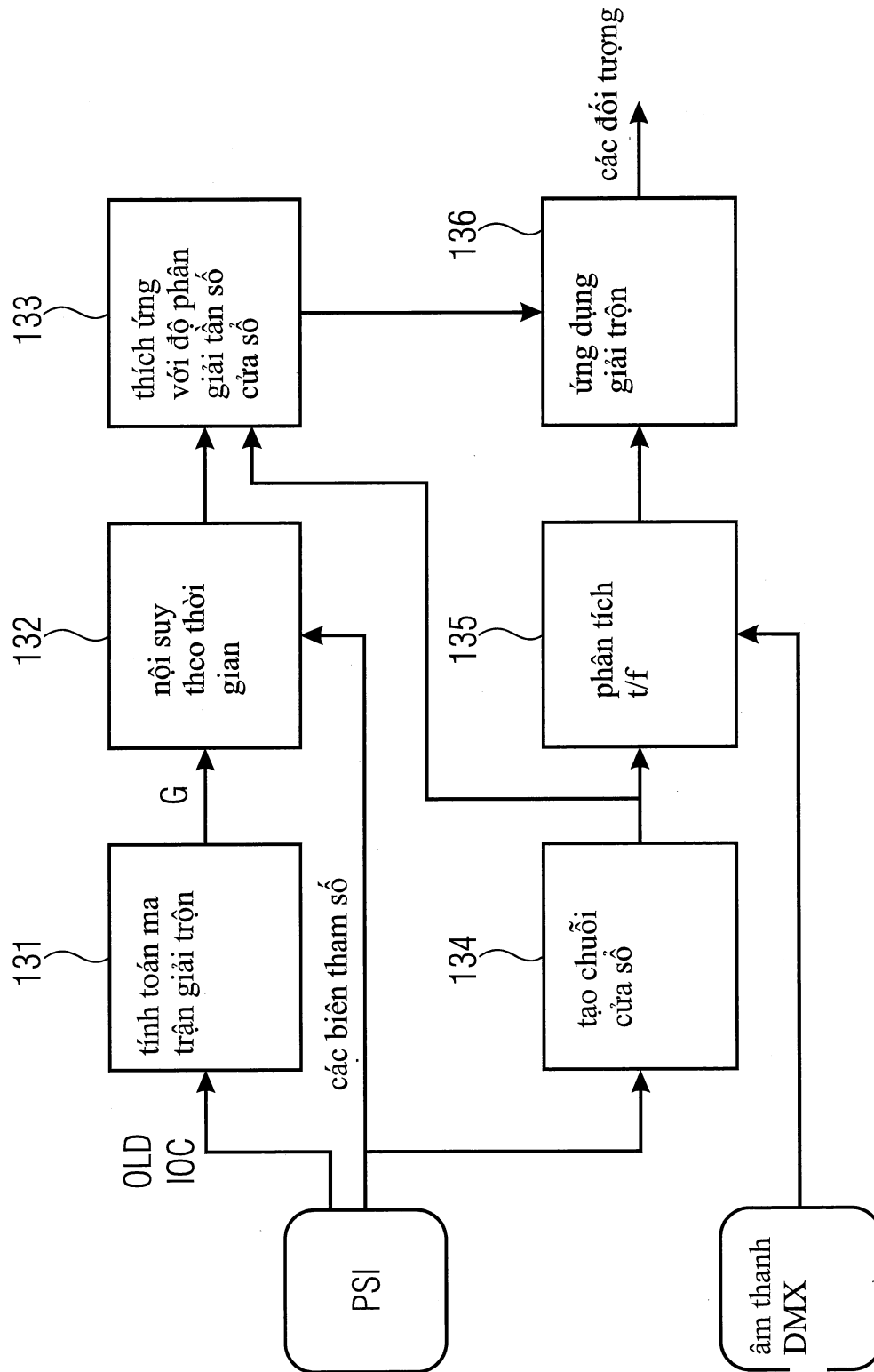


Fig.13

