



(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)



1-0025892

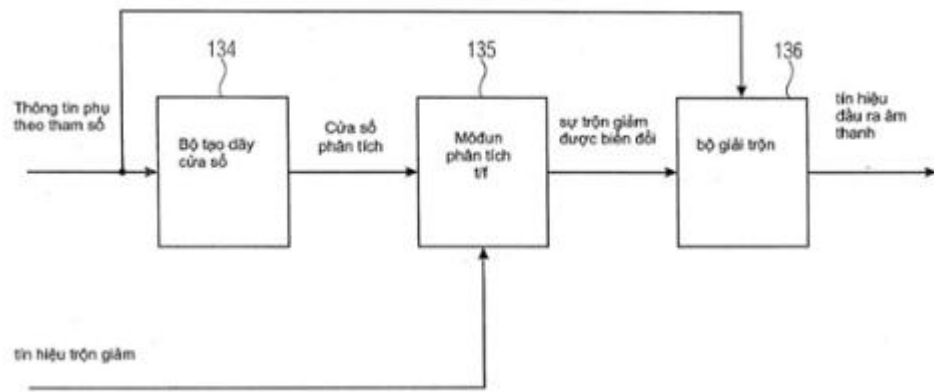
(51)⁷ G10L 19/025; G10L 19/008

(13) B

-
- (21) 1-2015-01497 (22) 02/10/2013
(86) PCT/EP2013/070551 02/10/2013 (87) WO/2014/053548 10/04/2014
(30) 61/710,133 05/10/2012 US; 13167481.4 13/05/2013 EP
(45) 26/10/2020 391 (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); PAULUS, Jouni (FI); EDLER, Bernd (DE); HELLMUTH,
Oliver (DE); HERRE, Juergen (DE); KASTNER, Thorsten (DE).
(74) CÔNG TY LUẬT TRÁCH NHIỆM HỮU HẠN AMBYS HÀ NỘI (AMBYS
HANOI)
-

(54) BỘ MÃ HÓA, BỘ GIẢI MÃ VÀ CÁC PHƯƠNG PHÁP MÃ HÓA VÀ GIẢI MÃ
DÙNG CHO VIỆC THÍCH ỨNG ĐỘNG TƯƠNG THÍCH NGƯỢC ĐỘ PHÂN
GIẢI THỜI GIAN/TẦN SỐ TRONG MÃ HÓA ĐỐI TƯỢNG ÂM THANH
TRONG KHÔNG GIAN

(57) Sáng chế đề cập đến bộ mã hóa, bộ giải mã và các phương pháp mã hóa và giải mã dùng cho việc thích ứng động tương thích ngược của độ phân giải thời gian/tần số trong mã hóa đối tượng âm thanh không gian. Bộ giải mã để tạo ra tín hiệu đầu ra âm thanh bao gồm một hoặc nhiều kênh đầu ra âm thanh từ tín hiệu trộn giảm bao gồm nhiều mẫu trộn giảm miền thời gian. Tín hiệu trộn giảm mã hóa hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh. Bộ giải mã bao gồm bộ tạo dãy cửa sổ (134) để xác định nhiều cửa sổ phân tích, trong đó mỗi cửa sổ phân tích bao gồm nhiều mẫu trộn giảm miền thời gian của tín hiệu trộn giảm. Mỗi cửa sổ phân tích trong số nhiều cửa sổ phân tích có chiều dài cửa sổ biểu thị số lượng mẫu trộn giảm miền thời gian của cửa sổ phân tích đã nêu. Bộ tạo dãy cửa sổ (134) được tạo cấu hình để xác định nhiều cửa sổ phân tích để cho chiều dài cửa sổ của mỗi cửa sổ phân tích phụ thuộc vào đặc tính tín hiệu của ít nhất hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh. Hơn nữa, bộ giải mã bao gồm môđun phân tích t/f (135) để biến đổi nhiều mẫu trộn giảm miền thời gian của mỗi cửa sổ phân tích trong số nhiều cửa sổ phân tích từ miền thời gian thành miền thời gian-tần số phụ thuộc vào chiều dài cửa sổ của cửa sổ phân tích đã nêu, để thu được sự trộn giảm được biến đổi. Hơn nữa, bộ giải mã bao gồm bộ giải trộn (136) để giải trộn sự trộn giảm được biến đổi dựa trên thông tin phụ theo tham số trên hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh để thu được tín hiệu đầu ra âm thanh. Hơn nữa, bộ mã hóa được đề xuất.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến việc mã hóa tín hiệu âm thanh, giải mã tín hiệu âm thanh và xử lý tín hiệu âm thanh và cụ thể là, đến bộ mã hóa, bộ giải mã và các phương pháp mã hóa và giải mã dùng cho việc thích ứng động tương thích ngược của độ phân giải thời gian/tần số trong mã hóa đối tượng âm thanh trong không gian (spatial-audio-object-coding - SAOC).

Tình trạng kỹ thuật của sáng chế

Trong các hệ thống âm thanh số hiện đại, có xu hướng chính cho phép các cải biên liên quan đến đối tượng âm thanh của nội dung được truyền tải ở phía bộ nhận. Các cải biên này bao gồm các cải biên tăng ích của các phần tín hiệu âm thanh và/hoặc sự tái định vị trong không gian của các đối tượng âm thanh được chỉ định trong trường hợp phát lại đa kênh và các loa được phân chia trong không gian. Điều này có thể đạt được bằng cách đưa các phần khác nhau của nội dung âm thanh đến các loa khác nhau.

Nói cách khác, trong lĩnh vực xử lý âm thanh, việc truyền tải âm thanh và lưu trữ âm thanh, có mong muốn tăng nhu cầu cho phép tương tác người dùng về việc phát lại nội dung âm thanh hướng đối tượng và cả nhu cầu ứng dụng các khả năng mở rộng phát lại đa kênh để kết xuất một cách riêng biệt các nội dung âm thanh và các phần của chúng để cải thiện ấn tượng nghe. Bằng cách này, việc sử dụng nội dung âm thanh đa kênh tạo ra các cải thiện đáng kể đối với người dùng. Ví dụ, ấn tượng nghe ba chiều có thể thu được, mà mang lại sự thỏa mãn cho người dùng được cải thiện trong các ứng dụng giải trí. Tuy nhiên, nội dung âm thanh đa kênh cũng hữu dụng trong các môi trường chuyên nghiệp, ví dụ, trong các ứng dụng hội thảo từ xa, do độ rõ của người nói có thể được cải thiện bằng cách sử dụng sự phát lại âm thanh đa kênh. Có thể có ứng dụng khác là đề xuất người nghe của phần âm nhạc để điều chỉnh một cách riêng biệt mức phát lại và/hoặc vị trí trong không gian của các phần khác nhau (còn được gọi là “đối tượng âm thanh”) hoặc các đường rãnh, như phần phát âm hoặc các dụng cụ khác nhau. Người dùng có thể thực hiện sự điều chỉnh này với các lý do về sở thích

cá nhân, để ghi lại dễ dàng hơn một hoặc nhiều phần từ phần âm nhạc, mục đích giáo dục, karaoke, buổi diễn tập, v.v.

Việc truyền rời rạc thẳng của tất cả nội dung đa kênh số hoặc âm thanh đa đối tượng, ví dụ, ở dạng dữ liệu mã xung (PCM) hoặc thậm chí các định dạng âm thanh được nén, yêu cầu tốc độ bit rất cao. Tuy nhiên, cũng mong muốn truyền tải và lưu trữ dữ liệu âm thanh theo cách có hiệu quả với tốc độ bit. Do đó, một là sẵn sàng chấp nhận sự trao đổi hợp lý giữa chất lượng âm thanh và các yêu cầu tốc độ bit để tránh tải trọng tài nguyên dư thừa do các ứng dụng đa kênh/đa đối tượng gây ra.

Gần đây, trong lĩnh vực mã hóa âm thanh, các kỹ thuật tham số hóa với việc truyền tải/lưu trữ có hiệu quả với tốc độ bit của các tín hiệu âm thanh đa kênh/đa đối tượng được giới thiệu bởi, ví dụ, nhóm chuyên gia dịch chuyển ảnh (MPEG) và các nhóm khác. Một ví dụ là bao quanh MPEG (MPEG Surround - MPS) là phương pháp hướng kênh [MPS, BCC], hoặc mã hóa đối tượng âm thanh trong không gian (SAOC) MPEG làm phương pháp hướng đối tượng [JSC, SAOC, SAOC1, SAOC2]. Phương pháp hướng đối tượng khác được gọi là “tách nguồn được thông tin” [ISS1, ISS2, ISS3, ISS4, ISS5, ISS6]. Các kỹ thuật này nhằm tái cấu trúc ngữ cảnh âm thanh đầu ra mong muốn hoặc đối tượng âm thanh mong muốn dựa trên sự hoà trộn của các kênh /đối tượng và thông tin phụ bổ sung mô tả ngữ cảnh âm thanh được truyền tải/lưu trữ và/hoặc các đối tượng nguồn âm thanh trong ngữ cảnh âm thanh.

Việc ước tính và ứng dụng đối tượng/kênh liên quan đến thông tin phụ theo cách chọn lọc thời gian-tần số. Do đó, các hệ thống này sử dụng các biến đổi thời gian-tần số như Phép biến đổi Fourier rời rạc (Discrete Fourier Transform - DFT), Phép biến đổi Fourier thời gian ngắn (Short Time Fourier Transform - STFT) hoặc các dàn lọc như các dàn lọc gương vuông góc (Quadrature Mirror Filter - QMF), v.v. Nguyên lý cơ bản của các hệ thống này được mô tả trên Fig.3, sử dụng ví dụ MPEG SAOC.

Trong trường hợp của STFT, kích thước thời gian được biểu diễn bằng số khối thời gian và kích thước phổ được giữ lại bằng hệ số phổ (“bin”). Trong trường hợp của QMF, kích thước thời gian được biểu diễn bằng số khe thời gian và kích thước phổ được giữ lại bằng số dải con. Nếu độ phân giải phổ của QMF được cải thiện bằng ứng

dụng kế tiếp của giai đoạn lọc thứ hai, toàn bộ dàn lọc được gọi là QMF lai và các dải con độ phân giải mịn được gọi là các dải con lai.

Như được kể đến trên đây, trong SAOC, việc xử lý chung được thực hiện theo cách chọn lọc thời gian-tần số và có thể được mô tả như sau nằm trong mỗi dải tần số, như được mô tả trên Fig.3:

- N tín hiệu đối tượng âm thanh đầu vào $s_1 \dots s_N$ được trộn giảm vào P kênh $x_1 \dots x_P$ như phần của việc xử lý 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 tách thông tin phụ mô tả các đặc điểm của các đối tượng âm thanh đầu vào (môđun ước tính thông tin phụ (side-information-estimator - SIE)). Đối với MPEG SAOC, các mối liên hệ của các công suất đối tượng đối với nhau là dạng cơ bản nhất của thông tin phụ này.
- (Các) tín hiệu trộn giảm và thông tin phụ được truyền tải/lưu trữ. Do đó, (các) tín hiệu âm thanh trộn giảm có thể được nén, ví dụ sử dụng các bộ mã hóa âm thanh giác quan đã biết như MPEG-1/2 lớp II hoặc III (còn được biết đến là .mp3), mã hóa âm thanh tiên tiến (Advanced Audio Coding – AAC) MPEG-2/4, v.v.
- Ở đầu nhận, bộ giải mã thử theo quan niệm để phục hồi các tín hiệu đối tượng gốc (“tách đối tượng”) từ các tín hiệu trộn giảm (được giải mã) sử dụng thông tin phụ được truyền tải. Tiếp đó, các tín hiệu đối tượng gần đúng $\hat{s}_1 \dots \hat{s}_N$ được trộn với ngữ cảnh đích được biểu diễn bằng M kênh đầu ra âm thanh $\hat{y}_1 \dots \hat{y}_M$ sử dụng ma trận kết xuất được mô tả bằng các hệ số $r_{1,1} \dots r_{N,M}$ trên Fig.3. Ngữ cảnh đích mong muốn có thể là, trong trường hợp đặc biệt, kết xuất chỉ một tín hiệu nguồn ra khỏi hỗn hợp (kịch bản tách nguồn), mà còn ngữ cảnh âm thanh tùy ý bất kỳ khác của các đối tượng được truyền tải. Ví dụ, đầu ra có thể là ngữ cảnh đích một kênh, 2 kênh lập thể hoặc đa kênh 5.1.

Các hệ thống dựa trên thời gian tần số có thể ứng dụng biến đổi thời gian-tần số (t/f) với độ phân giải thời gian và tần số tĩnh. Việc lựa chọn mạng lưới phân giải t/f được cố định nhất định thường bao gồm sự cân bằng giữa độ phân giải thời gian và tần số.

Tác dụng của độ phân giải của t/f cố định có thể được chứng minh ở ví dụ của các tín hiệu đối tượng thông thường trong hỗn hợp tín hiệu âm thanh. Ví dụ, phổ tiếng âm thể hiện cấu trúc liên quan hài hòa với tần số cơ bản và một số âm bội. Năng lượng của các tín hiệu này được tập trung ở các vùng tần số nhất định. Với các tín hiệu này, độ phân giải tần số cao của phép biểu diễn t/f được ứng dụng có lợi để tách các vùng âm dải hẹp từ hỗn hợp tín hiệu. Ngược lại, các tín hiệu tạm thời, như tiếng trống, thường có cấu trúc thời gian khác biệt: năng lượng chủ yếu chỉ có mặt trong khoảng thời gian ngắn và được lan rộng qua dải tần số rộng. Với các tín hiệu này, độ phân giải thời gian lớn của phép biểu diễn t/f được ứng dụng có lợi để tách phần tín hiệu tạm thời từ hỗn hợp tín hiệu.

Các sơ đồ mã hóa đối tượng âm thanh hiện thời chỉ đề xuất khả năng thay đổi giới hạn theo sự chọn lọc thời gian-tần số của xử lý SAOC. Chẳng hạn, MPEG SAOC [SAOC] [SAOC1] [SAOC2] được giới hạn ở độ phân giải thời gian-tần số mà có thể thu được bằng cách sử dụng dàn lọc gương vuông góc lai (QMF-Lai) và nhóm tiếp theo của nó vào các dải tham số. Do đó, việc phục hồi đối tượng theo SAOC tiêu chuẩn (MPEG SAOC, như được chuẩn hóa trong [SAOC]) thường phải chịu sự phân giải tần số thô của QMF-Lai dẫn đến sự giao tiếp chéo được điều biến có thể nghe được từ các đối tượng âm thanh khác (ví dụ, các thành phần lạ giao tiếp đôi trong bài diễn văn hoặc các thành phần lạ thô ráp có thể nghe được trong âm nhạc).

Các sơ đồ mã hóa đối tượng âm thanh, như mã hóa tín hiệu âm thanh nổi [BCC] (Binaury Cue Coding – BCC) và mã hóa kết hợp tham số của các nguồn âm thanh [JSC] (Parametric Joint-Coding of Audio Schemes – JSC), cũng được giới hạn đến việc sử dụng một dàn lọc phân giải cố định. Sự lựa chọn thực sự của dàn lọc phân giải cố định hoặc biến đổi luôn bao gồm sự cân bằng được xác định trước đối với khả năng tối ưu giữa các đặc tính thời gian và phổ của sơ đồ mã hóa.

Trong lĩnh vực tách nguồn được thông tin (informed source separation - ISS), đã được đề xuất để làm thích ứng động chiều dài biến đổi thời gian tần số với các đặc tính tín hiệu [ISS7] như đã được biết đến từ các sơ đồ mã hóa âm thanh giác quan, ví dụ, mã hóa âm thanh nâng cao (AAC) [AAC].

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 23003-2:2010.
- [AAC] Bosi, Marina; Brandenburg, Karlheinz; Quackenbush, Schuyler; Fielder, Louis; Akagiri, Kenzo; Fuchs, Hendrik; Dietz, Martin, "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] Andrew Nesbit, Emmanuel Vincent, and Mark 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.

[FB] B. Edler, "Aliasing reduction in subbands of cascaded filterbanks with decimation", Electronic Letters, vol. 28, No. 12, pp. 1104-1106, June 1992.

[MPEG-1] ISO/IEC JTC1/SC29/WG11 MPEG, International Standard ISO/IEC 11172, Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbit/s, 1993.

Bản chất kỹ thuật của sáng chế

Mục đích của sáng chế là đề xuất các khái niệm được cải thiện đối với việc mã hóa đối tượng âm thanh. Mục đích của sáng chế được giải quyết bởi bộ giải mã theo điểm 1 yêu cầu bảo hộ, bộ mã hóa theo điểm 7 yêu cầu bảo hộ, phương pháp giải mã theo điểm 13 yêu cầu bảo hộ, phương pháp để mã hóa theo điểm 14 yêu cầu bảo hộ và vật ghi đọc được bằng máy tính bao gồm chương trình máy tính theo điểm 15 yêu cầu bảo hộ.

Ngược với SAOC trong kỹ thuật hiện nay, các phương án được đề xuất để làm thích ứng động độ phân giải thời gian-tần số với tín hiệu theo cách tương hợp ngược, sao cho

- các dòng bit tham số SAOC bắt nguồn từ bộ mã hóa SAOC chuẩn (MPEG SAOC, như được tiêu chuẩn hóa trong [SAOC]) có thể vẫn được giải mã bằng bộ giải mã được tăng cường với chất lượng giác quan có thể so sánh với chất lượng thu được với bộ giải mã tiêu chuẩn,

- các dòng bit tham số SAOC được tăng cường có thể được giải mã với chất lượng tối ưu có bộ giải mã được tăng cường, và

- các dòng bit tham số SAOC tiêu chuẩn được tăng cường có thể được trộn, ví dụ, trong kịch bản bộ điều khiển đa điểm (Multi-point control unit - MCU), thành một dòng bit chung mà có thể được giải mã với bộ giải mã tiêu chuẩn hoặc được tăng cường.

Đối với các đặc tính nêu trên, sáng chế hữu dụng đề xuất dàn lọc/sự biểu diễn biến đổi chung mà có thể được thích hợp trong độ phân giải thời gian-tần số với việc hỗ trợ giải mã dữ liệu SAOC được tăng cường mới và, đồng thời, ánh xạ tương hợp ngược của dữ liệu SAOC tiêu chuẩn thông thường. Việc hợp nhất dữ liệu SAOC được tăng cường và dữ liệu SAOC tiêu chuẩn có thể đưa ra biểu diễn chung.

Chất lượng giác quan SAOC được tăng cường có thể thu được bằng cách thích ứng độ phân giải thời gian-tần số của dàn lọc hoặc biến đổi mà được sử dụng để ước tính hoặc được sử dụng để tổng hợp đối tượng âm thanh nhằm ám chỉ các đặc tính cụ thể của đối tượng âm thanh đầu vào. Chẳng hạn, nếu đối tượng âm thanh chuẩn dừng trong khoảng thời gian nhất định, ước tính và tổng hợp tham số được thực hiện trên độ phân giải thời gian thô và độ phân giải tần số mịn. Nếu đối tượng âm thanh chứa các phần tạm thời hoặc các phần không ổn định trong khoảng thời gian nhất định, thì việc ước tính và tổng hợp tham số được thực hiện một cách có lợi bằng cách sử dụng độ phân giải thời gian mịn và độ phân giải tần số thô. Theo đó, sự thích ứng động của dàn lọc hoặc biến đổi cho phép

- tính chọn lọc tần số cao trong tách phổ các tín hiệu chuẩn dừng để tránh sự giao tiếp chéo giữa các đối tượng, và
- tính chính xác thời gian cao với đối tượng lúc bắt đầu hoặc các sự kiện tạm thời để tối thiểu hóa các tiếng vang trước và sau.

Đồng thời, chất lượng SAOC thông thường có thể thu được bằng cách ánh xạ dữ liệu SAOC tiêu chuẩn lên mạng lưới thời gian-tần số được tạo ra bằng chuyển đổi thích ứng tín hiệu tương hợp ngược sáng tạo mà tùy thuộc vào thông tin phụ mô tả các đặc trưng tín hiệu của đối tượng.

Có thể giải mã cả hai dữ liệu SAOC tiêu chuẩn và tăng cường bằng cách sử dụng một biến đổi chung cho phép khả năng tương hợp ngược trực tiếp với các ứng dụng bao gồm hỗn hợp dữ liệu SAOC tiêu chuẩn và được tăng cường mới.

Bộ giải mã để tạo ra tín hiệu đầu ra âm thanh bao gồm một hoặc nhiều kênh đầu ra âm thanh từ tín hiệu trộn giảm bao gồm nhiều mẫu trộn giảm miền thời gian được tạo ra. Tín hiệu trộn giảm mã hóa hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh.

Bộ giải mã bao gồm bộ tạo dãy cửa sổ hoặc xác định nhiều cửa sổ phân tích, trong đó mỗi cửa sổ phân tích bao gồm nhiều mẫu trộn giảm miền thời gian của tín hiệu trộn giảm. Mỗi cửa sổ phân tích trong nhiều cửa sổ phân tích có chiều dài cửa sổ chỉ ra số mẫu trộn giảm miền thời gian của cửa sổ phân tích đã nêu. Bộ tạo dãy cửa sổ được tạo cấu hình để xác định nhiều cửa sổ phân tích sao cho chiều dài cửa sổ của mỗi cửa sổ phân tích phụ thuộc vào đặc tính tín hiệu của ít nhất một trong hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh.

Hơn nữa, bộ giải mã bao gồm môđun phân tích t/f để biến đổi nhiều mẫu trộn giảm miền thời gian của mỗi cửa sổ phân tích trong nhiều cửa sổ phân tích từ miền thời gian sang miền thời gian tần số phụ thuộc vào chiều dài cửa sổ của cửa sổ phân tích đã nêu, để thu được sự trộn giảm được biến đổi.

Hơn nữa, bộ giải mã bao gồm bộ giải trộn để giải trộn sự trộn giảm được biến đổi dựa trên thông tin phụ tham số trên hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh để thu được tín hiệu đầu ra âm thanh.

Theo một phương án, bộ tạo dãy cửa sổ có thể được tạo cấu hình để xác định nhiều cửa sổ phân tích, sao cho phần tạm thời, chỉ ra sự thay đổi tín hiệu của ít nhất một trong hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh được mã hóa bởi tín hiệu trộn giảm, được chứa bởi cửa sổ phân tích thứ nhất trong nhiều cửa sổ phân tích và cửa sổ phân tích thứ hai trong nhiều cửa sổ phân tích, trong đó tâm c_k của cửa sổ phân tích thứ nhất được xác định bằng vị trí t của phần tạm thời theo $c_k = t - l_b$, và tâm c_{k+1} của cửa sổ phân tích thứ nhất được xác định bằng vị trí t của phần tạm thời theo $c_{k+1} = t + l_a$, trong đó l_a và l_b là các số.