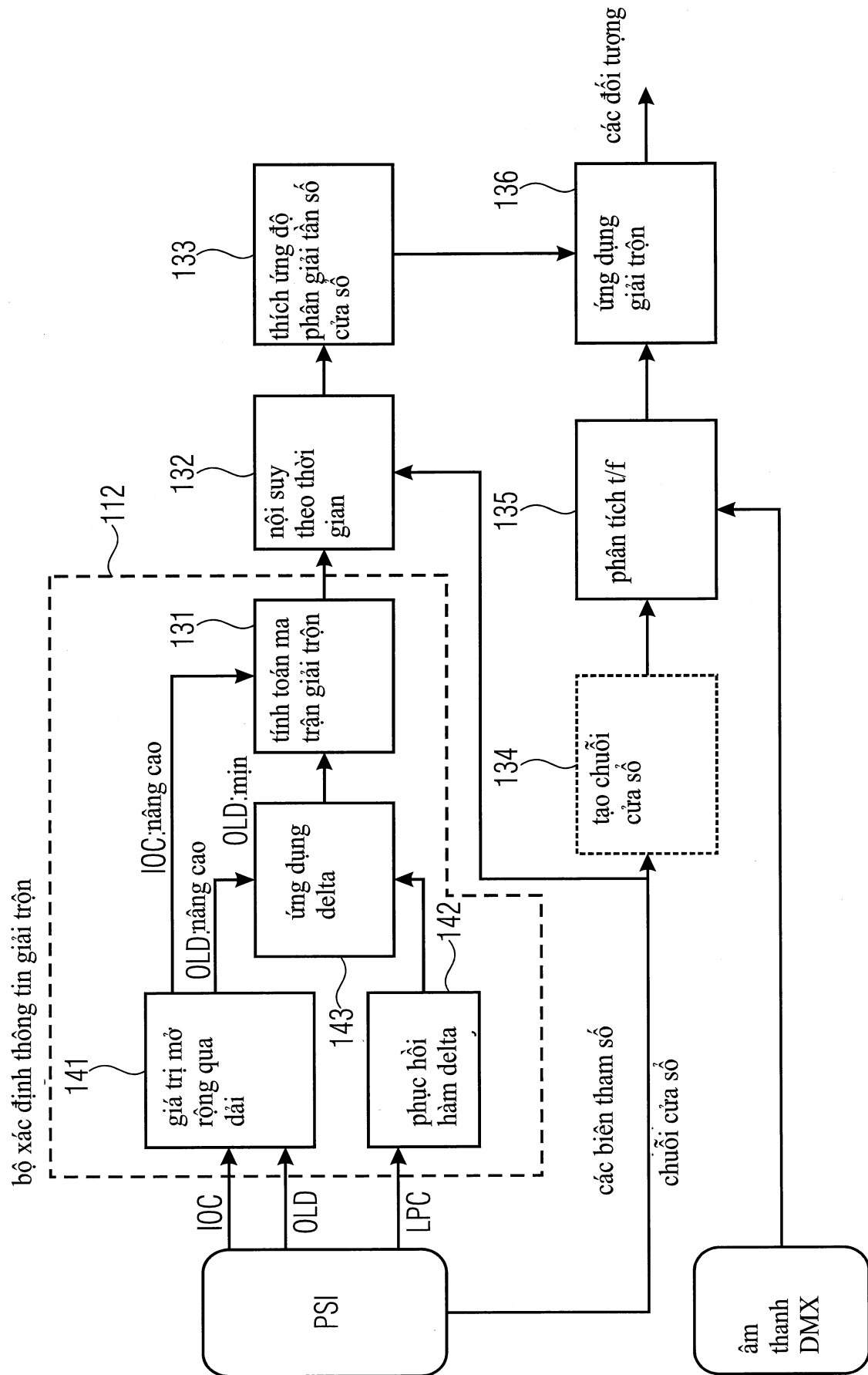


Fig.14