Theo một phương án, bộ tạo dãy cửa sổ có thể được tạo cấu hình để xác định nhiều cửa sổ phân tích, sao cho phần tạm thời chỉ ra sự thay đổi của ít nhất một trong hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh được mã hóa bằng tín hiệu trộn giảm, được chứa bởi cửa sổ phân tích thứ nhất trong nhiều cửa sổ phân tích, trong đó tâm c_k của cửa sổ phân tích thứ nhất được xác định bằng vị trí t của phần tạm thời theo $c_k = t$, trong đó tâm c_{k-1} của cửa sổ phân tích thứ hai trong nhiều cửa sổ phân tích được xác định bằng vị trí t của phần tạm thời theo $c_{k-1} = t - l_b$, và trong đó tâm c_{k+1} của cửa sổ phân tích thứ ba trong nhiều cửa sổ phân tích được xác định bằng vị trí t của phần tạm thời theo $c_{k+1} = t + l_a$, trong đó l_a và l_b là các số.

Theo một phương án, bộ tạo dãy cửa sổ có thể được tạo cấu hình để xác định nhiều cửa sổ phân tích, sao cho một trong nhiều cửa sổ phân tích bao gồm số lượng các mẫu tín hiệu miền thời gian thứ nhất hoặc số lượng các mẫu tín hiệu miền thời gian thứ hai, trong đó số lượng các mẫu tín hiệu miền thời gian thứ hai lớn hơn số lượng các mẫu tín hiệu miền thời gian thứ nhất, và trong đó mỗi cửa sổ phân tích trong số nhiều cửa sổ phân tích bao gồm số lượng các mẫu tín hiệu miền thời gian thứ nhất khi cửa sổ phân tích đã nêu bao gồm phần tạm thời, chỉ ra sự thay đổi tín hiệu của ít nhất một trong hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh được mã hóa bằng tín hiệu trộn giảm.

Theo một phương án, mô đun phân tích t/f có thể được tạo cấu hình để biến đổi các mẫu trộn giảm miền thời gian của mỗi cửa sổ phân tích từ miền thời gian sang miền tần số- thời gian bằng cách sử dụng dàn lọc QMF và dàn lọc Nyquist, trong đó bộ phân tích t/f được tạo cấu hình để biến đổi nhiều mẫu tín hiệu miền thời gian của mỗi cửa sổ phân tích tùy thuộc vào chiều dài cửa sổ của cửa sổ phân tích đã nêu.

Hơn nữa, bộ mã hóa để mã hóa hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh đầu vào được tạo ra. Mỗi tín hiệu trong hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh đầu vào bao gồm nhiều mẫu tín hiệu miền thời gian. Bộ mã hóa bao gồm bộ dãy cửa sổ để xác định nhiều cửa sổ phân tích. Mỗi cửa sổ phân tích bao gồm nhiều mẫu tín hiệu miền thời gian của một trong các tín hiệu đối tượng âm thanh đầu vào, trong đó mỗi các cửa sổ phân tích có chiều dài cửa sổ chỉ ra số các mẫu tín hiệu miền thời gian của cửa sổ phân tích đã nêu. Bộ dãy cửa sổ được tạo cấu hình để xác

định nhiều cửa sổ phân tích sao cho chiều dài cửa sổ của mỗi cửa sổ phân tích tùy thuộc vào đặc tính tín hiệu của ít nhất một trong hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh đầu vào.

Hơn nữa, bộ mã hóa bao gồm bộ phân tích t/f để biến đổi các mẫu tín hiệu miền thời gian của mỗi cửa sổ phân tích từ miền thời gian sang miền tần số- thời gian để thu được các mẫu tín hiệu được biến đổi. Bộ phân tích t/f có thể được tạo cấu hình để biến đổi nhiều mẫu tín hiệu miền thời gian của mỗi cửa sổ phân tích tùy thuộc vào chiều dài cửa sổ của cửa sổ phân tích đã nêu.

Hơn nữa, bộ mã hóa bao gồm bộ ước tính PSI để xác định thông tin phụ tham số tùy thuộc vào các mẫu tín hiệu được biến đổi.

Theo một phương án, bộ mã hóa có thể còn bao gồm bộ phát hiện tạm thời được tạo cấu hình để xác định nhiều sự khác biệt về mức đối tượng của hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh đầu vào, và được tạo cấu hình để xác định, nếu sự khác biệt giữa mỗi sự khác biệt về mức của đối tượng thứ nhất và mỗi khác biệt về mức của đối tượng thứ hai có lớn hơn giá trị ngưỡng, để xác định mỗi cửa sổ phân tích, nếu cửa sổ phân tích đã nêu bao gồm phần tạm thời chỉ ra sự thay đổi tín hiệu của ít nhất một trong hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh đầu vào.

Theo một phương án, bộ phát hiện tạm thời có thể được tạo cấu hình để sử dụng hàm phát hiện $d(n)$ để xác định nếu sự khác biệt giữa mỗi khác biệt về mức của đối tượng thứ nhất và mỗi khác biệt về mức của đối tượng thứ hai lớn hơn giá trị ngưỡng, trong đó hàm phát hiện $d(n)$ được xác định là:

$$d(n) = \sum_{i,j} |\log(OLD_{i,j}(b, n-1)) - \log(OLD_{i,j}(b, n))|$$

trong đó, n là chỉ số, trong đó i là đối tượng thứ nhất, trong đó j là đối tượng thứ hai, trong đó b là dải tham số. OLD có thể, ví dụ, là sự khác biệt về mức đối tượng.

Theo một phương án, bộ dãy cửa sổ có thể được tạo cấu hình để xác định nhiều cửa sổ phân tích, sao cho phần tạm thời chỉ ra sự thay đổi tín hiệu của ít nhất một trong hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh đầu vào, được chứa bởi cửa sổ phân tích thứ nhất trong nhiều cửa sổ phân tích và cửa sổ phân tích thứ hai trong nhiều cửa sổ phân tích, trong đó tâm c_k của cửa sổ phân tích thứ nhất được xác định bằng vị

trí t của phần tạm thời theo $c_k = t - l_b$, và tâm c_{k+1} của cửa sổ phân tích thứ nhất được xác định bằng vị trí t của phần tạm thời theo $c_{k+1} = t + l_a$, trong đó l_a và l_b là các số.

Theo một phương án, bộ dãy cửa sổ có thể được tạo cấu hình để xác định nhiều cửa sổ phân tích, sao cho phần tạm thời chỉ ra sự thay đổi tín hiệu của ít nhất một trong hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh đầu vào, được chứa bởi cửa sổ phân tích thứ nhất trong nhiều cửa sổ phân tích, trong đó tâm c_k của cửa sổ phân tích thứ nhất được xác định bằng vị trí t của phần tạm thời theo $c_k = t$, trong đó tâm c_{k-1} của cửa sổ phân tích thứ hai trong nhiều cửa sổ phân tích được xác định bằng vị trí t của phần tạm thời theo $c_{k-1} = t - l_b$, và trong đó tâm c_{k+1} của cửa sổ phân tích thứ ba trong nhiều cửa sổ phân tích được xác định bằng vị trí t của phần tạm thời theo $c_{k+1} = t + l_a$, trong đó l_a và l_b là các số.

Theo một phương án, bộ dãy cửa sổ có thể được tạo cấu hình để xác định nhiều cửa sổ phân tích, sao cho một trong nhiều cửa sổ phân tích bao gồm số lượng các mẫu tín hiệu miền thời gian thứ nhất hoặc số lượng các mẫu tín hiệu miền thời gian thứ hai, trong đó số lượng các mẫu tín hiệu miền thời gian thứ hai lớn hơn số lượng các mẫu tín hiệu miền thời gian thứ nhất, và trong đó mỗi cửa sổ phân tích trong nhiều cửa sổ phân tích bao gồm số lượng các mẫu tín hiệu miền thời gian thứ nhất khi cửa sổ phân tích đã nêu bao gồm phần tạm thời chỉ ra sự thay đổi tín hiệu của ít nhất một trong hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh đầu vào.

Theo một phương án, bộ phân tích t/f có thể được tạo cấu hình để biến đổi các mẫu tín hiệu miền thời gian của mỗi cửa sổ phân tích từ miền thời gian sang miền tần số- thời gian bằng cách sử dụng dàn lọc QMF và dàn lọc Nyquist, trong đó bộ phân tích t/f có thể được tạo cấu hình để biến đổi nhiều mẫu tín hiệu miền thời gian của mỗi cửa sổ phân tích tùy thuộc vào chiều dài cửa sổ của cửa sổ phân tích đã nêu.

Hơn nữa, bộ giải mã để tạo ra tín hiệu đầu ra âm thanh bao gồm một hoặc nhiều kênh đầu ra âm thanh từ tín hiệu trộn giảm bao gồm nhiều mẫu trộn giảm miền thời gian được tạo ra. Tín hiệu trộn giảm mã hóa hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh. Bộ giải mã bao gồm môđun con phân tích thứ nhất để biến đổi nhiều mẫu trộn giảm miền thời gian để thu được nhiều dải con bao gồm nhiều mẫu dải con. Hơn nữa, bộ giải mã bao gồm bộ tạo dãy cửa sổ để xác định nhiều cửa sổ phân tích, trong

đó mỗi cửa sổ phân tích bao gồm nhiều mẫu dải con của một trong nhiều dải con, trong đó mỗi cửa sổ phân tích trong nhiều cửa sổ phân tích có chiều dài cửa sổ chỉ ra số mẫu dải con của cửa sổ phân tích đã nêu, trong đó bộ tạo dãy cửa sổ được tạo cấu hình để xác định nhiều cửa sổ phân tích sao cho chiều dài cửa sổ của mỗi cửa sổ phân tích tùy thuộc vào đặc tính tín hiệu của ít nhất một trong hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh. Hơn nữa, bộ giải mã bao gồm môđun phân tích thứ hai để biến đổi nhiều mẫu dải con của mỗi cửa sổ phân tích trong nhiều cửa sổ phân tích tùy thuộc vào chiều dài cửa sổ của cửa sổ phân tích đã nêu để thu được sự trộn giảm được biến đổi. Hơn nữa, bộ giải mã bao gồm bộ giải trộn để giải trộn sự trộn giảm được biến đổi dựa trên thông tin phụ tham số trên hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh để thu được tín hiệu đầu ra âm thanh.

Hơn nữa, bộ mã hóa để mã hóa hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh đầu vào được đề xuất. Một trong số hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh đầu vào bao gồm nhiều mẫu tín hiệu miền thời gian. Bộ mã hóa bao gồm môđun con phân tích thứ nhất để biến đổi nhiều mẫu tín hiệu miền thời gian để thu được nhiều dải con bao gồm nhiều mẫu dải con. Hơn nữa, bộ mã hóa bao gồm bộ dãy cửa sổ để xác định nhiều cửa sổ phân tích, trong đó mỗi cửa sổ phân tích bao gồm nhiều mẫu dải con của một trong nhiều dải con, trong đó mỗi cửa sổ phân tích có chiều dài cửa sổ chỉ ra số mẫu dải con của cửa sổ phân tích đã nêu, trong đó bộ dãy cửa sổ được tạo cấu hình để xác định nhiều cửa sổ phân tích sao cho chiều dài cửa sổ của mỗi cửa sổ phân tích tùy thuộc vào đặc tính tín hiệu của ít nhất một trong hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh đầu vào. Hơn nữa, bộ mã hóa bao gồm môđun phân tích thứ hai để biến đổi nhiều mẫu dải con của mỗi cửa sổ phân tích trong nhiều cửa sổ phân tích tùy thuộc vào chiều dài cửa sổ của cửa sổ phân tích này để thu được các mẫu tín hiệu được biến đổi. Hơn nữa, bộ mã hóa bao gồm bộ ước tính PSI để xác định thông tin phụ tham số tùy thuộc vào các mẫu tín hiệu được biến đổi.

Hơn nữa, bộ giải mã để tạo ra tín hiệu đầu ra âm thanh bao gồm một hoặc nhiều kênh đầu ra âm thanh từ tín hiệu trộn giảm được đề xuất. Tín hiệu trộn giảm mã hóa một hoặc nhiều tín hiệu đối tượng âm thanh. Bộ giải mã bao gồm bộ điều khiển để cài đặt chỉ dẫn hoạt động sang trạng thái hoạt động tùy thuộc vào đặc tính tín hiệu của ít nhất một trong số một hoặc nhiều tín hiệu đối tượng âm thanh. Hơn nữa, bộ giải mã

bao gồm môđun phân tích thứ nhất để biến đổi tín hiệu trộn giảm để thu được sự trộn giảm được biến đổi thứ nhất bao gồm nhiều kênh dải con thứ nhất. Hơn nữa, bộ giải mã bao gồm môđun phân tích thứ hai để tạo ra, khi chỉ dẫn hoạt động được cài đặt sang trạng thái hoạt động, sự trộn giảm được biến đổi thứ hai bằng cách biến đổi ít nhất một trong các kênh dải con thứ nhất để thu được nhiều kênh dải con thứ hai, trong đó sự trộn giảm được biến đổi thứ hai bao gồm các kênh dải con thứ nhất mà chưa được biến đổi bằng môđun phân tích thứ hai và các kênh dải con thứ hai. Hơn nữa, bộ giải mã bao gồm bộ giải trộn, trong đó bộ giải trộn được tạo cấu hình để giải trộn sự trộn giảm được biến đổi thứ hai, khi chỉ dẫn hoạt động được cài đặt sang trạng thái hoạt động, dựa trên thông tin phụ tham số trên một hoặc nhiều tín hiệu đối tượng âm thanh để thu được tín hiệu đầu ra âm thanh, và giải trộn sự trộn giảm được biến đổi thứ nhất, khi chỉ dẫn hoạt động không được cài đặt sang trạng thái hoạt động, dựa trên thông tin phụ tham số trên một hoặc nhiều tín hiệu đối tượng âm thanh để thu được tín hiệu đầu ra âm thanh.

Hơn nữa, bộ mã hóa để mã hóa tín hiệu đối tượng âm thanh đầu vào được đề xuất. Bộ mã hóa bao gồm bộ điều khiển để cài đặt chỉ dẫn hoạt động sang trạng thái hoạt động tùy thuộc vào đặc tính tín hiệu của tín hiệu đối tượng âm thanh đầu vào. Hơn nữa, bộ mã hóa bao gồm môđun phân tích thứ nhất để biến đổi tín hiệu đối tượng âm thanh đầu vào để thu được tín hiệu đối tượng âm thanh được biến đổi thứ nhất, trong đó tín hiệu đối tượng âm thanh được biến đổi thứ nhất bao gồm nhiều kênh dải con thứ nhất. Hơn nữa, bộ mã hóa bao gồm môđun phân tích thứ hai để tạo ra, khi chỉ dẫn hoạt động được cài đặt sang trạng thái hoạt động, tín hiệu đối tượng âm thanh được biến đổi thứ hai bằng cách biến đổi ít nhất một trong nhiều kênh dải con thứ nhất để thu được nhiều kênh dải con thứ hai, trong đó tín hiệu đối tượng âm thanh được biến đổi thứ hai bao gồm các kênh dải con thứ nhất mà chưa được biến đổi bằng môđun phân tích thứ hai và các kênh dải con thứ hai. Hơn nữa, bộ mã hóa bao gồm bộ ước tính PSI, trong đó bộ ước tính PSI được tạo cấu hình để xác định thông tin phụ tham số dựa trên tín hiệu đối tượng âm thanh được biến đổi thứ hai, khi chỉ dẫn hoạt động được cài đặt sang trạng thái hoạt động, và xác định thông tin phụ tham số dựa trên tín hiệu đối tượng âm thanh được biến đổi thứ nhất, khi chỉ dẫn hoạt động được cài đặt sang trạng thái hoạt động.

Hơn nữa, phương pháp giải mã để tạo ra tín hiệu đầu ra âm thanh bao gồm một hoặc nhiều kênh đầu ra âm thanh từ tín hiệu trộn giảm bao gồm nhiều mẫu trộn giảm miền thời gian được đề xuất. Tín hiệu trộn giảm mã hóa hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh. Phương pháp này bao gồm các bước:

- xác định nhiều cửa sổ phân tích, trong đó mỗi cửa sổ phân tích bao gồm nhiều mẫu trộn giảm miền thời gian của tín hiệu trộn giảm, trong đó mỗi cửa sổ phân tích trong nhiều cửa sổ phân tích có chiều dài cửa sổ chỉ ra số các mẫu trộn giảm miền thời gian của cửa sổ phân tích này, trong đó việc xác định nhiều cửa sổ phân tích được tiến hành sao cho chiều dài cửa sổ của mỗi cửa sổ phân tích tùy thuộc vào đặc tính tín hiệu của ít nhất một trong hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh.
- biến đổi nhiều mẫu trộn giảm miền thời gian của mỗi cửa sổ phân tích trong nhiều cửa sổ phân tích từ miền thời gian sang miền thời gian tần số tùy thuộc vào chiều dài cửa sổ của cửa sổ phân tích này, để thu được sự trộn giảm được biến đổi, và
- giải trộn sự trộn giảm được biến đổi dựa trên thông tin phụ tham số trên hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh để thu được tín hiệu đầu ra âm thanh,

Hơn nữa, phương pháp mã hóa hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh đầu vào được đề xuất. Một trong hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh đầu vào bao gồm nhiều mẫu tín hiệu miền thời gian. Phương pháp này bao gồm các bước:

- xác định nhiều cửa sổ phân tích, trong đó mỗi cửa sổ phân tích bao gồm nhiều mẫu tín hiệu miền thời gian của mỗi tín hiệu đối tượng âm thanh đầu vào, trong đó mỗi cửa sổ phân tích có chiều dài cửa sổ chỉ ra số các mẫu tín hiệu miền thời gian của cửa sổ phân tích này, trong đó việc xác định nhiều cửa sổ phân tích được tiến hành sao cho chiều dài cửa sổ của mỗi cửa sổ phân tích tùy thuộc vào đặc tính tín hiệu của ít nhất một trong hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh đầu vào.
- biến đổi các mẫu tín hiệu miền thời gian của mỗi cửa sổ phân tích từ miền thời gian sang miền thời gian tần số để thu được các mẫu tín hiệu được biến đổi, trong đó việc biến đổi nhiều mẫu tín hiệu miền thời gian của mỗi cửa sổ phân tích tùy thuộc vào chiều dài cửa sổ của cửa sổ phân tích này. Và:

- xác định thông tin phụ tham số tùy thuộc vào các mẫu tín hiệu được biến đổi.

Hơn nữa, phương pháp giải mã bằng cách tạo ra tín hiệu đầu ra âm thanh bao gồm một hoặc nhiều kênh đầu ra âm thanh từ tín hiệu trộn giảm bao gồm nhiều mẫu trộn giảm miền thời gian, trong đó tín hiệu trộn giảm mã hóa hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh, được đề xuất. Phương pháp này bao gồm các bước:

- biến đổi nhiều mẫu trộn giảm miền thời gian để thu được nhiều dải con bao gồm nhiều mẫu dải con.
- xác định nhiều cửa sổ phân tích, trong đó mỗi cửa sổ phân tích bao gồm nhiều mẫu dải con của một trong nhiều dải con, trong đó mỗi cửa sổ phân tích của nhiều cửa sổ phân tích có chiều dài cửa sổ chỉ ra số mẫu dải con của cửa sổ phân tích này, trong đó việc xác định nhiều cửa sổ phân tích được tiến hành sao cho chiều dài cửa sổ của mỗi cửa sổ phân tích tùy thuộc vào đặc tính tín hiệu của ít nhất một trong hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh.
- biến đổi nhiều mẫu dải con của mỗi cửa sổ phân tích trong nhiều cửa sổ phân tích tùy thuộc vào chiều dài cửa sổ của cửa sổ phân tích này để thu được sự trộn giảm được biến đổi. Và:
- giải trộn sự trộn giảm được biến đổi dựa trên thông tin phụ tham số trên hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh để thu được tín hiệu đầu ra âm thanh.

Hơn nữa, phương pháp mã hóa hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh đầu vào, trong đó một trong hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh đầu vào bao gồm nhiều mẫu tín hiệu miền thời gian, được đề xuất. Phương pháp này bao gồm các bước:

- biến đổi nhiều mẫu tín hiệu miền thời gian để thu được nhiều dải con bao gồm nhiều mẫu dải con.
- xác định nhiều cửa sổ phân tích, trong đó mỗi cửa sổ phân tích bao gồm nhiều mẫu dải con của một trong nhiều dải con, trong đó mỗi cửa sổ phân tích có chiều dài cửa sổ chỉ ra số mẫu dải con của cửa sổ phân tích này, trong đó việc xác định nhiều cửa sổ phân tích được tiến hành sao cho chiều dài cửa sổ của mỗi cửa sổ phân tích tùy thuộc

vào đặc tính tín hiệu của ít nhất một trong hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh đầu vào.

- chuyển đổi nhiều mẫu dải con của mỗi cửa sổ phân tích trong nhiều cửa sổ phân tích tùy thuộc vào chiều dài cửa sổ của cửa sổ phân tích này để thu được các mẫu tín hiệu được biến đổi. Và

- xác định thông tin phụ tham số tùy thuộc vào các mẫu tín hiệu được biến đổi.

Hơn nữa, phương pháp giải mã bằng cách tạo ra tín hiệu đầu ra âm thanh bao gồm một hoặc nhiều kênh đầu ra âm thanh từ tín hiệu trộn giảm, trong đó tín hiệu trộn giảm mã hóa hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh, được đề xuất. Phương pháp này bao gồm các bước:

- cài đặt chỉ dẫn hoạt động sang trạng thái hoạt động tùy thuộc vào đặc tính tín hiệu của ít nhất một trong hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh.

- biến đổi tín hiệu trộn giảm để thu được sự trộn giảm được biến đổi thứ nhất bao gồm nhiều kênh dải con thứ nhất.

- tạo ra, khi chỉ dẫn hoạt động được cài đặt sang trạng thái hoạt động, sự trộn giảm được biến đổi thứ hai bằng cách biến đổi ít nhất mỗi kênh dải con thứ nhất để thu được nhiều kênh dải con thứ hai, trong đó sự trộn giảm được biến đổi thứ hai bao gồm các kênh dải con thứ nhất mà chưa được biến đổi bằng môđun phân tích thứ hai và các kênh dải con thứ hai. Và:

- giải trộn sự trộn giảm được biến đổi thứ hai, khi chỉ dẫn hoạt động được cài đặt sang trạng thái hoạt động, dựa trên thông tin phụ tham số trên hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh để thu được tín hiệu đầu ra âm thanh, và giải trộn sự trộn giảm được biến đổi thứ nhất, khi chỉ dẫn hoạt động không được cài đặt sang trạng thái hoạt động, dựa trên thông tin phụ tham số trên hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh để thu được tín hiệu đầu ra âm thanh.