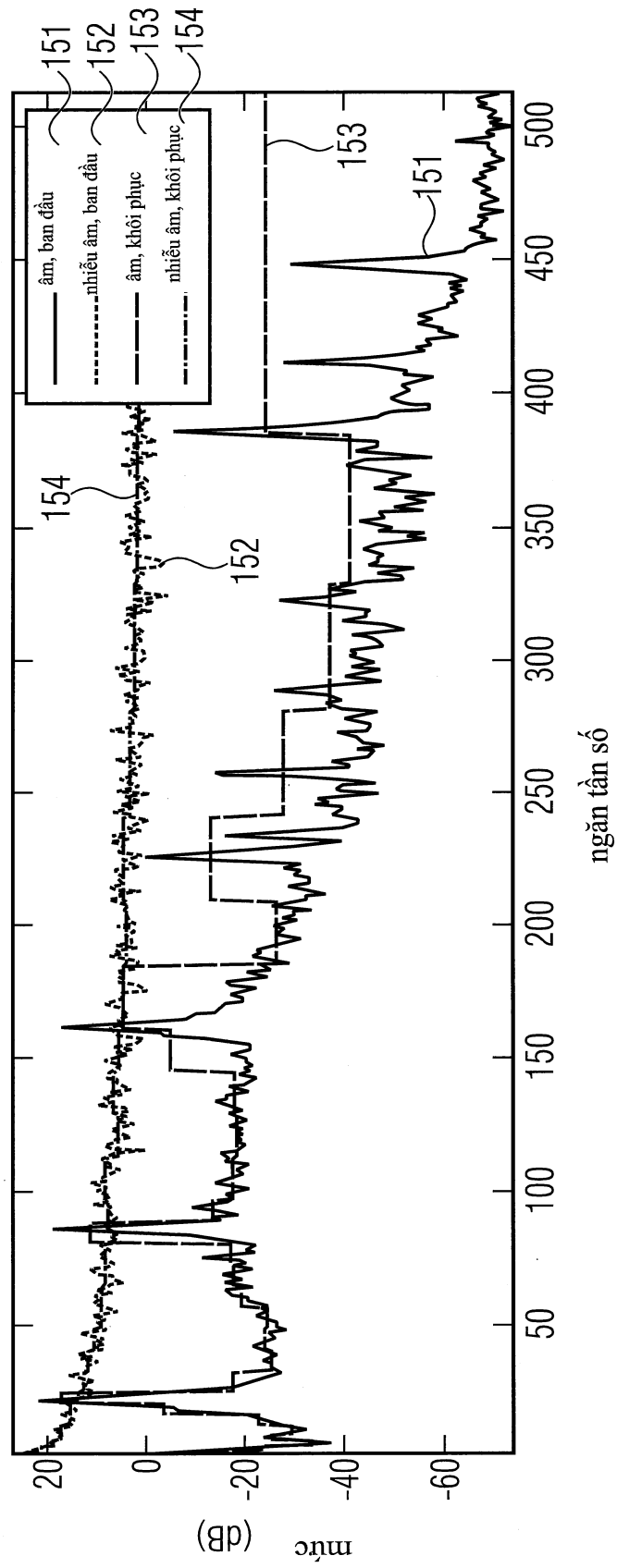


Fig.15

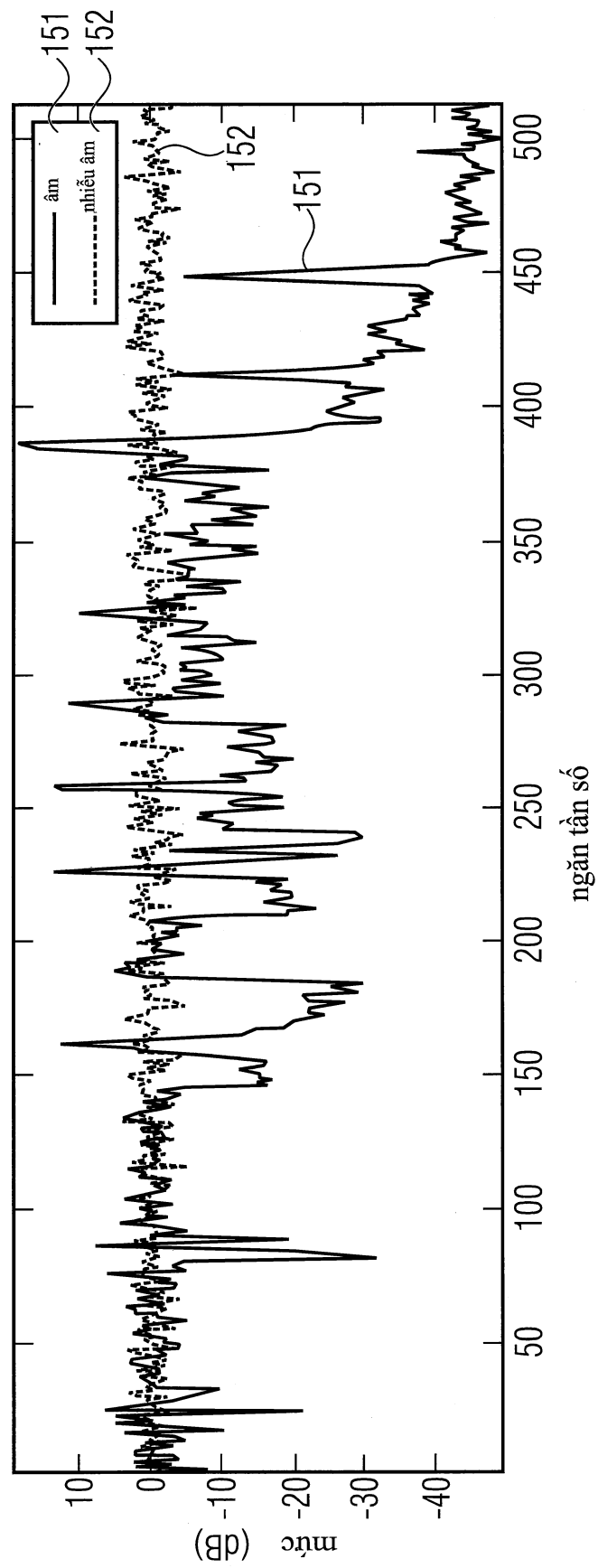