Hơn nữa, phương pháp mã hóa hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh đầu vào được đề xuất. Phương pháp này bao gồm các bước:

- cài đặt chỉ dẫn hoạt động sang trạng thái hoạt động tùy thuộc vào đặc tính tín hiệu của ít nhất một trong hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh đầu vào.

- biến đổi mỗi tín hiệu đối tượng âm thanh đầu vào để thu được tín hiệu đối tượng âm thanh được biến đổi thứ nhất của tín hiệu đối tượng âm thanh đầu vào này, trong đó tín hiệu đối tượng âm thanh được biến đổi thứ nhất bao gồm nhiều kênh dải con thứ nhất.

- tạo ra cho mỗi tín hiệu đối tượng âm thanh đầu vào, khi chỉ dẫn hoạt động được cài đặt sang trạng thái hoạt động, tín hiệu đối tượng âm thanh được biến đổi thứ hai bằng cách biến đổi ít nhất mỗi kênh dải con thứ nhất của tín hiệu đối tượng âm thanh được biến đổi thứ nhất của tín hiệu đối tượng âm thanh đầu vào này để thu được nhiều kênh dải con thứ hai, trong đó sự trộn giảm được biến đổi thứ hai bao gồm các kênh dải con thứ nhất mà chưa được biến đổi bằng môđun phân tích thứ hai và các kênh dải con thứ hai. Và:

- xác định thông tin phụ tham số dựa trên tín hiệu đối tượng âm thanh được biến đổi thứ hai của mỗi tín hiệu đối tượng âm thanh đầu vào, khi chỉ dẫn hoạt động được cài đặt sang trạng thái hoạt động, và xác định thông tin phụ tham số dựa trên tín hiệu đối tượng âm thanh được biến đổi thứ nhất của mỗi tín hiệu đối tượng âm thanh đầu vào, khi chỉ dẫn hoạt động không được cài đặt sang trạng thái hoạt động.

Hơn nữa, vật ghi đọc được bằng máy tính bao gồm chương trình máy tính để thực hiện mỗi phương pháp nêu trên khi được thực hiện trên máy tính hoặc bộ xử lý tín hiệu được đề xuất.

Mô tả vắn tắt các hình vẽ

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

Sau đây, các phương án theo 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 một phương án,

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

Fig.1c minh họa bộ giải mã theo phương án khác,

Fig.2a minh họa bộ mã hóa để mã hóa các tín hiệu đối tượng âm thanh đầu vào theo một phương án,

Fig.2b minh họa bộ mã hóa để mã hóa các tín hiệu đối tượng âm thanh đầu vào theo phương án khác,

Fig.2c minh họa bộ mã hóa để mã hóa các tín hiệu đối tượng âm thanh đầu vào theo phương án khác,

Fig.3 thể hiện sơ đồ khối theo tiến trình tổng quan khái niệm của hệ thống SAOC,

Fig.4 thể hiện sơ đồ theo tiến trình và minh họa của phép biểu diễn phổ thời gian của tín hiệu âm thanh một kênh,

Fig.5 thể hiện sơ đồ khối theo tiến trình của việc tính lựa chọn thời gian-tần số của thông tin phụ nằm trong bộ mã hóa SAOC,

Fig.6 mô tả sơ đồ khối của bộ giải mã SAOC được tăng cường theo một phương án, minh họa việc giải mã các dòng bit SAOC tiêu chuẩn,

Fig.7 mô tả sơ đồ khối của bộ giải mã theo một phương án,

Fig.8 minh họa sơ đồ khối của bộ mã hóa theo phương án cụ thể thực hiện con đường tham số của bộ mã hóa,

Fig.9 minh họa sự thích ứng của dãy cửa sổ thông thường để chứa điểm giao với cửa sổ lúc tạm thời,

Fig.10 minh họa sơ đồ biến đổi khối cách ly tạm thời theo một phương án,

Fig.11 minh họa tín hiệu với phần tạm thời dãy cửa sổ giống như AAC thu được theo một phương án,

Fig.12 minh họa việc lọc lại QMF mở rộng,

Fig.13 minh họa ví dụ mà các cửa sổ ngắn được sử dụng để biến đổi,

Fig.14 minh họa ví dụ mà các cửa sổ dài hơn được sử dụng để biến đổi trong ví dụ của Fig.13.

Fig.15 minh họa ví dụ, mà độ phân giải tần số cao và độ phân giải thời gian thấp được nhận biết,

Fig.16 minh họa ví dụ, trong đó độ phân giải thời gian cao và phân giải thời gian thấp được nhận biết,

Fig.17 minh họa ví dụ thứ nhất, trong đó độ phân giải thời gian trung gian và độ phân giải tần số trung gian được nhận biết, và

Fig.18 minh họa ví dụ thứ nhất, trong đó độ phân giải thời gian trung gian và độ phân giải tần số trung gian được nhận biết.

Mô tả chi tiết sáng chế

Trước khi mô tả các phương án theo sáng chế, các tình trạng của hệ thống SAOC trong kỹ thuật hiện nay được đề xuất.

Fig.3 thể hiện sự sắp xếp của bộ mã hóa SAOC 10 và bộ giải mã SAOC 12. Bộ mã hóa SAOC 10 nhận đầu vào N đối tượng, nghĩa là các tín hiệu âm thanh s_1 đến s_N . Cụ thể là, bộ mã hóa 10 bao gồm bộ trộn giảm 16 mà nhận các tín hiệu âm thanh s_1 đến s_N và trộn giảm chúng thành tín hiệu trộn giảm 18. Theo cách khác, sự trộn giảm có thể được tạo ra bên ngoài (“trộn giảm nghệ thuật”) và hệ thống ước tính thông tin phụ bổ sung để làm cho sự trộn giảm được tạo ra thích hợp với trộn giảm được tính toán. Trên Fig.3, tín hiệu trộn giảm được thể hiện là tín hiệu P kênh. Do đó, 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 giảm lập thể, các kênh tín hiệu trộn giảm 18 được ký hiệu là $L0$ và $R0$, trong trường hợp trộn giảm đơn chỉ đơn giản ký hiệu là $L0$. Để cho phép bộ giải mã SAOC 12 phục hồi các đối tượng riêng biệt s_1 đến s_N , bộ ước tính thông tin phụ 17 tạo ra bộ giải mã SAOC 12 với thông tin phụ bao gồm các tham số SAOC. Ví dụ, trong trường hợp trộn giảm lập thể, các tham số SAOC bao gồm các khác biệt về mức của đối tượng (OLD), các tương quan giữa các đối tượng (IOC) (các tham số tương quan giữa các đối tượng), các giá trị tăng ích trộn giảm (DMG) và các khác biệt về mức kênh trộn giảm (DCLD). Thông tin phụ 20, bao gồm các tham số SAOC, cùng với tín hiệu trộn giảm 18, tạo thành dòng dữ liệu đầu ra SAOC được nhận bởi bộ giải mã SAOC 12.

Bộ giải mã SAOC 12 bao gồm bộ trộn tăng nhận tín hiệu trộn giảm 18 cũng như thông tin phụ 20 để phục hồi và kết xuất các tín hiệu âm thanh s_1 và s_N lên trên tập

hợp được chọn lọc bởi người dùng bất kỳ của các kênh y_I đến y_M , với sự kết xuất được quy định bằng cách kết xuất thông tin 26 được nhập vào bộ giải mã SAOC 12.

Các tín hiệu âm thanh s_I đến s_N có thể được nhập vào 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 dẫn vào bộ mã hóa 10 trong miền thời gian, như PCM được mã hóa, bộ mã hóa 10 có thể sử dụng dàn lọc, như dàn QMF lai, để truyền tín hiệu vào miền phổ, mà tại đó các tín hiệu âm thanh được biểu diễn theo một vài dải con được kết hợp với các phần phổ khác nhau, tại độ phân giải dàn lọc cụ thể. Nếu các tín hiệu âm thanh s_I đến s_N đã theo biểu diễn được mong đợi bởi bộ mã hóa 10, thì chúng không phải thực hiện sự phân tích phổ.

Fig.4 thể hiện tín hiệu âm thanh trong miền phổ vừa nêu. Như có thể thấy, tín hiệu âm thanh được biểu diễn làm nhiều tín hiệu dải con. Mỗi tín hiệu dải con 30_I đến 30_K bao gồm thứ tự thời gian của các giá trị dải con được chỉ ra bởi các hộp nhỏ 32. Như có thể thấy, các giá trị dải con 32 của các tín hiệu dải con 30_I đến 30_K được đồng bộ hóa với nhau theo thời gian sao cho cho, với mỗi khe thời gian dàn lọc liên tiếp 30_I đến 30_K bao gồm chính xác một giá trị dải con 32. Như được minh họa bằng trục tần số 36, các tín hiệu dải con 30_I đến 30_K được kết hợp với các vùng tần số khác nhau, và như được minh họa bằng trục thời gian 38, các khe thời gian của dàn lọc 34 được sắp xếp liên tiếp theo thời gian.

Như được kể đến trên đây, bộ tách thông tin phụ 17 của Fig.3 tính các tham số SAOC từ các tín hiệu âm thanh đầu vào s_I đến s_N . Theo tiêu chuẩn SAOC được thực hiện hiện thời, bộ mã hóa 10 thực hiện việc tính toán này theo độ phân giải thời gian/tần số mà có thể được tăng so với độ phân giải thời gian/tần số ban đầu như được xác định bởi các khe thời gian dàn lọc 34 và sự phân tích dải con, bằng lượng nhất định, với lượng nhất định này được truyền tín hiệu đến phía bộ giải mã nằm trong thông tin phụ 20. Các nhóm của các khe thời gian 34 có thể tạo thành khung SAOC 41. Cả số dải tham số nằm trong khung SAOC 41 được vận chuyển 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.4 bằng các đường nét đứt 42. Trên Fig.4, các dải tham số được phân chia theo cùng cách trong các khung SAOC được mô tả khác nhau 41 sao cho sự sắp xếp thông

thường của các ô thời gian/tần số thu được. Nói chung, tuy nhiên, các dải tham số có thể khác nhau từ một khung SAOC 41 đến khung tiếp theo, tùy thuộc vào các nhu cầu khác nhau với độ phân giải phổ trong các khung SAOC tương ứng 41. Hơn nữa, chiều dài của các khung SAOC 41 cũng có thể khác nhau. Do đó, sự sắp xếp các ô thời gian/tần số có thể không theo thông thường. Tuy nhiên, các ô thời gian/tần số trong khung SAOC cụ thể 41 thường cùng khoảng thời gian và được sắp xếp theo hướng thời gian, nghĩa là tất cả các ô t/f trong khung SAOC 41 bắt đầu ở lúc bắt đầu của khung SAOC được đưa ra 41 và kết thúc ở lúc kết thúc của khung SAOC 41.

Bộ tách thông tin phụ 17 được mô tả trên Fig.3 tính các tham số SAOC theo các công thức sau đây. Cụ thể là, bộ tách thông tin phụ 17 tính các khác biệt về mức của đối tượng với mỗi đối tượng i là

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

trong đó, tổng và các chỉ số n và k lần lượt đi qua tất cả các chỉ số thời gian 34, và tất cả các chỉ số phổ 30 mà thuộc về ô thời gian/tần số nhất định 42, được tham chiếu bằng các chỉ số l với khung SAOC (hoặc khe thời gian xử lý) và m với dải tham số. Do đó, năng lượng của tất cả các giá trị dải con x_i của tín hiệu âm thanh hoặc đối tượng i được tính tổng và thông thường hóa đến giá trị năng lượng cao nhất của ô đó trong số tất cả các đối tượng hoặc tín hiệu âm thanh. $x_i^{n,k*}$ chỉ số liên hợp phức của $x_i^{n,k}$.

Hơn nữa, bộ tách thông tin phụ SAOC 17 có thể tính số đo tương tự của các ô thời gian/tần số tương ứng của các cặp của các đối tượng đầu vào khác nhau s_I đến s_N . Mặc dù bộ tách thông tin phụ SAOC 17 có thể tính số đo tương tự giữa tất cả các cặp của đối tượng đầu vào s_I đến s_N , bộ tách thông tin phụ 17 cũng có thể ngăn cản tín hiệu của các số đo tương tự hoặc giới hạn tính toán của các số đo tương tự với các đối tượng âm thanh s_I đến s_N tạo thành các kênh bên trái hoặc bên phải của kênh stereo

chung. Trong trường hợp bất kỳ, số đo tương tự được gọi là tham số tương quan chéo giữa các đối tượng $IOC_{i,j}^{l,m}$. Việc tính toán như sau:

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

Với các chỉ số nữa n và k đi qua tất cả các giá trị dải con thuộc về ô thời gian/tần số nhất định 42, i và j là cặp đối tượng âm thanh nhất định s_l đến s_N , và $\text{Re}\{ \}$ là thao tác loại bỏ phần tương tự của lập luận phức tạp.

Bộ trộn giảm 16 của Fig.3 trộn giảm các đối tượng s_l đến s_N bằng cách sử dụng các thừa số tăng ích được áp dụng với mỗi đối tượng s_l đến s_N . Cụ thể là, hệ số tăng ích d_i được áp dụng với đối tượng i và tiếp đó là tất cả các đối tượng trọng số s_l đến s_N được tính tổng để thu được tín hiệu trộn giảm mono, mà được minh họa trên Fig.3 nếu $P=1$. Trong trường hợp ví dụ khác của tín hiệu trộn giảm hai kênh, được mô tả trên Fig.3 nếu $P=2$, thừa số tăng ích $d_{1,i}$ được áp dụng với đối tượng i và tiếp đó tất cả các đối tượng được khuếch đại tăng ích được tính tổng để thu được kênh trộn giảm bên trái $L0$, và các thừa số tăng ích $d_{2,i}$ được áp dụng với đối tượng i và tiếp đó là các đối tượng được khuếch đại tăng ích được tính tổng để thu được kênh trộn giảm bên phải $R0$. Việc xử lý tương tự trên đây là được áp dụng trường hợp của trộn giảm đa kênh ($P>2$).

Mệnh lệnh trộn giảm này được truyền tín hiệu đến phía bộ giải mã bằng cả độ tăng ích trộn giảm DMG_i và, trường hợp tín hiệu trộn giảm stereo, các khác biệt về mức kênh trộn giảm $DCLD_i$.

Độ tăng ích trộn giảm được tính 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 như 10^{-9} .

Với các DCLD, công thức sau đây áp dụng:

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

Theo kiểu thông thường, bộ trộn giảm 16 lần lượt tạo ra tín hiệu trộn giảm theo:

$$(L0) = (d_i) \begin{pmatrix} s_1 \\ \vdots \\ s_N \end{pmatrix}$$

với trộn giảm đơn, hoặc

$$\begin{pmatrix} L0 \\ R0 \end{pmatrix} = \begin{pmatrix} d_{1,i} \\ d_{2,i} \end{pmatrix} \begin{pmatrix} s_1 \\ \vdots \\ s_N \end{pmatrix}$$

với trộn giảm lập thể.

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 *d*. Theo đó, lưu ý rằng *d* có thể khác nhau theo thời gian và tần số.

Do đó, theo kiểu thông thường, bộ trộn giảm 16 trộn tất cả các đối tượng *s₁* đến *s_N* mà không có sự ưu tiên, nghĩa là với thao tác tất cả các đối tượng *s₁* đến *s_N* ngang nhau.

Tại phía bộ giải mã, bộ trộn tăng thực hiện sự nghịch đảo thủ tục trộn giảm và thực hiện “thông tin kết xuất” 26 được biểu diễn bởi ma trận **R** (trong ngữ cảnh này đôi khi gọi là **A**) trong một bước tính toán, cụ thể là, trong trường hợp trộn giảm hai kênh

$$\begin{pmatrix} \hat{y}_1 \\ \vdots \\ \hat{y}_M \end{pmatrix} = \mathbf{RED}^* (\mathbf{DED}^*)^{-1} \begin{pmatrix} L0 \\ R0 \end{pmatrix},$$

trong đó ma trận **E** là hàm của các tham số *OLD* và *IOC*, và ma trận **D** chứa các hệ số trộn giảm là

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

Ma trận E là ma trận hiệp phương sai được ước tính của đối tượng âm thanh s_1 đến s_N . Trong các phương án SAOC hiện tại, việc tính toán ma trận hiệp phương sai được ước tính E thường được thực hiện với độ phân giải phổ /thời gian của các tham số SAOC, nghĩa là, với mỗi (l,m) , sao cho ma trận hiệp phương sai được ước tính có thể được viết là $E^{l,m}$. Ma trận hiệp phương sai được ước tính $E^{l,m}$ là kích cỡ là $N \times N$ có các hệ số được xác định là:

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

Do đó, ma trận $E^{l,m}$ với

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

Có cùng với các khác biệt về mức của đối tượng, nghĩa là, $e_{i,j}^{l,m} = OLD_i^{l,m}$ với $i=j$, do $OLD_i^{l,m} = OLD_j^{l,m}$ và $IOC_{i,j}^{l,m} = 1$ với $i=j$. Ngoài ma trận hiệp phương sai được ước tính E đường chéo của nó lần lượt có các hệ số ma trận biểu diễn trung bình hình học của các khác biệt về mức của đối tượng của các đối tượng i và j , được tạo trọng số với số đo tương quan chéo giữa các đối tượng $IOC_{i,j}^{l,m}$.

Fig.5 hiển thị một nguyên lý có thể có của phương án về ví dụ của bộ ước tính thông tin phụ (SIE) làm một phần của bộ mã hóa SAOC 10. Bộ mã hóa SAOC 10 bao gồm bộ trộn 16 và bộ ước tính thông tin phụ (SIE) 17. SIE theo quan niệm bao gồm hai môđun: Một môđun 45 để tính thời gian ngắn dựa trên biểu diễn t/f (ví dụ, STFT hoặc QMF) của mỗi tín hiệu. Biểu diễn t/f trong thời gian ngắn được tính được dẫn vào môđun thứ hai 46, môđun ước tính thông tin phụ chọn lọc với t/f (t/f-SIE). Môđun t/f-SIE 46 tính thông tin phụ với mỗi ô t/f. Trong các phương án SAOC hiện tại, biến đổi thời gian/tần số được cố định và giống nhau với tất cả các đối tượng âm thanh s_1 đến s_N . Hơn nữa, các tham số SAOC được xác định qua các khung SAOC là giống nhau với tất cả các đối tượng âm thanh và có cùng độ phân giải thời gian/tần số với tất cả các đối tượng âm thanh s_1 đến s_N , do đó loại bỏ các nhu cầu cụ thể với đối tượng với độ phân giải thời gian tốt ở một số trường hợp hoặc độ phân giải phổ tốt ở các trường hợp khác.

Sau đây, các phương án theo sáng chế được mô tả.

Fig.1a minh họa bộ giải mã để tạo ra tín hiệu đầu ra âm thanh bao gồm một hoặc nhiều kênh đầu ra âm thanh từ tín hiệu trộn giảm bao gồm nhiều mẫu trộn giảm miền thời gian theo một phương án. Tín hiệu trộn giảm mã hóa hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh.

Bộ giải mã bao gồm bộ tạo dãy cửa sổ 134 để xác định nhiều cửa sổ phân tích (ví dụ, dựa trên thông tin phụ tham số, ví dụ, các khác biệt về mức của đối tượng), trong đó mỗi cửa sổ phân tích bao gồm nhiều mẫu trộn giảm miền thời gian của tín hiệu trộn giảm. Mỗi cửa sổ phân tích trong nhiều cửa sổ phân tích có chiều dài cửa sổ chỉ ra số các mẫu trộn giảm miền thời gian của cửa sổ phân tích này. Bộ tạo dãy cửa sổ 134 được tạo cấu hình để xác định nhiều cửa sổ phân tích sao cho chiều dài cửa sổ trong mỗi cửa sổ phân tích tùy thuộc vào đặc tính tín hiệu của ít nhất một trong hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh. Ví dụ, chiều dài cửa sổ có thể tùy thuộc vào nếu cửa sổ phân tích này bao gồm phần tạm thời chỉ ra sự thay đổi tín hiệu của ít nhất một trong hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh mà được mã hóa bởi tín hiệu trộn giảm.

Để xác định nhiều cửa sổ phân tích, bộ tạo dãy cửa sổ 134 có thể, ví dụ, phân tích thông tin phụ tham số, ví dụ, các khác biệt về mức của đối tượng được truyền tải so với hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh, để xác định chiều dài cửa sổ của các cửa sổ phân tích, sao cho chiều dài cửa sổ của mỗi cửa sổ phân tích tùy thuộc vào đặc tính tín hiệu của ít nhất một trong hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh. Hoặc, ví dụ, để xác định nhiều cửa sổ phân tích, bộ tạo dãy cửa sổ 134 có thể phân tích hình dạng của cửa sổ hoặc bản thân các cửa sổ phân tích, trong đó các hình dạng cửa sổ hoặc các cửa sổ phân tích có thể, ví dụ, được truyền tải trong dòng bit từ bộ mã hóa đến bộ giải mã, và trong đó chiều dài cửa sổ của mỗi cửa sổ phân tích tùy thuộc vào đặc tính tín hiệu của ít nhất một trong hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh.

Hơn nữa, bộ giải mã bao gồm môđun phân tích t/f 135 để biến đổi nhiều mẫu trộn giảm miền thời gian của mỗi cửa sổ phân tích trong nhiều cửa sổ phân tích từ

miền thời gian sang miền thời gian tần số tùy thuộc vào chiều dài cửa sổ của cửa sổ phân tích này, để thu được sự trộn giảm được biến đổi.

Hơn nữa, bộ giải mã bao gồm bộ giải trộn 136 để giải trộn sự trộn giảm được biến đổi dựa trên thông tin phụ tham số trên hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh để thu được tín hiệu đầu ra âm thanh.

Các phương án sau sử dụng cơ chế cấu trúc dãy cửa sổ đặc biệt. Hàm cửa sổ nguyên mẫu $f(n, N_w)$ được xác định với chỉ số $0 \leq n \leq N_w - 1$ với chiều dài cửa sổ N_w . Thiết kế một cửa sổ $w_k(n)$, cần ba điểm điều khiển, cụ thể là các tâm của cửa sổ trước, hiện thời, tiếp theo, c_{k-1} , c_k , và c_{k+1} .

Sử dụng chúng, hàm cửa sổ được xác định là

$$w_k(n) = \begin{cases} f(n, 2(c_k - c_{k-1})), & \text{for } 0 \leq n < c_k - c_{k-1} \\ f(n - 2c_k + c_{k-1} + c_{k+1}, 2(c_{k+1} - c_k)), & \text{for } c_k - c_{k-1} \leq n < c_{k+1} - c_{k-1} \end{cases}$$

Tiếp đó, vị trí cửa sổ thực tế là $\lceil c_{k-1} \rceil \leq m \leq \lfloor c_{k+1} \rfloor$ với $n = m - \lceil c_{k-1} \rceil$ ($\lceil \cdot \rceil$ là toán tử bao quanh lập luận đến số nguyên tiếp theo phía trên, và $\lfloor \cdot \rfloor$ là toán tử tương ứng bao quanh lập luận đến số nguyên tiếp theo xuống). Hàm cửa sổ nguyên mẫu được sử dụng trong các minh họa là cửa sổ dạng hình sin được xác định là:

$$f(n, N) = \sin\left(\frac{\pi(2n+1)}{2N}\right),$$

nhưng các dạng khác cũng có thể được sử dụng. Vị trí tạm thời t xác định các tâm với ba cửa sổ $c_{k-1} = t - l_b$, $c_k = t$, và $c_{k+1} = t + l_a$, trong đó các số l_b và l_a xác định dãy cửa sổ mong muốn trước và sau phần tạm thời.

Như được giải thích sau đối với Fig.9, bộ tạo dãy cửa sổ 134 có thể, ví dụ, được tạo cấu hình để xác định nhiều cửa sổ phân tích, sao cho phần tạm thời được chứa bởi cửa sổ phân tích thứ nhất trong nhiều cửa sổ phân tích và cửa sổ phân tích thứ hai trong nhiều cửa sổ phân tích, trong đó tâm c_k của cửa sổ phân tích thứ nhất được xác định bằng vị trí t của phần tạm thời theo $c_k = t - l_b$, và tâm c_{k+1} của cửa sổ phân tích thứ nhất được xác định bằng vị trí t của phần tạm thời theo $c_{k+1} = t + l_a$, trong đó l_a và l_b là các số.

Như được giải thích sau đối với Fig.10, bộ tạo dãy cửa sổ 134 có thể, ví dụ, được tạo cấu hình để xác định nhiều cửa sổ phân tích, sao cho phần tạm thời được chứa bởi cửa sổ phân tích thứ nhất trong nhiều cửa sổ phân tích, trong đó tâm c_k của cửa sổ phân tích thứ nhất được xác định bằng vị trí t của phần tạm thời theo $c_k = t$, trong đó tâm c_{k-1} của cửa sổ phân tích thứ hai trong nhiều cửa sổ phân tích được xác định bằng vị trí t của phần tạm thời theo $c_{k-1} = t - l_b$, và trong đó tâm c_{k+1} của cửa sổ phân tích thứ ba trong nhiều cửa sổ phân tích được xác định bằng vị trí t của phần tạm thời theo $c_{k+1} = t + l_a$, trong đó l_a và l_b là các số.

Như được giải thích sau với Fig.11, bộ tạo dãy cửa sổ 134 có thể, ví dụ, được tạo cấu hình để xác định nhiều cửa sổ phân tích, sao cho một trong nhiều cửa sổ phân tích bao gồm số lượng các mẫu tín hiệu miền thời gian thứ nhất hoặc số lượng các mẫu tín hiệu miền thời gian thứ hai, trong đó số lượng các mẫu tín hiệu miền thời gian thứ hai lớn hơn số lượng các mẫu tín hiệu miền thời gian thứ nhất, và trong đó mỗi cửa sổ phân tích trong nhiều cửa sổ phân tích bao gồm số lượng các mẫu tín hiệu miền thời gian thứ nhất khi cửa sổ phân tích này bao gồm phần tạm thời.

Theo một phương án, mô đun phân tích t/f 135 được tạo cấu hình để biến đổi các mẫu trộn giảm miền thời gian của mỗi cửa sổ phân tích từ miền thời gian sang miền thời gian tần số bằng cách sử dụng dàn lọc QMF và dàn lọc Nyquist, trong đó bộ phân tích t/f (135) được tạo cấu hình để biến đổi nhiều mẫu tín hiệu miền thời gian của mỗi cửa sổ phân tích tùy thuộc vào chiều dài cửa sổ của cửa sổ phân tích này.

Fig.2a minh họa bộ mã hóa để mã hóa hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh đầu vào. Một trong hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh đầu vào bao gồm nhiều mẫu tín hiệu miền thời gian.

Bộ mã hóa bao gồm bộ dãy cửa sổ 102 để xác định nhiều cửa sổ phân tích. Mỗi cửa sổ phân tích bao gồm nhiều mẫu tín hiệu miền thời gian của mỗi tín hiệu đối tượng âm thanh đầu vào, trong đó mỗi cửa sổ phân tích có chiều dài cửa sổ chỉ ra số các mẫu tín hiệu miền thời gian của cửa sổ phân tích này. Bộ dãy cửa sổ 102 được tạo cấu hình để xác định nhiều cửa sổ phân tích sao cho chiều dài cửa sổ của mỗi cửa sổ phân tích tùy thuộc vào đặc tính tín hiệu của ít nhất một trong hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh đầu vào. Ví dụ, chiều dài cửa sổ có thể tùy thuộc vào nếu cửa sổ phân

tích này bao gồm phần tạm thời chỉ ra sự thay đổi tín hiệu của ít nhất một trong hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh đầu vào.

Hơn nữa, bộ mã hóa bao gồm bộ phân tích t/f 103 để biến đổi các mẫu tín hiệu miền thời gian của mỗi cửa sổ phân tích từ miền thời gian sang miền thời gian tần số để thu được các mẫu tín hiệu được biến đổi. Bộ phân tích t/f 103 có thể được tạo cấu hình để biến đổi nhiều mẫu tín hiệu miền thời gian của mỗi cửa sổ phân tích tùy thuộc vào chiều dài cửa sổ của cửa sổ phân tích này.

Hơn nữa, bộ mã hóa bao gồm bộ ước tính PSI 104 để xác định thông tin phụ tham số tùy thuộc vào các mẫu tín hiệu được biến đổi.

Theo phương án, bộ mã hóa có thể, ví dụ, còn bao gồm bộ phát hiện tạm thời 101 được tạo cấu hình để xác định nhiều khác biệt về mức của đối tượng của hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh đầu vào, và được tạo cấu hình để xác định, nếu sự khác biệt giữa một trong khác biệt về mức của đối tượng thứ nhất và mỗi khác biệt về mức của đối tượng thứ hai lớn hơn giá trị ngưỡng, để xác định mỗi cửa sổ phân tích, nếu cửa sổ phân tích này bao gồm phần tạm thời chỉ ra sự thay đổi tín hiệu của ít nhất một trong hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh đầu vào.

Theo một phương án, bộ phát hiện tạm thời 101 được tạo cấu hình để sử dụng hàm phát hiện $d(n)$ để xác định nếu sự khác biệt giữa mỗi khác biệt về mức của đối tượng thứ nhất và mỗi khác biệt về mức của đối tượng thứ hai lớn hơn giá trị ngưỡng, trong đó hàm phát hiện $d(n)$ được xác định là:

$$d(n) = \sum_{i,j} \left| \log(OLD_{i,j}(b, n-1)) - \log(OLD_{i,j}(b, n)) \right|$$

trong đó n là chỉ số thời gian, trong đó i là đối tượng thứ nhất, trong đó j là đối tượng thứ hai, trong đó b là dải tham số. OLD có thể, ví dụ, là sự khác biệt về mức của đối tượng.

Như được giải thích sau với Fig.9, bộ dãy cửa sổ 102 có thể, ví dụ, được tạo cấu hình để xác định nhiều cửa sổ phân tích, sao cho phần tạm thời chỉ ra sự thay đổi tín hiệu của ít nhất một trong hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh đầu vào, được chứa bởi cửa sổ phân tích thứ nhất trong nhiều cửa sổ phân tích và cửa sổ phân tích thứ hai trong nhiều cửa sổ phân tích, trong đó tâm c_k của cửa sổ phân tích thứ nhất

được xác định bằng vị trí t của phần tạm thời theo $c_k = t - l_b$, và tâm c_{k+1} của cửa sổ phân tích thứ nhất được xác định bằng vị trí t của phần tạm thời theo $c_{k+1} = t + l_a$, trong đó l_a và l_b là các số.

Như được giải thích sau với Fig.10, bộ dãy cửa sổ 102 có thể, ví dụ, được tạo cấu hình để xác định nhiều cửa sổ phân tích, sao cho phần tạm thời chỉ ra sự thay đổi tín hiệu của ít nhất một trong hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh đầu vào, được chứa bởi cửa sổ phân tích thứ nhất trong nhiều cửa sổ phân tích, trong đó tâm c_k của cửa sổ phân tích thứ nhất được xác định bằng vị trí t của phần tạm thời theo $c_k = t$, trong đó tâm c_{k-1} của cửa sổ phân tích thứ hai trong nhiều cửa sổ phân tích được xác định bằng vị trí t của phần tạm thời theo $c_{k-1} = t - l_b$, và trong đó tâm c_{k+1} của cửa sổ phân tích thứ ba trong nhiều cửa sổ phân tích được xác định bằng vị trí t của phần tạm thời theo $c_{k+1} = t + l_a$, trong đó l_a và l_b là các số.

Như được giải thích sau với Fig.11, bộ dãy cửa sổ 102 có thể, ví dụ, được tạo cấu hình để xác định nhiều cửa sổ phân tích, sao cho một trong nhiều cửa sổ phân tích bao gồm số lượng các mẫu tín hiệu miền thời gian thứ nhất hoặc số lượng các mẫu tín hiệu miền thời gian thứ hai, trong đó số lượng các mẫu tín hiệu miền thời gian thứ hai lớn hơn số lượng các mẫu tín hiệu miền thời gian thứ nhất, và trong đó mỗi cửa sổ phân tích trong nhiều cửa sổ phân tích bao gồm số lượng các mẫu tín hiệu miền thời gian thứ nhất khi cửa sổ phân tích này bao gồm phần tạm thời chỉ ra sự thay đổi tín hiệu của ít nhất một trong hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh đầu vào.

Theo một phương án, bộ phân tích t/f 103 được tạo cấu hình để biến đổi các mẫu tín hiệu miền thời gian của mỗi cửa sổ phân tích từ miền thời gian sang miền thời gian tần số bằng cách sử dụng dàn lọc QMF và dàn lọc Nyquist, trong đó bộ phân tích t/f 103 được tạo cấu hình để biến đổi nhiều mẫu tín hiệu miền thời gian của mỗi cửa sổ phân tích tùy thuộc vào chiều dài cửa sổ của cửa sổ phân tích này.

Sau đây, SAOC được tăng cường sử dụng dàn lọc thích ứng tương hợp ngược được mô tả.

Thứ nhất, giải thích việc giải mã các dòng bit SAOC tiêu chuẩn bằng bộ giải mã SAOC được tăng cường.

Bộ giải mã SAOC được tăng cường được thiết kế sao cho nó có thể giải mã các dòng bit từ bộ mã hóa SAOC tiêu chuẩn với chất lượng tốt. Việc giải mã được giới hạn chỉ đến tái cấu trúc tham số, và có thể có các dòng dư được bỏ qua.

Fig.6 mô tả sơ đồ khối của bộ giải mã SAOC được tăng cường theo một phương án, minh họa việc giải mã các dòng bit SAOC. Các khối chức năng màu đen đậm (132, 133, 134, 135) biểu thị việc xử lý sáng tạo. Thông tin phụ tham số (PSI) bao gồm các tập hợp của các khác biệt về mức của đối tượng (OLD), các tương quan giữa các đối tượng (IOC), và ma trận trộn giảm $\mathbf{D}$ được sử dụng để tạo ra tín hiệu trộn giảm (âm thanh DMX) từ các đối tượng riêng biệt trong bộ giải mã. Mỗi tập hợp tham số được kết hợp với biên tham số mà xác định vùng thời gian mà các tham số được kết hợp ở đó. Trong SAOC chuẩn, các ngăn tần số của phép biểu diễn thời gian/tần số trải qua được nhóm vào các dải tham số. Khoảng cách của các dải giống của các dải quan trọng trong hệ thống thính giác của người. Hơn nữa, nhiều khung biểu diễn t/f có thể được nhóm vào khung tham số. Cả hai thao tác này tạo ra việc giảm lượng thông tin phụ yêu cầu với cái giá là sự thiếu chính xác mô hình hóa.

Như được mô tả trong tiêu chuẩn SAOC, các OLD và IOC được sử dụng để tính ma trận giải trộn $\mathbf{G} = \mathbf{E}\mathbf{D}^T\mathbf{J}$, trong đó các phần tử của $\mathbf{E}$ là $E(i, j) = IOC_{i,j} \sqrt{OLD_i OLD_j}$ xấp xỉ ma trận tương quan chéo của đối tượng, i và j là các chỉ số của đối tượng, $\mathbf{J} \approx (\mathbf{D}\mathbf{E}\mathbf{D}^T)^{-1}$, và $\mathbf{D}^T$ là hoán vị của $\mathbf{D}$. Do đó, bộ tính ma trận giải trộn 131 có thể được tạo cấu hình để tính ma trận giải trộn.

Tiếp đó, ma trận giải trộn được nội suy tuyến tính bằng bộ nội suy thời gian 132 từ ma trận giải trộn của khung đi trước qua khung tham số lên đến biên tham số mà các giá trị được ước tính đạt đến, theo SAOC tiêu chuẩn. Việc này tạo ra các ma trận giải trộn với mỗi cửa sổ phân tích thời gian/thời gian và dải tham số.

Độ phân giải tần số của dải tham số của các ma trận giải trộn được mở rộng đến độ phân giải của phép biểu diễn thời gian-tần số trong cửa sổ phân tích bằng bộ thích ứng với độ phân giải tần số cửa sổ 133. Khi ma trận giải trộn được nội suy với dải tham số b trong khung thời gian được xác định là $\mathbf{G}(b)$, cùng các hệ số giải trộn được sử dụng với tất cả các ngăn tần số trong dải tham số đó.

Bộ tạo dãy cửa sổ 134 được tạo cấu hình để sử dụng thông tin dải đặt theo tham số từ PSI để xác định dãy cửa sổ thích hợp để phân tích tín hiệu âm thanh trộn giảm đầu vào. Yêu cầu chính là ở chỗ khi có biên tập hợp tham số trong PSI, điểm giao nhau giữa các cửa sổ phân tích nên thích hợp với nó. Cửa sổ cũng xác định độ phân giải tần số của dữ liệu trong mỗi cửa sổ (được sử dụng trong mở rộng dữ liệu giải trộn, như được mô tả trước đây).

Tiếp đó, dữ liệu được cửa sổ hóa được biến đổi bằng môđun phân tích t/f 135 vào biểu diễn miền tần số sử dụng biến đổi thời gian-tần số thích hợp, ví dụ, Phép biến đổi Fourier rời rạc (DFT), Phép biến đổi cosin rời rạc được cải biên phức hợp (CMDCT), hoặc Phép biến đổi Fourier rời rạc được xếp chồng lẻ (ODFT).

Cuối cùng, bộ giải trộn 136 áp dụng các ma trận giải trộn ngắn theo khung theo tần số theo trên biểu diễn phổ của tín hiệu trộn giảm $\mathbf{X}$ để thu được các tái cấu trúc $\mathbf{Y}$. Kênh đầu ra j là tổ 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 quy trình này là với hầu hết các mục đích không khác biệt về giác quan từ kết quả thu được với bộ giải mã SAOC tiêu chuẩn.

Lưu ý rằng ngữ cảnh trên đây mô tả sự tái cấu trúc của các đối tượng riêng biệt, nhưng trong SAOC tiêu chuẩn, việc kết xuất chứa trong ma trận giải trộn, nghĩa là, nó bao gồm trong việc nội suy tham số. Do thao tác tuyến tính, thứ tự của các thao tác không phải là vấn đề, nhưng sự chênh lệch không đáng để lưu ý.

Sau đây, việc giải mã các dòng bit SAOC được tăng cường bằng bộ giải mã SAOC được tăng cường được mô tả.

Chức năng chính của bộ giải mã SAOC được tăng cường đã được mô tả trước đây trong giải mã các dòng bit SAOC tiêu chuẩn. Phần này sẽ được giải thích làm thế nào các tăng cường SAOC được đưa vào trong PSI có thể được sử dụng để thu được chất lượng giác quan tốt hơn.

Fig.7 mô tả các khối chức năng chính của bộ giải mã theo một phương án minh họa việc giải mã các tăng cường phân giải tần số. Các khối chức năng màu đen đậm (132, 133, 134, 135) chỉ ra việc xử lý sáng tạo.

Thứ nhất, bộ giá trị mở rộng qua dải 141 làm thích ứng các giá trị OLD và IOC với mỗi dải tham số đến độ phân giải tần số được sử dụng trong các tăng cường, ví dụ, đến 1024 ngấn. Việc này được thực hiện bằng cách tái tạo giá trị qua các ngấn tần số tương ứng với dải tham số. Việc này tạo ra các OLD mới $OLD_i^{enh}(f) = \mathbf{K}(f, b)OLD_i(b)$ và các IOC $IOC_{i,j}^{enh}(f) = \mathbf{K}(f, b)IOC_{i,j}(b)$. $\mathbf{K}(f, b)$ là ma trận kernel xác định việc chuyển các ngấn tần số f thành các dải tham số b bằng

$$\mathbf{K}(f, b) = \begin{cases} 1, & \text{if } f \in b \\ 0, & \text{otherwise} \end{cases}.$$

Đồng thời, bộ phức hồi hàm delta 142 nghịch đảo tham số hóa thừa số chính xác để thu được hàm delta $C_i^{rec}(f)$ có cùng kích cỡ như OLD và IOC được mở rộng.

Tiếp đó, bộ ứng dụng delta 143 áp dụng delta trên các giá trị OLD được mở rộng, và các giá trị OLD phân giải mịn thu được bằng $OLD_i^{fine}(f) = \hat{\mathbf{C}}_i(f)OLD_i^{enh}(f)$.

Theo phương án cụ thể, việc tính các ma trận giải trộn, có thể, ví dụ, được thực hiện bằng bộ tính ma trận giải trộn 131 như với việc giải mã dòng bit SAOC tiêu chuẩn: $\mathbf{G}(f) = \mathbf{E}(f)\mathbf{D}^T(f)\mathbf{J}(f)$, với $\mathbf{E}_{i,j}(f) = IOC_{i,j}^{enh}(f)\sqrt{OLD_i^{fine}(f)OLD_j^{fine}(f)}$, và $\mathbf{J}(f) \approx (\mathbf{D}(f)\mathbf{E}(f)\mathbf{D}^T(f))^{-1}$. Nếu muốn, ma trận kết xuất có thể được nhân với ma trận giải trộn $\mathbf{G}(f)$. Nội suy thời gian bởi bộ nội suy thời gian 132 theo như SAOC tiêu chuẩn.

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

Thông tin dây cửa sổ từ dòng bit có thể được sử dụng để thu được phân tích thời gian-tần số bổ sung hoàn toàn vào thông tin được sử dụng trong bộ mã hóa, hoặc

dãy cửa sổ có thể được cấu trúc dựa trên các biên tham số, như được thực hiện trong giải mã dòng bit SAOC tiêu chuẩn. Do đó, bộ tạo dãy cửa sổ 134 có thể được sử dụng.

Phân tích thời gian-tần số của âm thanh trộn giảm được tiến hành bằng môđun phân tích t/f 135 sử dụng các cửa sổ được cho.

Cuối cùng, các ma trận phân lớp được nội suy và được thích ứng (có khả năng) phổ được áp dụng bằng bộ 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 là sự tổ hợp tuyến tính của các kênh đầu vào $\mathbf{Y}_j(f) = \sum_i \mathbf{G}_{j,i}^{low}(f) \mathbf{X}_i(f)$.

Sau đây, việc mã hóa SAOC được tăng cường tương hợp ngược được mô tả.

Bây giờ, bộ mã hóa SAOC được tăng cường mà tạo ra dòng bit chứa phần thông tin phụ tương hợp ngược và các tăng cường bổ sung được mô tả. Các bộ giải mã SAOC tiêu chuẩn hiện tại có thể giải mã phần tương hợp ngược của PSI và tạo ra các tái cấu trúc của các đối tượng. Thông tin được bổ sung được sử dụng bằng bộ giải mã SAOC được tăng cường cải thiện chất lượng giác quan của các tái cấu trúc trong hầu hết các trường hợp. Ngoài ra, nếu bộ giải mã SAOC được tăng cường chạy trên các tài nguyên có giới hạn, các tăng cường có thể được bỏ qua và sự tái cấu trúc chất lượng cơ bản vẫn đạt được. Lưu ý rằng các tái cấu trúc từ các bộ giải mã SAOC tiêu chuẩn và SAOC được tăng cường chỉ sử dụng PSI tương hợp với SAOC tiêu chuẩn khác nhau, nhưng được xét thấy rất tương tự về giác quan (sự khác biệt là bản chất tương tự như trong giải mã các dòng bit SAOC với bộ giải mã SAOC được tăng cường).

Fig.8 minh họa sơ đồ khối của bộ mã hóa theo phương án cụ thể thực hiện con đường tham số hóa của bộ mã hóa được mô tả trên đây. Các khối chức năng màu đen đậm (102, 103) chỉ ra việc xử lý sáng tạo. Cụ thể là, Fig.8 minh họa sơ đồ khối của dòng bit tương hợp ngược mã hóa hai giai đoạn với các tăng cường với các bộ giải mã có khả năng hơn.

Thứ nhất, tín hiệu được chia nhỏ thành các khung phân tích, mà tiếp đó được biến đổi thành miền tần số. Nhiều khung phân tích được nhóm thành khung tham số chiều dài cố định sử dụng, ví dụ, theo chiều dài MPEG SAOC là 16 và 32 khung phân tích phổ biến. Xem xét rằng các đặc tính tín hiệu duy trì chuẩn dừng trong khung tham

số và do đó có thể được đặc trưng chỉ với một tập hợp tham số. Nếu các đặc trưng tín hiệu thay đổi trong khung tham số, việc mô hình hóa lỗi mắc phải, và sẽ có lợi trong việc chia nhỏ khung tham số lâu hơn thành các phần mà giải thiết gần như dừng lại kết thúc. Với mục đích này, cần phát hiện tạm thời.