Fig.16

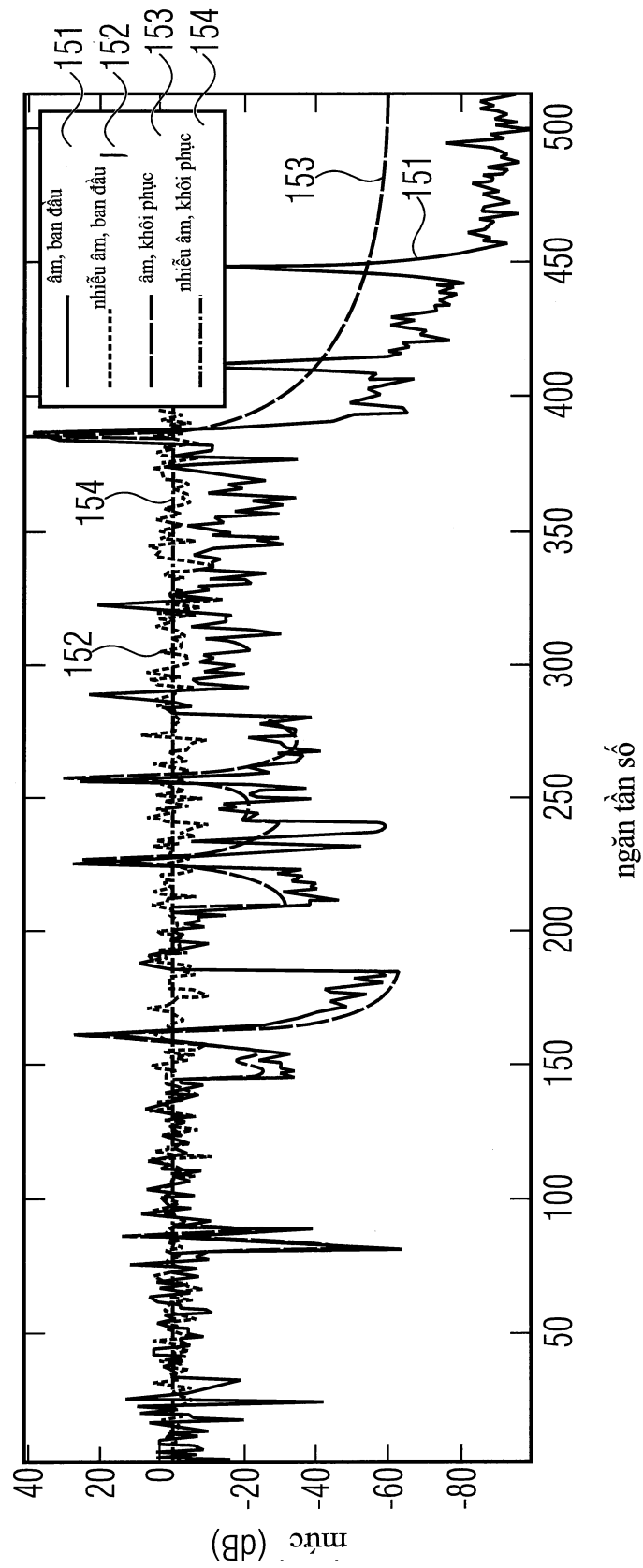