Các phần tạm thời có thể được phát hiện bằng bộ phát hiện tạm thời 101 từ tất cả các đối tượng đầu vào riêng biệt, và khi có các sự kiện tạm thời chỉ trong mỗi đối tượng mà vị trí được tuyên bố là vị trí tạm thời toàn cầu. Thông tin của các vị trí tạm thời được sử dụng để cấu trúc dãy cửa sổ thích hợp. Cấu trúc này có thể được sử dụng, ví dụ, dựa trên logic sau đây:

- Cài đặt chiều dài cửa sổ mặc định, nghĩa là, chiều dài của khối biến đổi tín hiệu mặc định, ví dụ, 2048 mẫu.
- Cài đặt chiều dài khung tham số, ví dụ, 4096 mẫu, tương ứng với 4 cửa sổ mặc định với 50% chồng lấp. Các khung tham số nhóm nhiều cửa sổ với nhau và một tập hợp phần mô tả tín hiệu được sử dụng với toàn bộ khối thay vì có các phần mô tả với mỗi cửa sổ riêng biệt. Việc này cho phép giảm lượng PSI.
- Nếu không có phần tạm thời được phát hiện, sử dụng các cửa sổ mặc định và chiều dài khung tham số đầu đủ.
- Nếu phần tạm thời được phát hiện, thích hợp cửa sổ để tạo ra độ phân giải thời gian tốt hơn tại vị trí của phần tạm thời.

Trong khi cấu trúc dãy cửa sổ, bộ dãy cửa sổ 102 chịu trách nhiệm cho nó cũng tạo ra các khung con tham số từ một hoặc nhiều cửa sổ phân tích. Mỗi tập hợp con được phân tích làm thực thể và chỉ một tập hợp của các tham số PSI được truyền tải cho mỗi khối con. Để tạo ra PSI tương hợp với SAOC tiêu chuẩn, chiều dài khối tham số được xác định được sử dụng làm chiều dài khối tham số chính, và các phần tạm thời được định vị có thể có nằm trong khối xác định các tập hợp con tham số.

Dãy cửa sổ được cấu trúc được xuất ra để phân tích thời gian-tần số của các tín hiệu âm thanh đầu vào được tiến hành bởi bộ phân tích t/f 103, và được truyền tải trong phần tăng cường SAOC được tăng cường của PSI.

Dữ liệu phổ của mỗi cửa sổ phân tích được sử dụng bằng bộ ước tính PSI 104 để ước tính PSI với phần SAOC (ví dụ, MPEG) tương hợp ngược. Việc này được thực hiện bằng cách nhóm các ngăn phổ thành các dải tham số của MPEG SAOC và ước tính các IOC, OLD và năng lượng của đối tượng tuyệt đối (NRG) trong các dải. Việc ký hiệu không chặt chẽ của MPEG SAOC sau đây, sản phẩm thông thường hóa của hai phổ đối tượng $S_i(f, n)$ và $S_j(f, n)$ trong ô tham số hóa được xác định là:

$$nrg_{i,j}(b) = \frac{\sum_{n=0}^{N-1} \sum_{f=0}^{F_n-1} \mathbf{K}(b, f, n) S_i(f, n) S_j^*(f, n)}{\sum_{n=0}^{N-1} \sum_{f=0}^{F_n-1} \mathbf{K}(b, f, n)},$$

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

$$\mathbf{K}(b, f, n) = \begin{cases} 1, & \text{if } f \in b \\ 0, & \text{otherwise} \end{cases}, \text{ và}$$

S^* là liên hợp phức của S . Độ phân giải phổ có thể khác nhau giữa các khung trong một khối tham số, nên ma trận ánh xạ biến đổi dữ liệu thành cơ sở phân giải chung. Năng lượng đối tượng tối đa trong ô tham số này được xác định là năng lượng đối tượng tối đa $NRG(b) = \max_i(nrg_{i,i}(b))$. Có giá trị này, tiếp đó, các OLD được xác định là các năng lượng đối tượng được thông thường hóa

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

Và cuối cùng IOC có thể thu được từ các công suất ngang là

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

Việc này kết luận rằng việc ước tính các phần tương hợp SAOC tiêu chuẩn của dòng bit.

Bộ tái cấu trúc phổ công suất kém 105 được tạo cấu hình để sử dụng các OLD và NRG để tái cấu trúc ước tính thô của sự bao phủ phổ trong khối phân tích tham số. Sự bao phủ được cấu trúc trong độ phân giải tần số cao nhất trong khối đó.

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

Các phổ công suất thu được được biến đổi thành biểu diễn độ phân giải tần số cao chung bằng bộ thích ứng phân giải tần số 107. Việc này được thực hiện, ví dụ, bằng cách nội suy các giá trị phổ công suất. Tiếp đó, tập hợp tích phổ công suất trung bình được tính bằng trung bình phổ trong khối tham số. Việc này tương đối tương ứng với ước tính OLD bỏ qua sự kết hợp dải tham số. Biên dạng phổ thu được được xem xét là OLD phân giải mịn.

Bộ ước tính delta 108 được tạo cấu hình để ước tính thừa số chính xác, “delta”, ví dụ, bằng cách chia OLD phân giải mịn bằng tái cấu trúc phổ công suất thô. Do đó, việc này cung cấp cho mỗi ngăn tần số thừa số chính xác (nhân) mà có thể được sử dụng làm tròn OLD phân giải mịn đã cho phổ thô.

Cuối cùng, bộ mô hình hóa delta 109 được tạo cấu hình để mô hình hóa thừa số chính xác được ước tính theo cách có hiệu quả để truyền tải.

Hiệu quả là, các cải biên SAOC được tăng cường với dòng bit bao gồm thông tin dãy cửa sổ và các tham số để truyền tải “delta”.

Sau đây, việc phát hiện tạm thời được mô tả.

Khi các đặc trưng tín hiệu duy trì chuẩn dừng, mã hóa độ tăng ích (đối với lượng thông tin phụ) có thể thu được bằng cách kết hợp một vài khung thời gian thành các khối tham số. Ví dụ, trong SAOC tiêu chuẩn, các giá trị được sử dụng thường là 16 và 32 khung QMF/ một khối tham số. Chúng lần lượt tương ứng với 1024 và 2048 mẫu. Chiều dài của khối tham số có thể được cài đặt trước với giá trị cố định. Một tác dụng trực tiếp mà nó có là độ trễ codec (bộ mã hóa phải có khung đầy đủ để có thể mã hóa nó). Khi sử dụng các khối tham số dài, sẽ có lợi để phát hiện các thay đổi đáng kể trong các đặc trưng tín hiệu, chủ yếu là khi giả thiết dừng bị vi phạm. Sau khi tìm ra sự

thay đổi đáng kể, tín hiệu miền thời gian có thể được chia ở đó và các phần này có thể lại kết thúc giả thiết gần như dừng tốt hơn.

Ở đây, phương pháp phát hiện tạm thời mới được mô tả là được sử dụng kết hợp với SAOC. Một cách mô phạm, nó không nhằm phát hiện các phần tạm thời, nhưng thay vì các thay đổi trong các tham số hóa tín hiệu mà có thể cũng được được gây ra, ví dụ, bằng cách bù tiếng nói.

Tín hiệu đầu vào được chia thành các khung chồng lấp, ngắn và các khung được biến đổi thành miền tần số, ví dụ, với Phép biến đổi Fourier rời rạc (DFT). Phổ phức được biến đổi thành phổ công suất bằng cách nhân các giá trị với các liên hợp phức của chúng (nghĩa là, bình phương các giá trị tuyệt đối của chúng). Tiếp đó, nhóm dải tham số, tương tự với dải được sử dụng trong SAOC tiêu chuẩn, được sử dụng, và năng lượng của mỗi dải tham số trong mỗi khung thời gian ở mỗi đối tượng được tính toán. Các toán tử tóm lại là

$$\mathbf{P}_i(b, n) = \sum_{f \in b} \mathbf{S}_i(f, n) \mathbf{S}_i^*(f, n),$$

trong đó $\mathbf{S}_i(f, n)$ là phổ phức của đối tượng i trong khung thời gian n . Tổng chạy qua các ngăn tần số f trong dải b . Để loại bỏ hiệu ứng ồn từ dữ liệu, các giá trị được lọc qua thấp với bộ lọc IIR thứ tự thứ nhất:

$$\mathbf{P}_i^{LP}(b, n) = a_{LP} \mathbf{P}_i^{LP}(b, n-1) + (1 - a_{LP}) \mathbf{P}_i(b, n),$$

trong đó $0 \leq a_{LP} \leq 1$ là hệ số phản hồi lọc, ví dụ, $a_{LP} = 0.9$.

Sự tham số hóa chính trong SAOC là các khác biệt về mức của đối tượng (OLD). Phương pháp phát hiện được đề xuất nhằm phát hiện khi các OLD sẽ thay đổi. Do đó, tất cả các cặp đối tượng được quan sát với $OLD_{i,j}(b, n) = \mathbf{P}_i^{LP}(b, n) / \mathbf{P}_j^{LP}(b, n)$. Sự thay đổi trong tất cả các cặp đối tượng duy nhất là tính tổng thành hàm phát hiện là

$$d(n) = \sum_{i,j} \left| \log(OLD_{i,j}(b, n-1)) - \log(OLD_{i,j}(b, n)) \right|.$$

Các giá trị thu được được so sánh với ngưỡng T để lọc các độ lệch mức nhỏ ra, và khoảng cách tối thiểu L giữa các phát hiện liên tiếp có hiệu lực. Do đó, hàm phát hiện là

$$\delta(n) = \begin{cases} 1, & \text{if } (d(n) > T) \& (\delta(m) = 0, \forall m : n-L < m < n) \\ 0 & \end{cases}$$

Sau đây, độ phân giải tần số SAOC tăng cường được mô tả.

Độ phân giải tần số thu được từ phân tích SAOC tiêu chuẩn được giới hạn thành số các dải tham số, có giá trị tối đa là 28 theo SAOC tiêu chuẩn. Chúng thu được từ dàn lọc lai bao gồm phân tích QMF 64 dải, sau đó là giai đoạn lọc lai trên các dải thấp mà còn phân chia chúng lên đến 4 dải con phức. Các dải tần số thu được được nhóm thành các dải tham số mô phỏng độ phân giải quan trọng của hệ nghe được của người. Nhóm này cho phép làm giảm tốc độ dữ liệu thông tin phụ yêu cầu.

Hệ thống hiện nay tạo ra chất lượng tách hợp lý đưa ra tốc độ dữ liệu thấp hợp lý. Vấn đề chính là độ phân giải tần số không đủ với việc tách sạch các tiếng âm. Điều này được biểu thị là “quảng” của các đối tượng khác quanh các thành phần âm của đối tượng. Nhận thấy là, việc này được quan sát sơ bộ hoặc thành phần lạ giống bộ mã tiếng nói. Tác dụng có hại của quảng này có thể được giảm bằng cách tăng độ phân giải tần số tham số. Lưu ý rằng độ phân giải bằng hoặc cao hơn 512 dải (ở tốc độ lấy mẫu 44,1 kHz) tạo ra việc tách tốt về các tín hiệu thử nghiệm. Độ phân giải này có thể thu được bằng cách mở rộng giai đoạn lọc lai của hệ thống hiện nay, nhưng các bộ lọc lai sẽ cần thứ tự khá cao với việc tách vừa đủ dẫn đến giá thành tính toán cao.

Một cách đơn giản thu được độ phân giải tần số theo yêu cầu là sử dụng biến đổi thời gian-tần số dựa trên DFT. Chúng có thể được thực hiện có hiệu quả thông qua thuật toán biến đổi Fourier nhanh (Fast Fourier Transform - FFT). Thay vì DFT, CMDCT hoặc ODFT thông thường được xem như các thay thế. Sự khác nhau là ở chỗ hai thuật toán sau là lẻ và phổ thu được chứa các tần số thuần dương và âm. So với DFT, các ngăn tần số nhảy theo độ rộng 0,5 ngăn. Trong DFT, mỗi ngăn ở trung tâm 0Hz và ngăn còn lại ở tần số Nyquist. Sự khác biệt giữa ODFT và CMDCT ở chỗ CMDCT chứa toán tử sau điều biến bổ sung ảnh hưởng đến phổ theo pha. Ưu điểm là ở chỗ phổ phức thu được bao gồm Phép biến đổi cosin rời rạc được cải biên (MDCT) và Phép biến đổi sin rời rạc được cải biên (MDST).

Biến đổi dựa trên DFT của chiều dài N tạo ra phổ phức có N giá trị. Khi thứ tự được biến đổi là giá trị thực, chỉ $N/2$ các giá trị này cần để tái cấu trúc hoàn hảo;

các giá trị $N/2$ còn lại có thể thu được từ các giá trị đưa ra với các thao tác đơn giản. Phân tích này thường thao tác trên chụp khung của N mẫu miền thời gian từ tín hiệu, áp dụng hàm cửa sổ trên các giá trị, và tiếp đó tính biến đổi thực tế trên dữ liệu cửa sổ. Các khối liên tiếp chồng lấp 50% theo thời gian và các hàm cửa sổ được thiết kế sao cho bình phương của các cửa sổ liên tiếp sẽ được tính tổng thống nhất. Việc này đảm bảo rằng khi hàm cửa sổ được áp dụng hai lần trên dữ liệu (mỗi lần phân tích tín hiệu miền thời gian, và lần thứ hai sau khi biến đổi phân tích trước khi bổ sung chồng lấp), chuỗi tổng hợp cùng với phân tích mà không có các cải biên tín hiệu là không mất đi.

Căn cứ vào 50% chồng lấp giữa các khung liên tiếp và chiều dài khung là 2048 mẫu, độ phân giải thời gian có hiệu quả là 1024 mẫu (tương ứng với 23,2ms ở tốc độ lấy mẫu 44,1kHz). Điều này không đủ nhỏ với hai lý do: thứ nhất, sẽ mong muốn có thể giải mã các dòng bit được tạo ra bằng bộ mã hóa SAOC tiêu chuẩn, và thứ hai là, phân tích các tín hiệu trong bộ mã hóa SAOC được tăng cường với các độ phân giải thời gian tốt hơn, nếu cần.

Trong SAOC, có thể nhóm nhiều khối thành các khung tham số. Giả định rằng các đặc tính tín hiệu giữ đủ tương tự so với khung tham số với nó để được đặc trưng với một tập hợp tham số. Các chiều dài khung tham số thường được đếm trong SAOC tiêu chuẩn là 16 hoặc 32 khung QMF (chiều dài lên đến 72 được phép theo tiêu chuẩn). Việc nhóm tương tự có thể được thực hiện khi sử dụng dàn lọc với độ phân giải tần số cao. Khi các đặc tính tín hiệu không thay đổi trong khung tham số, việc nhóm tạo ra hiệu quả mà không làm giảm chất lượng. Tuy nhiên, khi các đặc tính tín hiệu thay đổi trong khung tham số, việc nhóm này gây ra lỗi. SAOC tiêu chuẩn cho phép xác định chiều dài nhóm mặc định, mà được sử dụng với các tín hiệu chuẩn dừng, nhưng cũng xác định các khối con tham số. Các khối con xác định các nhóm ngắn hơn chiều dài mặc định, và việc tham số hóa được thực hiện trên mỗi khối con riêng biệt. Do độ phân giải thời gian của dàn QMF nằm dưới, độ phân giải thời gian thu được là 64 mẫu miền thời gian, độ phân giải này là mịn hơn nhiều so với độ phân giải thu được nhờ sử dụng dàn lọc được cố định với độ phân giải tần số cao. Yêu cầu này ảnh hưởng đến bộ giải mã SAOC được tăng cường.

Việc sử dụng dàn lọc với chiều dài biến đổi lớn tạo ra độ phân giải tần số tốt, nhưng độ phân giải thời gian giảm đồng thời (còn gọi là nguyên lý bất định). Nếu các đặc tính tín hiệu thay đổi trong một khung phân tích, độ phân giải thời gian thấp có thể gây ra mờ ở đầu ra tổng hợp. Do đó, sẽ có lợi để thu được độ phân giải thời gian khung con ở các vị trí của các thay đổi tín hiệu đáng kể. Độ phân giải thời gian khung con thường dẫn một cách tự nhiên đến độ phân giải tần số thấp hơn, nhưng xem xét rằng trong độ phân giải thời gian thay đổi tín hiệu là khía cạnh quan trọng hơn mà được giữ lại chính xác. Yêu cầu phân giải thời gian khung con này chủ yếu ảnh hưởng đến bộ mã hóa SAOC được tăng cường (và do đó cũng là bộ giải mã).

Cùng nguyên lý giải pháp có thể được sử dụng ở cả hai trường hợp: sử dụng các khung phân tích dài khi tín hiệu chuẩn dừng (không có các phần tạm thời được phát hiện) và khi không có các biên tham số. Khi một trong hai điều kiện hỏng, thì sử dụng sơ đồ biến đổi chiều dài khối. Loại trừ với điều kiện này có thể được thực hiện ở các biên tham số mà nằm ở giữa các nhóm khung không được phân chia và trùng với điểm giao giữa hai cửa sổ dài (trong khi giải mã dòng bit SAOC tiêu chuẩn). Xem xét rằng trong trường hợp mà các đặc tính tín hiệu duy trì dừng đủ với dàn lọc có độ phân giải cao. Khi biên tham số được truyền tín hiệu (từ dòng bit hoặc bộ phát hiện tạm thời), việc tạo khung được điều chỉnh để sử dụng chiều dài khung nhỏ hơn, do đó cải thiện độ phân giải thời gian cục bộ.

Hai phương án thứ nhất sử dụng cùng cơ chế cấu trúc dãy cửa sổ nằm dưới. Hàm cửa sổ nguyên mẫu $f(n, N)$ được xác định với chỉ số $0 \leq n \leq N-1$ với chiều dài cửa sổ N . Việc thiết kế cửa sổ $w_k(n)$, ba điểm điều khiển là cần thiết, cụ thể là các tâm của cửa sổ trước, hiện thời và tiếp theo, c_{k-1} , c_k , và c_{k+1} .

Sử dụng chúng, hàm cửa sổ được xác định là

$$w_k(n) = \begin{cases} f(n, 2(c_k - c_{k-1})), & \text{for } 0 \leq n < c_k - c_{k-1} \\ f(n - 2c_k + c_{k-1} + c_{k+1}, 2(c_{k+1} - c_k)), & \text{for } c_k - c_{k-1} \leq n < c_{k+1} - c_{k-1} \end{cases}$$

Tiếp đó, vị trí cửa sổ thực tế là $\lceil c_{k-1} \rceil \leq m \leq \lfloor c_{k+1} \rfloor$ với $n = m - \lceil c_{k-1} \rceil$. Hàm cửa sổ nguyên mẫu được sử dụng trong các minh họa là cửa sổ dạng hình sin được xác định là

$$f(n, N) = \sin\left(\frac{\pi(2n+1)}{2N}\right),$$

nhưng các dạng khác cũng có thể được sử dụng.

Sau đây, sự giao nhau tại phần tạm thời theo phương án được mô tả.

Fig.9 là minh họa về nguyên lý của sơ đồ biến đổi khối "giao nhau tại phần tạm thời". Cụ thể là, Fig.9 minh họa sự thích ứng của dãy cửa sổ thông thường để chứa điểm giao nhau của sổ tại phần tạm thời. Đường 111 thể hiện các mẫu tín hiệu miền thời gian, đường thẳng đứng 112 vị trí t của phần tạm thời được phát hiện (hoặc biên tham số từ dòng bit), và các đường 113 minh họa các hàm cửa sổ và các khoảng thời gian của chúng. Sơ đồ này yêu cầu quyết định lượng chồng lấp giữa hai cửa sổ w_k và w_{k+1} quanh phần tạm thời, xác định độ dốc của cửa sổ. Khi chiều dài chồng lấp được cài đặt đến trị số nhỏ, các cửa sổ có các điểm tối đa của chúng gần phần tạm thời và các phần giao nhau với phần tạm thời phân rã nhanh. Các chiều dài chồng lấp cũng có thể khác nhau trước và sau phần tạm thời. Trong phương pháp này, hai cửa sổ hoặc khung quanh phần tạm thời sẽ được điều chỉnh theo chiều dài. Vị trí của phần tạm thời xác định các tâm của cửa sổ xung quanh sẽ là $c_k = t - l_b$ và $c_{k+1} = t + l_a$, trong đó l_b và l_a lần lượt là chiều dài chồng lấp trước và sau phần tạm thời. Với việc xác định này, công thức nêu trên có thể được sử dụng.

Sau đây, việc tách tạm thời theo một phương án được mô tả.

Fig.10 minh họa nguyên lý của sơ đồ biến đổi khối tách tạm thời theo một phương án. Cửa sổ ngắn w_k nằm ở tâm của phần tạm thời, và hai cửa sổ lân cận w_{k-1} và w_{k+1} được điều chỉnh để hỗ trợ cho cửa sổ ngắn này. Có hiệu quả là, các cửa sổ lân cận được giới hạn đến vị trí tạm thời, nên cửa sổ trước chỉ chứa tín hiệu trước phần tạm thời, và cửa sổ sau chỉ chứa tín hiệu sau phần tạm thời. Trong phương pháp này, phần tạm thời xác định các tâm cho ba cửa sổ $c_{k-1} = t - l_b$, $c_k = t$, và $c_{k+1} = t + l_a$, trong đó l_b và l_a xác định khoảng cửa sổ mong muốn trước và sau phần tạm thời. Với việc xác định này, công thức nêu trên có thể được sử dụng.

Sau đây, việc tạo khung giống như AAC theo một phương án được mô tả.

Mức độ tự do của hai sơ đồ cửa sổ trước có thể không luôn cần thiết. Việc xử lý tạm thời khác nhau cũng được sử dụng trong lĩnh vực mã hóa âm thanh giác quan. Mục đích ở đây là làm giảm sự lan rộng theo thời gian của phần tạm thời mà sẽ gây ra cái gọi là điềm báo. Trong MPEG-2/4 AAC [AAC], hai chiều dài cửa sổ cơ bản được sử dụng: DÀI (có chiều dài 2048 mẫu), và NGẮN (có chiều dài 256 mẫu). Ngoài hai chiều dài này, còn hai cửa sổ chuyển tiếp cũng được xác định để cho phép chuyển tiếp từ DÀI sang NGẮN và ngược lại. Do điều kiện bổ sung, các cửa sổ NGẮN yêu cầu diễn ra trong các nhóm gồm 8 cửa sổ. Cách này, bước dài giữa các cửa sổ và nhóm cửa sổ duy trì ở giá trị không đổi là 1024 mẫu.

Nếu hệ thống SAOC sử dụng codec dựa trên AAC với các tín hiệu của đối tượng, sự trộn giảm hoặc các phần dư của đối tượng, sẽ có lợi để có sơ đồ tạo khung mà có thể dễ dàng được đồng bộ hóa với bộ mã hóa-giải mã. Với lý do này, sơ đồ biến đổi khối dựa trên các cửa sổ AAC được mô tả.

Fig.11 mô tả ví dụ biến đổi khối giống như AAC. Cụ thể là, Fig.11 minh họa cùng tín hiệu với phần tạm thời và dãy cửa sổ giống như AAC thu được. Có thể thấy rằng vị trí thời gian của phần tạm thời được bao phủ bằng 8 cửa sổ NGẮN, mà bao quanh bằng các cửa sổ chuyển tiếp từ và đến các cửa sổ DÀI. Có thể thấy rằng từ việc minh họa thì bản thân phần tạm thời không nằm ở trung tâm ở một cửa sổ cũng nhưng không ở điểm giao nhau giữa hai cửa sổ. Điều này là do các vị trí cửa sổ được cố định vào mạng lưới, nhưng mạng lưới này đảm bảo rằng bước dài không đổi tại cùng thời điểm. Lỗi quanh thời gian thu được được xem xét là đủ nhỏ để không liên quan so với các lỗi gây ra do chỉ sử dụng các cửa sổ DÀI.

Các cửa sổ được xác định là:

- Cửa sổ DÀI: $w_{LONG}(n) = f(n, N_{LONG})$, với $N_{LONG} = 2048$.
- Cửa sổ NGẮN: $w_{SHORT}(n) = f(n, N_{SHORT})$, với $N_{SHORT} = 256$.
- Cửa sổ chuyển tiếp từ DÀI sang NGẮN

$$\mathbf{w}_{START}(n) = \begin{cases} f(n, N_{LONG}), & \text{for } 0 \leq n < \frac{N_{LONG}}{2} \\ 1, & \text{for } \frac{N_{LONG}}{2} \leq n < \frac{2N_{LONG} + 7N_{SHORT}}{4} \\ f(n, N_{SHORT}), & \text{for } \frac{2N_{LONG} + 7N_{SHORT}}{4} \leq n < \frac{2N_{LONG} + 9N_{SHORT}}{4} \\ 0, & \text{for } \frac{2N_{LONG} + 9N_{SHORT}}{4} \leq n < N_{LONG} \end{cases}$$

- Cửa sổ chuyển tiếp từ NGẮN sang DÀI $\mathbf{w}_{STOP}(n) = \mathbf{w}_{START}(N_{LONG} - n - 1)$.

Sau đây, các thay đổi của phương án theo các phương án được mô tả.

Bất chấp sơ đồ biến đổi khối, sự lựa chọn thiết kế khác là chiều dài của biến đổi t/f thực tế. Nếu mục đích chính là giữ các thao tác miền tần số chỉ đơn giản vượt qua các khung phân tích, thì chiều dài biến đổi không đổi có thể được sử dụng. Chiều dài đặt cài đặt đến giá trị lớn thích hợp, ví dụ, tương ứng với chiều dài của khung cho phép lớn nhất. Nếu khung miền thời gian ngắn hơn giá trị này, thì nó được ghi số 0 vào chiều dài đầy đủ. Lưu ý rằng thậm chí sau khi ghi số 0 thì phổ có số ngắn lớn hơn, thì lượng thông tin thực tế không tăng so với biến đổi ngắn hơn. Trong trường hợp này, các ma trận kernel $\mathbf{K}(b, f, n)$ có cùng kích thước với tất cả các giá trị của n .

Theo cách khác là biến đổi khung cửa sổ mà không ghi số 0. Điều này có độ phức tạp tính toán nhỏ hơn với chiều dài biến đổi không đổi. Tuy nhiên, các độ phân giải tần số khác nhau giữa các khung liên tiếp cần được xem xét với các ma trận kernel $\mathbf{K}(b, f, n)$.

Sau đây, việc lọc lai được mở rộng theo một phương án được mô tả.

Khả năng khác để thu được độ phân giải tần số cao hơn sẽ cải biên dần lọc lai được sử dụng trong SAOC tiêu chuẩn với độ phân giải mịn hơn. Trong SAOC tiêu chuẩn, chỉ ba dải 64 QMF thấp nhập được đi qua dần lọc Nyquist chia nhỏ các nội dung dải hơn nữa.

Fig.12 minh họa việc lọc lai QMF được mở rộng. Các bộ lọc Nyquist được lặp lại với mỗi dải QMF riêng biệt, và đầu ra được kết hợp với một phổ có độ phân giải cao. Cụ thể là, Fig.12 minh họa làm thế nào để thu được độ phân giải tần số có thể so sánh với phương pháp dựa trên DFT sẽ yêu cầu chia nhỏ mỗi dải QMF thành, ví dụ, 16

dải con (yêu cầu lọc phức thành 32 dải con). Nhược điểm của phương pháp này là các nguyên mẫu lọc yêu cầu dài do độ hẹp của các dải. Việc này gây ra một số độ trễ xử lý và làm tăng độ phức tạp tính toán.

Cách khác là thực hiện lọc lại được mở rộng bằng cách thay thế các tập hợp của bộ lọc Nyquist bằng các dàn lọc/biến đổi có hiệu quả (ví dụ, “phóng to” DFT, Biến đổi cosin rời rạc, v.v). Hơn nữa, sai số chứa trong các hệ số phổ phân giải cao thu được mà gây ra do các hiệu ứng nửa của giai đoạn lọc thứ nhất (ở đây là: QMF), có thể được giảm về cơ bản bằng cách xử lý sau hủy sai số của các hệ số phổ phân giải cao tương tự với dàn lọc lại MPEG-1/2 Lớp 3 đã biết [FB] [MPEG-1].

Fig.1b minh họa bộ giải mã để tạo ra tín hiệu đầu ra âm thanh bao gồm một hoặc nhiều kênh đầu ra âm thanh từ tín hiệu trộn giảm bao gồm nhiều mẫu trộn giảm miền thời gian theo phương án tương ứng. Tín hiệu trộn giảm mã hóa hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh.

Bộ giải mã bao gồm môđun con phân tích thứ nhất 161 để biến đổi nhiều mẫu trộn giảm miền thời gian để thu được nhiều dải con bao gồm nhiều mẫu dải con.

Hơn nữa, bộ giải mã bao gồm bộ tạo dãy cửa sổ 162 để xác định nhiều cửa sổ phân tích, trong đó mỗi cửa sổ phân tích bao gồm nhiều mẫu dải con của một trong nhiều dải con, trong đó mỗi cửa sổ phân tích trong nhiều cửa sổ phân tích có chiều dài cửa sổ chỉ ra số mẫu dải con của cửa sổ phân tích này. Bộ tạo dãy cửa sổ 162 được tạo cấu hình để xác định nhiều cửa sổ phân tích, ví dụ, dựa trên thông tin phụ tham số, sao cho chiều dài cửa sổ của mỗi cửa sổ phân tích tùy thuộc vào đặc tính tín hiệu của ít nhất một trong hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh.

Hơn nữa, bộ giải mã bao gồm môđun phân tích thứ hai 163 để biến đổi nhiều mẫu dải con của mỗi cửa sổ phân tích trong nhiều cửa sổ phân tích tùy thuộc vào chiều dài cửa sổ của cửa sổ phân tích này để thu được sự trộn giảm được biến đổi.

Hơn nữa, bộ giải mã bao gồm bộ giải trộn 164 để giải trộn sự trộn giảm được biến đổi dựa trên thông tin phụ tham số trên hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh để thu được tín hiệu đầu ra âm thanh.

Nói cách khác: việc biến đổi được tiến hành trong hai pha. Trong pha biến đổi thứ nhất, nhiều dải con mà mỗi dải bao gồm nhiều mẫu dải con được tạo ra. Tiếp đó, trong pha thứ hai, sự biến đổi khác được tiến hành. Không kể những thứ khác, các cửa sổ phân tích được sử dụng trong pha thứ hai xác định độ phân giải thời gian và độ phân giải tần số của sự trộn giảm được biến đổi thu được.

Fig.13 minh họa ví dụ mà các cửa sổ ngắn được sử dụng cho biến đổi này. Sử dụng cửa sổ ngắn tạo ra độ phân giải tần số thấp, nhưng độ phân giải thời gian cao. Sử dụng các cửa sổ ngắn có thể, ví dụ, thích hợp, khi phân tậm thời có mặt trong các tín hiệu đối tượng âm thanh được mã hóa ($u_{i,j}$ là các mẫu dải con, và $v_{s,r}$ là các mẫu của sự trộn giảm được biến đổi trong miền tần số- thời gian.)

Fig.14 minh họa ví dụ mà các cửa sổ dài hơn được sử dụng với biến đổi hơn trong ví dụ của Fig.13. Sử dụng các cửa sổ dài dẫn đến độ phân giải tần số cao, nhưng độ phân giải thời gian thấp. Việc sử dụng các cửa sổ dài có thể, ví dụ, thích hợp, khi phân tậm thời không có mặt trong các tín hiệu đối tượng âm thanh được giải mã. (Lần nữa, $u_{i,j}$ là các mẫu dải con, và $v_{s,r}$ là các mẫu của sự trộn giảm được biến đổi trong miền tần số- thời gian.)

Fig.2b minh họa bộ mã hóa tương ứng để mã hóa hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh đầu vào theo một phương án. Mỗi tín hiệu trong số hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh đầu vào bao gồm nhiều mẫu tín hiệu miền thời gian.

Bộ mã hóa bao gồm môđun con phân tích thứ nhất 171 để biến đổi nhiều mẫu tín hiệu miền thời gian để thu được nhiều dải con bao gồm nhiều mẫu dải con.

Hơn nữa, bộ mã hóa bao gồm bộ dãy cửa sổ 172 để xác định nhiều cửa sổ phân tích, trong đó mỗi cửa sổ phân tích bao gồm nhiều mẫu dải con của một trong nhiều dải con, trong đó mỗi cửa sổ phân tích có chiều dài cửa sổ chỉ ra số mẫu dải con của cửa sổ phân tích này, trong đó bộ dãy cửa sổ 172 được tạo cấu hình để xác định nhiều cửa sổ phân tích, sao cho chiều dài cửa sổ của mỗi cửa sổ phân tích tùy thuộc vào đặc tính tín hiệu của ít nhất một trong số hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh đầu vào. Ví dụ, bộ phát hiện tạm thời (tùy ý) 175 có thể tạo ra thông tin nếu

phần tạm thời có mặt trong mỗi tín hiệu đối tượng âm thanh đầu vào sang bộ dây của số 172.

Hơn nữa, bộ mã hóa bao gồm môđun phân tích thứ hai 173 để biến đổi nhiều mẫu dải con của mỗi cửa sổ phân tích trong nhiều cửa sổ phân tích tùy thuộc vào chiều dài cửa sổ của cửa sổ phân tích này để thu được các mẫu tín hiệu được biến đổi.

Hơn nữa, bộ mã hóa bao gồm bộ ước tính PSI 174 để xác định thông tin phụ tham số tùy thuộc vào các mẫu tín hiệu được biến đổi.

Theo các phương án khác, hai môđun phân tích để tiến hành phân tích trong hai pha có thể có mặt, nhưng môđun thứ hai có thể bật hoặc tắt tùy thuộc vào đặc tính tín hiệu.

Ví dụ, nếu độ phân giải tần số cao yêu cầu và độ phân giải thời gian thấp có thể chấp nhận được, thì môđun phân tích thứ hai được bật.

Ngược lại, nếu độ phân giải thời gian cao yêu cầu và độ phân giải thời gian thấp có thể chấp nhận được, thì môđun phân tích thứ hai được tắt.

Fig.1c minh họa bộ giải mã để tạo ra tín hiệu đầu ra âm thanh bao gồm một hoặc nhiều kênh đầu ra âm thanh từ tín hiệu trộn giảm theo phương án này. Tín hiệu trộn giảm mã hóa một hoặc nhiều tín hiệu đối tượng âm thanh.

Bộ giải mã bao gồm bộ điều khiển 181 để cài đặt chỉ dẫn hoạt động sang trạng thái hoạt động tùy thuộc vào đặc tính tín hiệu của ít nhất một trong số một hoặc nhiều tín hiệu đối tượng âm thanh.

Hơn nữa, bộ giải mã bao gồm môđun phân tích thứ nhất 182 để biến đổi tín hiệu trộn giảm để thu được sự trộn giảm được biến đổi thứ nhất bao gồm nhiều kênh dải con thứ nhất.

Hơn nữa, bộ giải mã bao gồm môđun phân tích thứ hai 183 để tạo ra, khi chỉ dẫn hoạt động được cài đặt sang trạng thái hoạt động, sự trộn giảm được biến đổi thứ hai bằng cách biến đổi ít nhất mỗi kênh dải con thứ nhất để thu được nhiều kênh dải con thứ hai, trong đó sự trộn giảm được biến đổi thứ hai bao gồm các kênh dải con thứ nhất mà chưa được biến đổi bằng môđun phân tích thứ hai và các kênh dải con thứ hai.

Hơn nữa, bộ giải mã bao gồm bộ giải trộn 184, trong đó bộ giải trộn 184 được tạo cấu hình để giải trộn sự trộn giảm được biến đổi thứ hai, khi chỉ dẫn hoạt động được cài đặt sang trạng thái hoạt động, dựa trên thông tin phụ tham số trên một hoặc nhiều tín hiệu đối tượng âm thanh để thu được tín hiệu đầu ra âm thanh, và giải trộn trộn giảm được biến đổi thứ nhất, khi chỉ dẫn hoạt động không được cài đặt sang trạng thái hoạt động, dựa trên thông tin phụ tham số trên một hoặc nhiều tín hiệu đối tượng âm thanh để thu được tín hiệu đầu ra âm thanh.

Fig.15 minh họa ví dụ, trong đó độ phân giải tần số cao yêu cầu và độ phân giải thời gian thấp chấp nhận được. Do đó, bộ điều khiển 181 bật môđun phân tích thứ hai bằng cách cài đặt chỉ dẫn hoạt động sang trạng thái hoạt động (ví dụ, bằng cách cài đặt biến logic “activation_indication” sang “activation_indication = true”). Tín hiệu trộn giảm được biến đổi bằng môđun phân tích thứ nhất 182 (không được thể hiện trên Fig.15) để thu được sự trộn giảm được biến đổi thứ nhất. Trong ví dụ này, của Fig.15, sự trộn giảm được biến đổi có ba dải con. Trong các kịch bản ứng dụng thật sự, sự trộn giảm được biến đổi có thể, ví dụ, có, ví dụ, 32 hoặc 64 dải con. Tiếp đó, sự trộn giảm được biến đổi thứ nhất được biến đổi bằng môđun phân tích thứ hai 183 (không được thể hiện trên Fig.15) để thu được sự trộn giảm được biến đổi thứ hai. Trong ví dụ này, của Fig.15, sự trộn giảm được biến đổi có chín dải con. Trong các kịch bản ứng dụng thật sự hơn, sự trộn giảm được biến đổi có thể, ví dụ, có, ví dụ, 512, 1024 hoặc 2048 dải con. Bộ giải trộn 184 tiếp đó sẽ giải trộn sự trộn giảm được biến đổi thứ hai để thu được tín hiệu đầu ra âm thanh.

Ví dụ, bộ giải trộn 184 có thể nhận chỉ dẫn hoạt động từ bộ điều khiển 181. Hoặc, ví dụ, bất kỳ khi nào bộ giải trộn 184 nhận sự trộn giảm được biến đổi thứ hai từ môđun phân tích thứ hai 183, bộ giải trộn 184 kết luận rằng sự trộn giảm được biến đổi thứ hai phải được giải trộn; bất kỳ khi nào bộ giải trộn 184 không nhận sự trộn giảm được biến đổi thứ hai từ môđun phân tích thứ hai 183, bộ giải trộn 184 kết luận rằng sự trộn giảm được biến đổi thứ nhất phải được giải trộn.

Fig.16 minh họa ví dụ, trong đó độ phân giải thời gian cao được yêu cầu và độ phân giải thời gian thấp là có thể chấp nhận được. Do đó, bộ điều khiển 181 tắt môđun phân tích thứ hai bằng cách cài đặt chỉ dẫn hoạt động sang trạng thái khác với trạng

thái hoạt động (ví dụ, bằng cách cài đặt biến logic “activation_indication” sang “activation_indication = false”). Tín hiệu trộn giảm được biến đổi bằng môđun phân tích thứ nhất 182 (không được thể hiện trên Fig.16) để thu được sự trộn giảm được biến đổi thứ nhất. Tiếp đó, ngược với Fig.15, sự trộn giảm được biến đổi thứ nhất không được biến đổi hơn một lần bằng môđun phân tích thứ hai 183. Thay vào đó, bộ giải trộn 184 sẽ trước tiên giải trộn sự trộn giảm được biến đổi thứ hai để thu được tín hiệu đầu ra âm thanh.

Theo một phương án, bộ điều khiển 181 được tạo cấu hình để cài đặt chỉ dẫn hoạt động sang trạng thái hoạt động tùy thuộc vào nếu ít nhất một trong số một hoặc nhiều tín hiệu đối tượng âm thanh bao gồm phần tạm thời chỉ ra sự thay đổi tín hiệu của ít nhất một trong số một hoặc nhiều tín hiệu đối tượng âm thanh.

Theo phương án khác, chỉ dẫn biến đổi dải con được gán cho mỗi kênh dải con thứ nhất. Bộ điều khiển 181 được tạo cấu hình để cài đặt chỉ dẫn biến đổi dải con của mỗi kênh dải con thứ nhất sang trạng thái biến đổi dải con tùy thuộc vào đặc tính tín hiệu của ít nhất một trong số một hoặc nhiều tín hiệu đối tượng âm thanh. Hơn nữa, môđun phân tích thứ hai 183 được tạo cấu hình để biến đổi mỗi kênh dải con thứ nhất, chỉ dẫn biến đổi dải con của mỗi kênh dải con thứ nhất này được cài đặt sang trạng thái biến đổi dải con, để thu được nhiều kênh dải con thứ hai, và không biến đổi mỗi kênh dải con thứ hai, chỉ dẫn biến đổi dải con của mỗi kênh dải con thứ hai này không được cài đặt sang trạng thái biến đổi dải con.

Fig.17 minh họa ví dụ, trong đó bộ điều khiển 181 (không được thể hiện trên Fig.17) của đặt chỉ dẫn biến đổi dải con của dải con thứ hai sang trạng thái biến đổi dải con (ví dụ, bằng cách cài đặt biến logic “subband_transform_indication_2” sang “subband transform_indication_2 = true”). Do đó, môđun phân tích thứ hai 183 (không được thể hiện trên Fig.17) biến đổi dải con thứ hai để thu được ba dải con “phân giải mịn” mới. Trong ví dụ của Fig.17, bộ điều khiển 181 không cài đặt chỉ dẫn biến đổi dải con của dải con thứ nhất và thứ ba sang trạng thái biến đổi dải con (ví dụ, điều này có thể được chỉ ra bằng bộ điều khiển 181 bằng cách cài đặt các biến logic “subband_transform_indication_1” và “subband_transform_indication_3” sang “subband transform_indication_1 = false” và “subband transform_indication_3 =

false”). Do đó, môđun phân tích thứ hai 183 không biến đổi dải con thứ nhất và thứ ba. Thay vào đó, bản thân dải con thứ nhất và dải con thứ ba được sử dụng làm các dải con của sự trộn giảm được biến đổi thứ hai.

Fig.18 minh họa ví dụ, trong đó bộ điều khiển 181 (không được thể hiện trên Fig.18) cài đặt chỉ dẫn biến đổi dải con của dải con thứ nhất và thứ hai sang trạng thái biến đổi dải con (ví dụ, bằng cách cài đặt biến logic “subband_transform_indication_1” sang “subband transform_indication_1 = true” và, ví dụ, bằng cách cài đặt biến logic “subband_transform_indication_2” sang “subband transform_indication_2 = true”). Do đó, môđun phân tích thứ hai 183 (không được thể hiện trên Fig.18) biến đổi dải con thứ nhất và thứ hai để thu được sáu dải con “phân giải mịn” mới. Trong ví dụ của Fig.18, bộ điều khiển 181 không cài đặt chỉ định chuyển định dạng dải con của dải con thứ ba sang trạng thái biến đổi dải con (ví dụ, điều này có thể được chỉ ra bằng bộ điều khiển 181 bằng cách cài đặt biến logic “subband_transform_indication_3” sang “subband transform_indication_3 = false”). Do đó, môđun phân tích thứ hai 183 không biến đổi dải con thứ ba. Thay vào đó, bản thân dải con thứ ba được sử dụng làm dải con của sự trộn giảm được biến đổi thứ hai.

Theo một phương án, môđun phân tích thứ nhất 182 được tạo cấu hình để biến đổi tín hiệu trộn giảm để thu được sự trộn giảm được biến đổi thứ nhất bao gồm nhiều kênh dải con thứ nhất bằng cách sử dụng bộ lọc gương vuông góc (QMF).

Theo một phương án, môđun phân tích thứ nhất 182 được tạo cấu hình để biến đổi tín hiệu trộn giảm tùy thuộc vào chiều dài cửa sổ phân tích thứ nhất, trong đó chiều dài cửa sổ phân tích thứ nhất tùy thuộc vào đặc tính tín hiệu, và/hoặc môđun phân tích thứ hai 183 được tạo cấu hình để tạo ra, khi chỉ dẫn hoạt động được cài đặt sang trạng thái hoạt động, sự trộn giảm được biến đổi thứ hai bằng cách biến đổi ít nhất mỗi kênh dải con thứ nhất tùy thuộc vào chiều dài cửa sổ phân tích thứ hai, trong đó chiều dài cửa sổ phân tích thứ hai tùy thuộc vào đặc tính tín hiệu. Phương án này nhận biết để bật và tắt môđun phân tích thứ hai 183, và cài đặt chiều dài của cửa sổ phân tích.

Theo một phương án, bộ giải mã được tạo cấu hình để tạo ra tín hiệu đầu ra âm thanh bao gồm một hoặc nhiều kênh đầu ra âm thanh từ tín hiệu trộn giảm, trong đó tín

hiệu trộn giảm mã hóa hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh. Bộ điều khiển 181 được tạo cấu hình để cài đặt chỉ dẫn hoạt động sang trạng thái hoạt động tùy thuộc vào đặc tính tín hiệu của ít nhất một trong hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh. Hơn nữa, bộ giải trộn 184 được tạo cấu hình để giải trộn sự trộn giảm được biến đổi thứ hai, khi chỉ dẫn hoạt động được cài đặt sang trạng thái hoạt động, dựa trên thông tin phụ tham số trên một hoặc nhiều tín hiệu đối tượng âm thanh để thu được tín hiệu đầu ra âm thanh, và giải trộn sự trộn giảm được biến đổi thứ nhất, khi chỉ dẫn hoạt động không được cài đặt sang trạng thái hoạt động, dựa trên thông tin phụ tham số trên hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh để thu được tín hiệu đầu ra âm thanh.