Fig.17

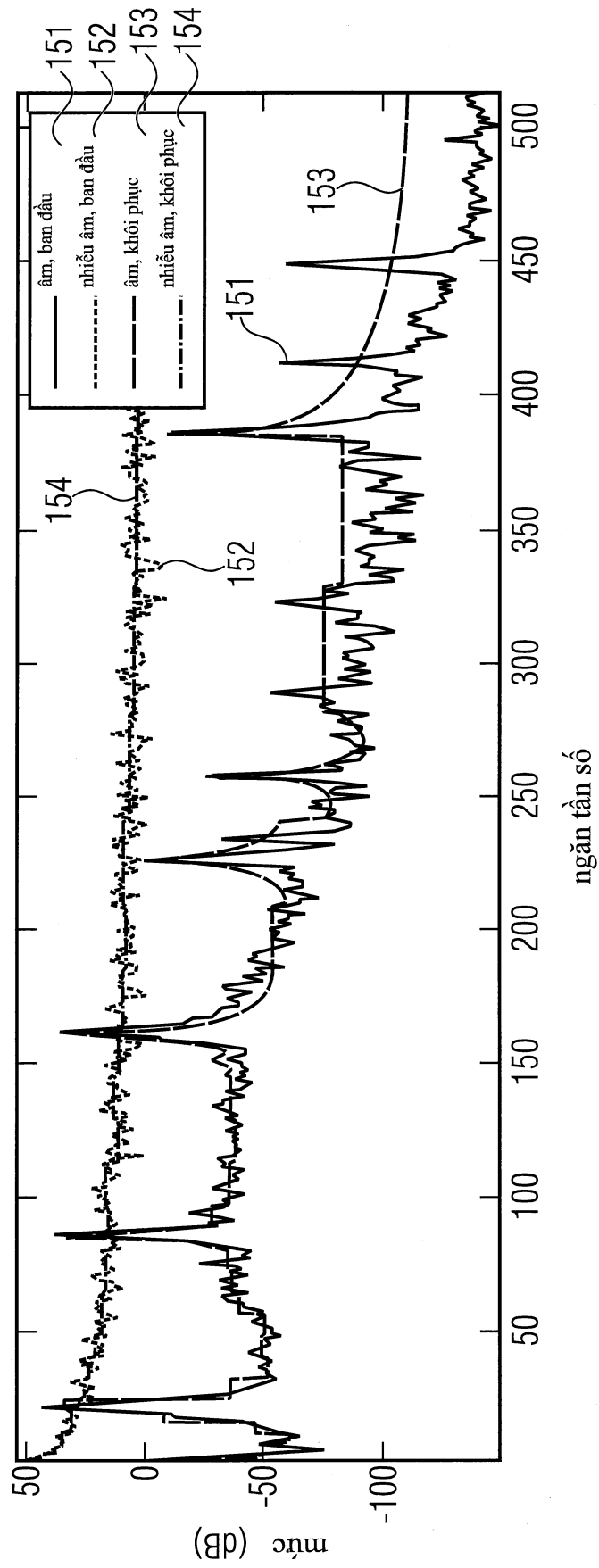


Fig.18