Fig.2c minh họa bộ mã hóa để mã hóa tín hiệu đối tượng âm thanh đầu vào theo một phương án.

Bộ mã hóa bao gồm bộ điều khiển 191 để cài đặt chỉ dẫn hoạt động sang trạng thái hoạt động tùy thuộc vào đặc tính tín hiệu của tín hiệu đối tượng âm thanh đầu vào.

Hơn nữa, bộ mã hóa bao gồm môđun phân tích thứ nhất 192 để biến đổi tín hiệu đối tượng âm thanh đầu vào để thu được tín hiệu đối tượng âm thanh được biến đổi thứ nhất, trong đó tín hiệu đối tượng âm thanh được biến đổi thứ nhất bao gồm nhiều kênh dải con thứ nhất.

Hơn nữa, bộ mã hóa bao gồm môđun phân tích thứ hai 193 để tạo ra, khi chỉ dẫn hoạt động được cài đặt sang trạng thái hoạt động, tín hiệu đối tượng âm thanh được biến đổi thứ hai bằng cách biến đổi ít nhất một trong nhiều kênh dải con thứ nhất để thu được nhiều kênh dải con thứ hai, trong đó tín hiệu đối tượng âm thanh được biến đổi thứ hai bao gồm các kênh dải con thứ nhất mà chưa được biến đổi bằng môđun phân tích thứ hai và các kênh dải con thứ hai.

Hơn nữa, bộ mã hóa bao gồm bộ ước tính PSI 194, trong đó bộ ước tính PSI 194 được tạo cấu hình để xác định thông tin phụ tham số dựa trên tín hiệu đối tượng âm thanh được biến đổi thứ hai, khi chỉ dẫn hoạt động được cài đặt sang trạng thái hoạt động, và xác định thông tin phụ tham số dựa trên tín hiệu đối tượng âm thanh được biến đổi thứ nhất, khi chỉ dẫn hoạt động không được cài đặt sang trạng thái hoạt động.

Theo một phương án, bộ điều khiển 191 được tạo cấu hình để cài đặt chỉ dẫn hoạt động sang trạng thái hoạt động tùy thuộc vào nút tín hiệu đối tượng âm thanh đầu vào bao gồm phần tạm thời chỉ ra sự thay đổi tín hiệu của tín hiệu đối tượng âm thanh đầu vào.

Theo phương án khác, chỉ dẫn biến đổi dải con được gán cho mỗi kênh dải con thứ nhất. Bộ điều khiển 191 được tạo cấu hình để cài đặt chỉ dẫn biến đổi dải con của mỗi kênh dải con thứ nhất sang trạng thái biến đổi dải con tùy thuộc vào đặc tính tín hiệu của tín hiệu đối tượng âm thanh đầu vào. Môđun phân tích thứ hai 193 được tạo cấu hình để biến đổi mỗi kênh dải con thứ nhất, chỉ dẫn biến đổi dải con của mỗi kênh dải con thứ nhất này được cài đặt sang trạng thái biến đổi dải con, để thu được nhiều kênh dải con thứ hai, và không biến đổi mỗi kênh dải con thứ hai, chỉ dẫn biến đổi dải con của mỗi kênh dải con thứ hai này không được cài đặt sang trạng thái biến đổi dải con.

Theo một phương án, môđun phân tích thứ nhất 192 được tạo cấu hình để biến đổi một trong các tín hiệu đối tượng âm thanh đầu vào bằng cách sử dụng bộ lọc gương vuông góc.

Theo phương án khác, môđun phân tích thứ nhất 192 được tạo cấu hình để biến đổi tín hiệu đối tượng âm thanh đầu vào tùy thuộc vào chiều dài cửa sổ phân tích thứ nhất, trong đó chiều dài cửa sổ phân tích thứ nhất tùy thuộc vào đặc tính tín hiệu, và/hoặc môđun phân tích thứ hai 193 được tạo cấu hình để tạo ra, khi chỉ dẫn hoạt động được cài đặt sang trạng thái hoạt động, tín hiệu đối tượng âm thanh được biến đổi thứ hai bằng cách biến đổi ít nhất một trong nhiều kênh dải con thứ nhất tùy thuộc vào chiều dài cửa sổ phân tích thứ hai, trong đó chiều dài cửa sổ phân tích thứ hai tùy thuộc vào đặc tính tín hiệu.

Theo phương án khác, bộ mã hóa được tạo cấu hình để mã hóa tín hiệu đối tượng âm thanh đầu vào và ít nhất một tín hiệu đối tượng âm thanh đầu vào khác. Bộ điều khiển 191 được tạo cấu hình để cài đặt chỉ dẫn hoạt động sang trạng thái hoạt động tùy thuộc vào đặc tính tín hiệu của tín hiệu đối tượng âm thanh đầu vào và tùy thuộc vào đặc tính tín hiệu của ít nhất một tín hiệu đối tượng âm thanh đầu vào khác. Môđun phân tích thứ nhất 192 được tạo cấu hình để biến đổi ít nhất một tín hiệu đối

tượng âm thanh đầu vào khác để thu được ít nhất một tín hiệu đối tượng âm thanh được biến đổi thứ nhất khác, trong đó một trong số ít nhất một tín hiệu đối tượng âm thanh được biến đổi thứ nhất khác bao gồm nhiều kênh dải con thứ nhất. Môđun phân tích thứ hai 193 được tạo cấu hình để biến đổi, khi chỉ dẫn hoạt động được cài đặt sang trạng thái hoạt động, ít nhất một trong nhiều kênh dải con thứ nhất của ít nhất một trong số ít nhất một tín hiệu đối tượng âm thanh được biến đổi thứ nhất khác để thu được nhiều kênh dải con thứ hai khác. Hơn nữa, bộ ước tính PSI 194 được tạo cấu hình để xác định thông tin phụ tham số dựa trên nhiều kênh dải con thứ hai khác, khi chỉ dẫn hoạt động được cài đặt sang trạng thái hoạt động.

Phương pháp sáng tạo và thiết bị giảm các nhược điểm được đề cập trước đây của việc xử lý SAOC trong tình trạng kỹ thuật sử dụng dàn lọc cố định hoặc phép biến đổi thời gian-tần số. Chất lượng âm thanh tốt hơn có thể thu được bằng cách thích ứng độ phân giải thời gian /tần số của các biến đổi hoặc các dàn lọc được sử dụng để phân tích và tổng hợp các đối tượng trong SAOC. Đồng thời, các thành phần lạ như các phản hồi trước và sau gây ra do thiếu độ chính xác thời gian và các thành phần lạ như độ thô nghe được và tiếng nói kép gây ra do độ chính xác phổ không đủ có thể tối thiểu hóa trong cùng hệ SAOC. Quan trọng nhất, hệ thống SAOC được tăng cường biến đổi thích ứng sáng tạo duy trì khả năng tương hợp ngược với SAOC tiêu chuẩn vẫn tạo ra chất lượng giác quan tốt so với SAOC tiêu chuẩn.

Các phương án đề xuất bộ mã hóa âm thanh hoặc phương pháp mã hóa âm thanh hoặc chương trình máy tính liên quan như được mô tả trên đây. Hơn nữa, các phương án đề xuất bộ mã hóa âm thanh hoặc phương pháp giải mã âm thanh hoặc chương trình máy tính liên quan như được mô tả trên đây. Hơn nữa, các phương án đề xuất tín hiệu âm thanh được mã hóa hoặc vật ghi lưu trữ lưu trữ tín hiệu âm thanh như được mô tả trên đây.

Mặc dù một số khía cạnh được mô tả trong ngữ cảnh của thiết bị, nhưng rõ ràng rằng các khía cạnh này cũng thể hiện phần mô tả của phương pháp tương ứng, trong đó khối hoặc thiết bị tương ứng với bước của phương pháp hoặc dấu hiệu của bước của phương pháp. Tương tự, các khía cạnh được mô tả trong ngữ cảnh của bước của

phương pháp cũng thể hiện phần mô tả của khối tương ứng hoặc mục hoặc dấu hiệu của thiết bị tương ứng.

Tín hiệu được phân tích sáng tạo có thể được lưu trữ trên vật ghi lưu trữ số hoặc có thể được truyền tải trên vật ghi truyền tải như vật ghi truyền tải không dây hoặc vật ghi truyền tải có dây như Internet.

Tùy thuộc vào một số yêu cầu thực hiện, các phương án theo sáng chế có thể được thực hiện trong phần cứng hoặc trong phần mềm. Phương án có thể được thực hiện bằng cách sử dụng vật ghi lưu trữ số, ví dụ, đĩa mềm, DVD, CD, ROM, PROM, EPROM, EEPROM, hoặc bộ nhớ FLASH, có thể đọc các tín hiệu điều khiển có thể đọc điện tử được lưu trữ trên đó, mà kết hợp (hoặc có thể kết hợp) với hệ máy tính có thể lập trình sao cho phương pháp tương ứng được thực hiện.

Một số phương án theo sáng chế bao gồm vật mang dữ liệu không tạm thời có các tín hiệu điều khiển có thể đọc điện tử, sao cho một trong các phương pháp này được mô tả được thực hiện.

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

Các phương án khác bao gồm chương trình máy tính để thực hiện một trong số các phương án được mô tả ở đây, được lưu trữ trên vật mang có thể đọc được bằng máy.

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

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

Do đó, phương án khác của phương pháp sáng tạo là dòng dữ liệu hoặc dãy gồm các tín hiệu biểu diễn chương trình máy tính để thực hiện một trong số các phương pháp được mô tả ở đây. Dòng dữ liệu hoặc thứ tự của các tín hiệu có thể, ví dụ, được tạo cấu hình để được truyền qua kết nối giao tiếp dữ liệu, ví dụ, qua Internet.

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

Phương án khác bao gồm máy tính cài đặt trên đó chương trình máy tính để thực hiện mỗi phương pháp được mô tả ở đây.

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

Các phương án được mô tả trên đây chỉ đơn thuần minh họa cho các nguyên lý của sáng chế. Cần hiểu rằng các cải biên và biến thể của các cách sắp xếp và phần mô tả chi tiết ở đây sẽ rõ ràng đối với các chuyên gia trong lĩnh vực này. Do đó, mục đích chỉ bị giới hạn bởi phạm vi của các điểm yêu cầu bảo hộ và không bởi các phần mô tả chi tiết và giải thích của các phương án trong bản mô tả này.

YÊU CẦU BẢO HỘ

1. Bộ giải mã để tạo ra tín hiệu đầu ra âm thanh bao gồm một hoặc nhiều kênh đầu ra âm thanh từ tín hiệu trộn giảm bao gồm nhiều mẫu trộn giảm miền thời gian, trong đó tín hiệu trộn giảm mã hóa hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh, trong đó bộ giải mã bao gồm:

bộ tạo dãy cửa sổ (134) để xác định nhiều cửa sổ phân tích trong đó mỗi cửa sổ phân tích bao gồm nhiều mẫu trộn giảm miền thời gian của tín hiệu trộn giảm, trong đó mỗi cửa sổ phân tích trong số nhiều cửa sổ phân tích có chiều dài cửa sổ biểu thị số lượng các mẫu trộn giảm miền thời gian của cửa sổ phân tích đã nêu, trong đó bộ tạo dãy cửa sổ (134) được tạo cấu hình để xác định nhiều cửa sổ phân tích để cho chiều dài cửa sổ của mỗi cửa sổ phân tích phụ thuộc vào đặc tính tín hiệu của ít nhất một tín hiệu trong số hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh,

môđun phân tích thời gian - tần số (135) để biến đổi nhiều mẫu trộn giảm miền thời gian của mỗi cửa sổ phân tích trong số nhiều cửa sổ phân tích từ miền thời gian đến miền thời gian-tần số phụ thuộc vào chiều dài cửa sổ của cửa sổ phân tích đã nêu, để thu được sự trộn giảm được biến đổi, và

bộ giải trộn (136) để giải trộn sự trộn giảm được biến đổi được dựa trên thông tin phụ tham số trên hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh để thu được tín hiệu đầu ra âm thanh.

2. Bộ giải mã theo điểm 1, trong đó bộ tạo dãy cửa sổ (134) được tạo cấu hình để xác định nhiều cửa sổ phân tích, để cho phần tạm thời, biểu thị sự thay đổi tín hiệu của ít nhất một tín hiệu trong số hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh được mã hóa bởi tín hiệu trộn giảm, được bao gồm bởi cửa sổ phân tích thứ nhất trong số nhiều cửa sổ phân tích và bởi cửa sổ phân tích thứ hai trong số nhiều cửa sổ phân tích, trong đó tâm c_k của cửa sổ phân tích thứ nhất được xác định bởi vị trí t của phần tạm thời theo $c_k = t - l_b$, và tâm c_{k+1} của cửa sổ phân tích thứ hai được xác định bởi vị trí t của phần tạm thời theo $c_{k+1} = t + l_a$, trong đó l_a và l_b là các số.

3. Bộ giải mã theo điểm 1, trong đó bộ tạo dãy cửa sổ (134) được tạo cấu hình để xác định nhiều cửa sổ phân tích, để cho phần tạm thời biểu thị sự thay đổi tín hiệu của ít

nhất một tín hiệu trong số hai hoặc nhiều hơn hai tín hiệu âm thanh được mã hóa bởi tín hiệu trộn giảm, được bao gồm bởi cửa sổ phân tích thứ nhất trong số nhiều cửa sổ phân tích, trong đó tâm c_k của cửa sổ phân tích thứ nhất được xác định bởi vị trí t của phần tạm thời theo $c_k = t$, trong đó tâm c_{k-1} của cửa sổ phân tích thứ hai trong số nhiều cửa sổ phân tích được xác định bởi vị trí t của phần tạm thời theo $c_{k-1} = t - l_b$, và trong đó tâm c_{k+1} của cửa sổ phân tích thứ ba trong số nhiều cửa sổ phân tích được xác định bởi vị trí t của phần tạm thời theo $c_{k+1} = t + l_a$, trong đó l_a và l_b là các số.

4. Bộ giải mã theo điểm 1, trong đó bộ tạo dãy cửa sổ (134) được tạo cấu hình để xác định nhiều cửa sổ phân tích, để cho mỗi cửa sổ trong số nhiều cửa sổ phân tích bao gồm số lượng các mẫu tín hiệu miền thời gian thứ nhất hoặc số lượng các mẫu tín hiệu miền thời gian thứ hai, trong đó số lượng các mẫu tín hiệu miền thời gian thứ hai là lớn hơn số lượng các mẫu tín hiệu miền thời gian thứ nhất, và trong đó mỗi cửa sổ phân tích trong số nhiều cửa sổ phân tích bao gồm số lượng các mẫu tín hiệu miền thời gian thứ nhất khi cửa sổ phân tích đã nêu bao gồm phần tạm thời, biểu thị sự thay đổi tín hiệu của ít nhất một tín hiệu trong số hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh được mã hóa bởi tín hiệu trộn giảm.

5. Bộ giải mã để tạo ra tín hiệu đầu ra âm thanh bao gồm một hoặc nhiều kênh đầu ra âm thanh từ tín hiệu trộn giảm bao gồm nhiều mẫu trộn giảm miền thời gian, trong đó tín hiệu trộn giảm mã hóa hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh, trong đó bộ giải mã bao gồm:

môđun phụ phân tích thứ nhất (161) để biến đổi nhiều mẫu trộn giảm miền thời gian để thu được nhiều dải con bao gồm nhiều mẫu dải con,

bộ tạo dãy cửa sổ (162) để xác định nhiều cửa sổ phân tích, trong đó mỗi cửa sổ phân tích bao gồm nhiều mẫu dải con của một trong số nhiều dải con, trong đó mỗi cửa sổ phân tích trong số nhiều cửa sổ phân tích có chiều dài cửa sổ biểu thị số lượng các mẫu dải con của cửa sổ phân tích đã nêu, trong đó bộ tạo dãy cửa sổ (162) được tạo cấu hình để xác định nhiều cửa sổ phân tích để cho chiều dài cửa sổ của mỗi cửa sổ phân tích phụ thuộc vào đặc tính tín hiệu của ít nhất một tín hiệu trong số hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh,

môđun phân tích thứ hai (163) để biến đổi nhiều mẫu dải con của mỗi cửa sổ phân tích trong số nhiều cửa sổ phân tích phụ thuộc vào chiều dài cửa sổ của cửa sổ phân tích đã nêu để thu được sự trộn giảm được biến đổi, và

bộ giải trộn (164) để giải trộn sự trộn giảm được biến đổi dựa trên thông tin phụ theo tham số trên hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh để thu được tín hiệu đầu ra âm thanh.

6. Bộ mã hóa để mã hóa hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh đầu vào, trong đó mỗi tín hiệu trong số hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh đầu vào bao gồm nhiều mẫu tín hiệu miền thời gian, trong đó bộ mã hóa bao gồm:

bộ tạo dãy cửa sổ (102) để xác định nhiều cửa sổ phân tích, trong đó mỗi cửa sổ phân tích bao gồm nhiều mẫu tín hiệu miền thời gian của một trong số các tín hiệu đối tượng âm thanh đầu vào, trong đó mỗi cửa sổ phân tích có chiều dài cửa sổ biểu thị số lượng mẫu tín hiệu miền thời gian của cửa sổ phân tích đã nêu, trong đó bộ tạo dãy cửa sổ (102) được tạo cấu hình để xác định nhiều cửa sổ phân tích để cho chiều dài cửa sổ của mỗi cửa sổ phân tích phụ thuộc vào đặc tính tín hiệu của ít nhất một tín hiệu trong số hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh đầu vào,

bộ phân tích thời gian - tần số (103) để biến đổi các mẫu tín hiệu miền thời gian của mỗi cửa sổ phân tích từ miền thời gian thành miền thời gian-tần số để thu được mẫu tín hiệu được biến đổi, trong đó bộ phân tích thời gian - tần số (103) được tạo cấu hình để biến đổi nhiều mẫu tín hiệu miền thời gian của mỗi cửa sổ phân tích phụ thuộc vào chiều dài cửa sổ của cửa sổ phân tích đã nêu, và

bộ ước tính thông tin phụ tham số (104) để xác định thông tin phụ theo tham số phụ thuộc vào các mẫu tín hiệu được biến đổi.

7. Bộ mã hóa theo điểm 6, trong đó bộ mã hóa còn bao gồm bộ phát hiện phần tạm thời (101) được tạo cấu hình để xác định nhiều sự chênh lệch mức đối tượng của hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh đầu vào, và được tạo cấu hình để xác định, liệu sự chênh lệch giữa sự chênh lệch thứ nhất trong số các sự chênh lệch mức đối tượng và sự chênh lệch thứ hai trong số các sự chênh lệch mức đối tượng là lớn hơn giá trị ngưỡng hay không, để xác định cho mỗi cửa sổ phân tích, liệu cửa sổ phân

tích đã nêu bao gồm phần tạm thời, biểu thị sự thay đổi tín hiệu của ít nhất một tín hiệu trong số hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh đầu vào hay không.

8. Bộ mã hóa theo điểm 7,

trong đó bộ phát hiện phần tạm thời (101) được tạo cấu hình để sử dụng hàm phát hiện $d(n)$ để xác định liệu sự chênh lệch giữa sự chênh lệch thứ nhất trong số các sự chênh lệch mức đối tượng và sự chênh lệch thứ hai trong số các sự chênh lệch mức đối tượng là lớn hơn giá trị ngưỡng hay không,

trong đó hàm phát hiện $d(n)$ được xác định như:

$$d(n) = \sum_{i,j} \left| \log(OLD_{i,j}(b, n-1)) - \log(OLD_{i,j}(b, n)) \right|$$

trong đó n biểu thị chỉ số,

trong đó i biểu thị đối tượng thứ nhất,

trong đó j biểu thị đối tượng thứ hai, và

trong đó b biểu thị dải tham số.

9. Bộ mã hóa theo điểm 6, trong đó bộ tạo dãy cửa sổ (102) được tạo cấu hình để xác định nhiều cửa sổ phân tích, để cho các phần tạm thời biểu thị sự thay đổi tín hiệu của ít nhất một tín hiệu trong số hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh đầu vào, được bao gồm bởi cửa sổ phân tích thứ nhất trong số nhiều cửa sổ phân tích và bởi cửa sổ phân tích thứ hai trong số nhiều cửa sổ phân tích, trong đó tâm c_k của cửa sổ phân tích thứ nhất được xác định bởi vị trí t của phần tạm thời theo $c_k = t - l_b$, và tâm c_{k+1} của cửa sổ phân tích thứ hai được xác định bởi vị trí t của phần tạm thời theo $c_{k+1} = t + l_a$, trong đó l_a và l_b là các số.

10. Bộ mã hóa theo điểm 6, trong đó bộ tạo dãy cửa sổ (102) được tạo cấu hình để xác định nhiều cửa sổ phân tích, để cho phần tạm thời, chỉ ra sự thay đổi tín hiệu của ít nhất một tín hiệu trong số hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh đầu vào, được bao gồm bởi cửa sổ phân tích thứ nhất trong số nhiều cửa sổ phân tích, trong đó tâm c_k của cửa sổ phân tích thứ nhất được định nghĩa bởi vị trí t của phần tạm thời theo $c_k = t$, trong đó tâm c_{k-1} của cửa sổ phân tích thứ hai trong số nhiều cửa sổ phân tích được định nghĩa bởi vị trí t của phần tạm thời theo $c_{k-1} = t - l_b$, và trong đó tâm c_{k+1} của

cửa sổ phân tích thứ ba trong số nhiều cửa sổ phân tích được xác định bởi vị trí t của phần tạm thời theo $c_{k+1} = t + l_a$, trong đó l_a và l_b là các số.

11. Bộ mã hóa theo điểm 6, trong đó bộ tạo dãy cửa sổ (102) được tạo cấu hình để xác định nhiều cửa sổ phân tích, để cho mỗi cửa sổ trong số nhiều cửa sổ phân tích bao gồm hoặc số lượng các mẫu tín hiệu miền thời gian thứ nhất hoặc số lượng các mẫu tín hiệu miền thời gian thứ hai, trong đó số lượng các mẫu tín hiệu miền thời gian thứ hai là lớn hơn số lượng các mẫu tín hiệu miền thời gian thứ nhất, và trong đó mỗi cửa sổ phân tích trong số nhiều cửa sổ phân tích bao gồm số lượng các mẫu tín hiệu miền thời gian thứ nhất khi cửa sổ phân tích đã nêu bao gồm phần tạm thời, biểu thị sự thay đổi tín hiệu của ít nhất một tín hiệu trong số hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh đầu vào.

12. Bộ mã hóa để mã hóa hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh đầu vào, trong đó mỗi tín hiệu trong số hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh đầu vào bao gồm nhiều mẫu tín hiệu miền thời gian, trong đó bộ mã hóa bao gồm:

môđun phụ phân tích thứ nhất (171) để biến đổi nhiều tín hiệu miền thời gian để thu được nhiều dải con bao gồm nhiều mẫu dải con,

bộ tạo dãy cửa sổ (172) để xác định nhiều cửa sổ phân tích, trong đó mỗi cửa sổ phân tích bao gồm nhiều mẫu dải con của một trong số nhiều dải con, trong đó mỗi cửa sổ phân tích có chiều dài cửa sổ biểu thị số lượng mẫu dải con của cửa sổ phân tích đã nêu, trong đó bộ tạo dãy cửa sổ (172) được tạo cấu hình để xác định nhiều cửa sổ phân tích để cho chiều dài cửa sổ của mỗi cửa sổ phân tích phụ thuộc vào đặc tính tín hiệu của ít nhất một tín hiệu trong số hai hoặc nhiều hơn hai đối tượng âm thanh đầu vào,

môđun phân tích thứ hai (173) để biến đổi nhiều mẫu dải con của mỗi cửa sổ phân tích trong số nhiều cửa sổ phân tích phụ thuộc vào chiều dài cửa sổ của cửa sổ phân tích đã nêu để thu được các mẫu tín hiệu được biến đổi, và

bộ ước tính thông tin phụ tham số (174) để xác định thông tin phụ theo tham số phụ thuộc vào các mẫu tín hiệu được biến đổi.

13. Phương pháp giải mã để tạo ra tín hiệu đầu ra âm thanh bao gồm một hoặc nhiều kênh đầu ra âm thanh từ tín hiệu trộn giảm bao gồm nhiều mẫu trộn giảm miền thời gian, trong đó tín hiệu trộn giảm mã hóa hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh, trong đó phương pháp bao gồm:

xác định nhiều cửa sổ phân tích, trong đó mỗi cửa sổ phân tích bao gồm nhiều mẫu trộn giảm miền thời gian của tín hiệu trộn giảm, trong đó mỗi cửa sổ phân tích trong số nhiều cửa sổ phân tích có chiều dài cửa sổ biểu thị số lượng mẫu trộn giảm miền thời gian trong số cửa sổ phân tích đã nêu, trong đó việc xác định nhiều cửa sổ phân tích được thực hiện để cho chiều dài cửa sổ của mỗi cửa sổ phân tích phụ thuộc vào đặc tính tín hiệu của ít nhất một tín hiệu trong số hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh,

biến đổi nhiều mẫu trộn giảm miền thời gian của mỗi cửa sổ phân tích trong số nhiều cửa sổ phân tích từ miền thời gian thành miền thời gian – tần số phụ thuộc vào chiều dài cửa sổ của cửa sổ phân tích đã nêu, để thu được sự trộn giảm được biến đổi, và

giải trộn sự trộn giảm được biến đổi dựa trên thông tin phụ theo tham số trên hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh để thu được tín hiệu đối tượng âm thanh.

14. Phương pháp mã hóa hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh, trong đó mỗi tín hiệu trong số hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh đầu vào bao gồm nhiều mẫu tín hiệu miền thời gian, trong đó phương pháp bao gồm:

xác định nhiều cửa sổ phân tích, trong đó mỗi cửa sổ phân tích bao gồm nhiều mẫu tín hiệu miền thời gian của một trong số các tín hiệu đối tượng âm thanh đầu vào, trong đó mỗi cửa sổ phân tích có chiều dài cửa sổ biểu thị số lượng mẫu tín hiệu miền thời gian của cửa sổ phân tích đã nêu, trong đó việc xác định nhiều cửa sổ phân tích được thực hiện để cho chiều dài cửa sổ của mỗi cửa sổ phân tích phụ thuộc vào đặc tính tín hiệu của ít nhất một tín hiệu trong số hai hoặc nhiều hơn hai đối tượng âm thanh đầu vào,

biến đổi mẫu tín hiệu miền thời gian theo mỗi cửa sổ phân tích từ miền thời gian thành miền thời gian-tần số để thu được các mẫu tín hiệu biến đổi, trong đó sự biến đổi nhiều mẫu tín hiệu miền thời gian của mỗi cửa sổ phân tích phụ thuộc vào chiều dài cửa sổ của cửa sổ phân tích đã nêu,

xác định thông tin phụ theo tham số phụ thuộc vào mẫu tín hiệu được biến đổi.

15. Phương pháp giải mã bằng cách tạo ra tín hiệu đầu ra âm thanh bao gồm một hoặc nhiều kênh đầu ra âm thanh từ tín hiệu trộn giảm bao gồm nhiều mẫu trộn giảm miền thời gian, trong đó tín hiệu trộn giảm mã hóa hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh, trong đó phương pháp bao gồm:

biến đổi nhiều mẫu trộn giảm miền thời gian để thu được nhiều dải con bao gồm nhiều mẫu dải con,

xác định nhiều cửa sổ phân tích, trong đó mỗi cửa sổ phân tích bao gồm nhiều mẫu dải con của một trong số nhiều dải con, trong đó mỗi cửa sổ phân tích trong số nhiều cửa sổ phân tích có chiều dài cửa sổ biểu thị số lượng mẫu dải con của cửa sổ phân tích đã nêu, trong đó việc xác định nhiều cửa sổ phân tích được thực hiện để cho chiều dài cửa sổ của mỗi cửa sổ phân tích phụ thuộc vào đặc tính của ít nhất một tín hiệu trong số hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh,

biến đổi nhiều mẫu dải con của mỗi cửa sổ phân tích trong số nhiều cửa sổ phân tích phụ thuộc vào chiều dài cửa sổ của cửa sổ phân tích đã nêu để thu được sự trộn giảm được biến đổi, và

giải trộn sự trộn giảm được biến đổi dựa trên thông tin phụ theo tham số trên hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh để thu được tín hiệu đầu ra âm thanh.

16. Phương pháp mã hóa hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh đầu vào, trong đó mỗi tín hiệu trong số hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh đầu vào bao gồm nhiều mẫu tín hiệu miền thời gian, trong đó phương pháp bao gồm:

biến đổi nhiều mẫu tín hiệu miền thời gian để thu được nhiều dải con bao gồm nhiều mẫu dải con,

xác định nhiều cửa sổ phân tích, trong đó mỗi cửa sổ phân tích bao gồm nhiều mẫu dải con của một trong số nhiều dải con, trong đó mỗi cửa sổ phân tích có chiều dài cửa sổ biểu thị số mẫu dải con của cửa sổ phân tích đã nêu, trong đó việc xác định nhiều cửa sổ phân tích được thực hiện để cho chiều dài cửa sổ của mỗi cửa sổ phân tích phụ thuộc vào đặc tính tín hiệu của ít nhất một tín hiệu trong số hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh đầu vào,

biến đổi nhiều mẫu dải con của mỗi cửa sổ phân tích trong số nhiều cửa sổ phân tích phụ thuộc vào chiều dài cửa sổ của cửa sổ phân tích đã nêu để thu được các mẫu tín hiệu được biến đổi, và

xác định thông tin phụ theo tham số phụ thuộc vào các mẫu tín hiệu được biến đổi.

17. Vật ghi có thể đọc được bằng máy tính không tạm thời bao gồm chương trình máy tính để thực hiện phương pháp theo điểm 13 khi được thực hiện trên máy tính hoặc bộ xử lý tín hiệu.

18. Vật ghi có thể đọc được bằng máy tính không tạm thời bao gồm chương trình máy tính để thực hiện phương pháp theo điểm 14 khi được thực hiện trên máy tính hoặc bộ xử lý tín hiệu.

19. Vật ghi có thể đọc được bằng máy tính không tạm thời bao gồm chương trình máy tính để thực hiện phương pháp theo điểm 15 khi được thực hiện trên máy tính hoặc bộ xử lý tín hiệu.

20. Vật ghi có thể đọc được bằng máy tính không tạm thời bao gồm chương trình máy tính để thực hiện phương pháp theo điểm 16 khi được thực hiện trên máy tính hoặc bộ xử lý tín hiệu.

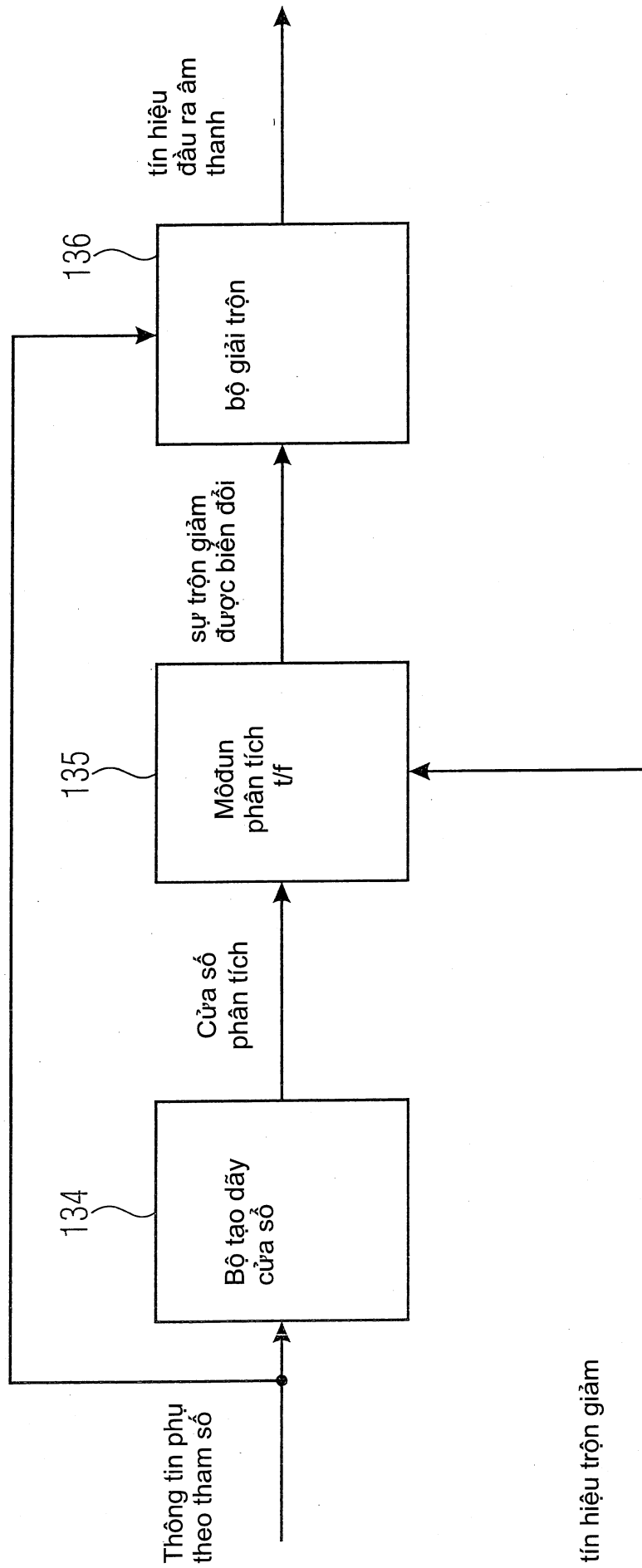


FIG. 1A

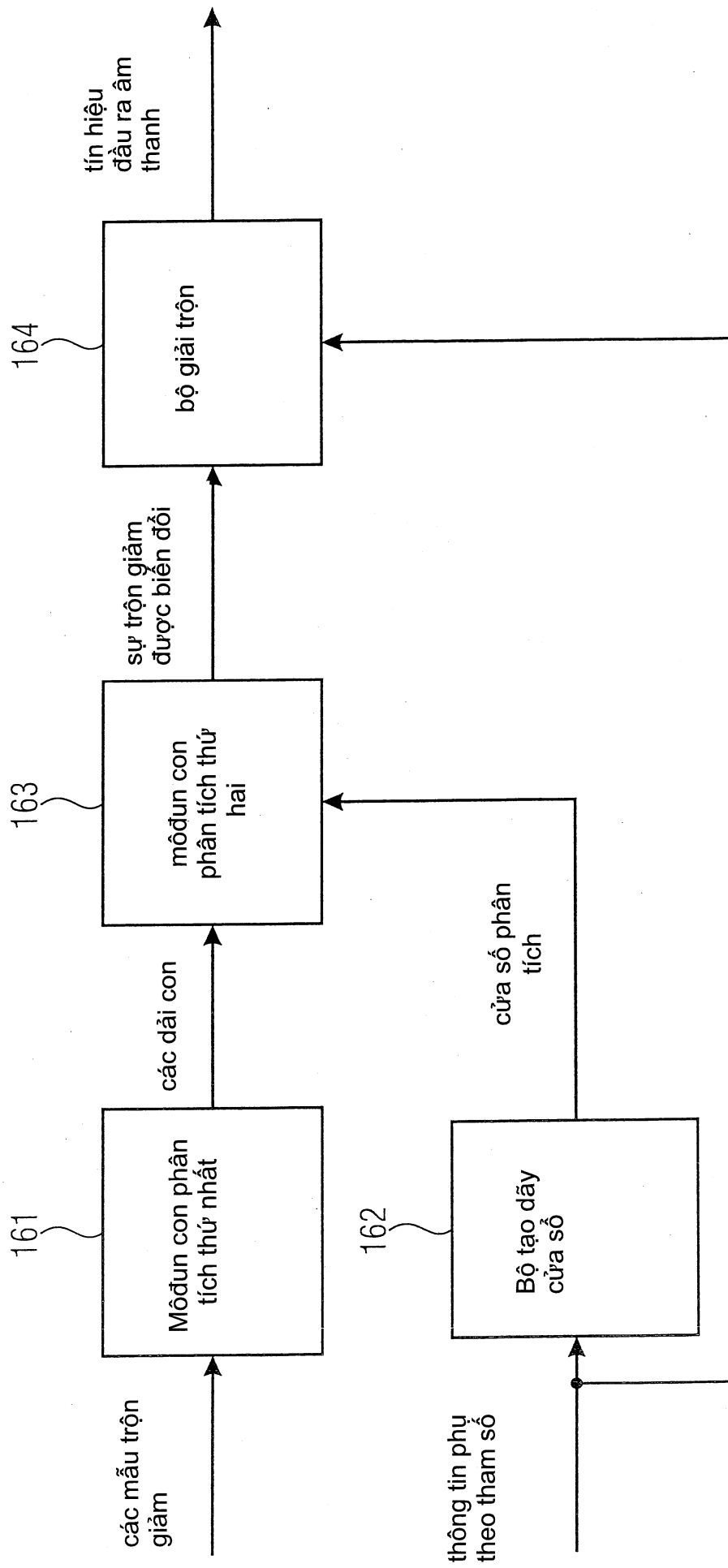


FIG.1B

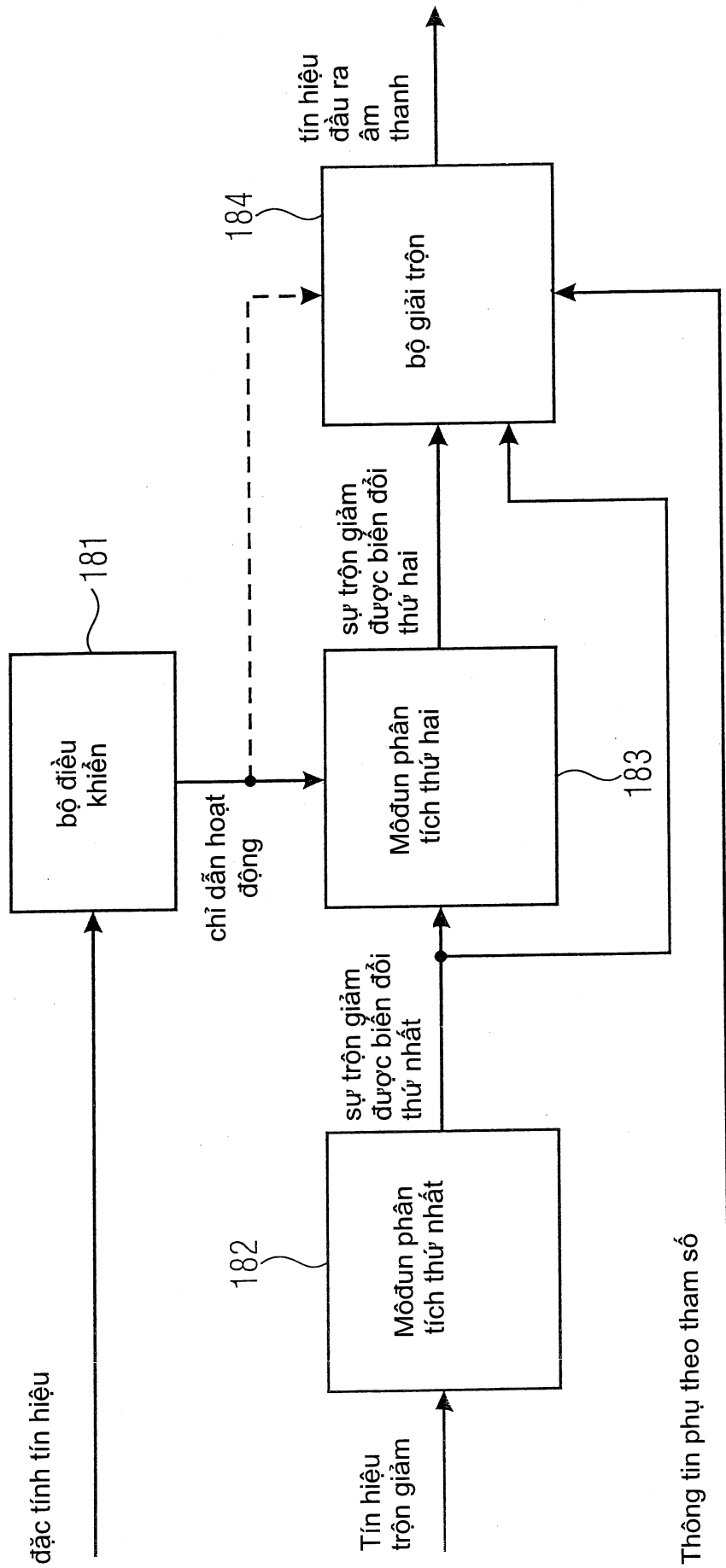


FIG.1C

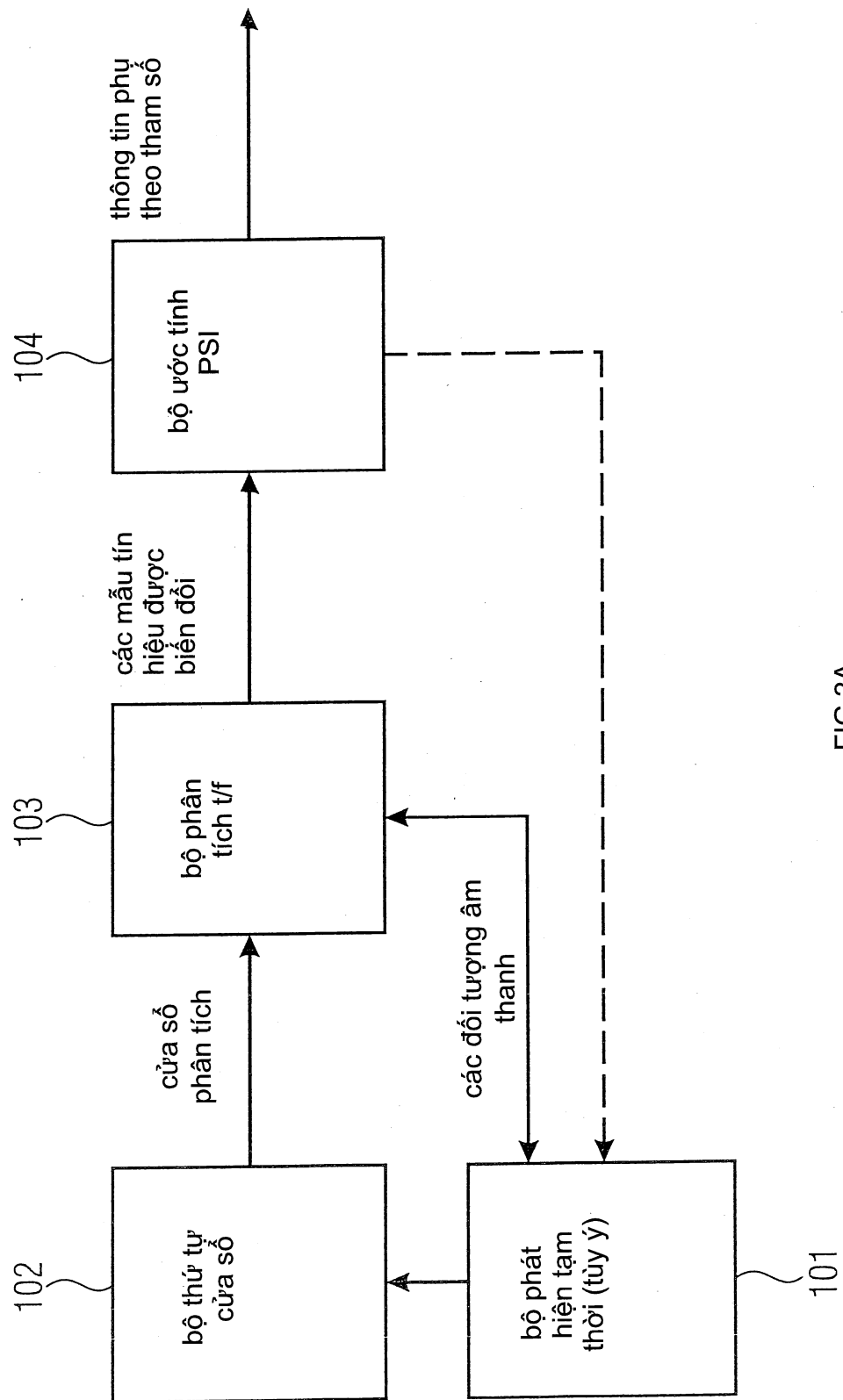


FIG. 2A

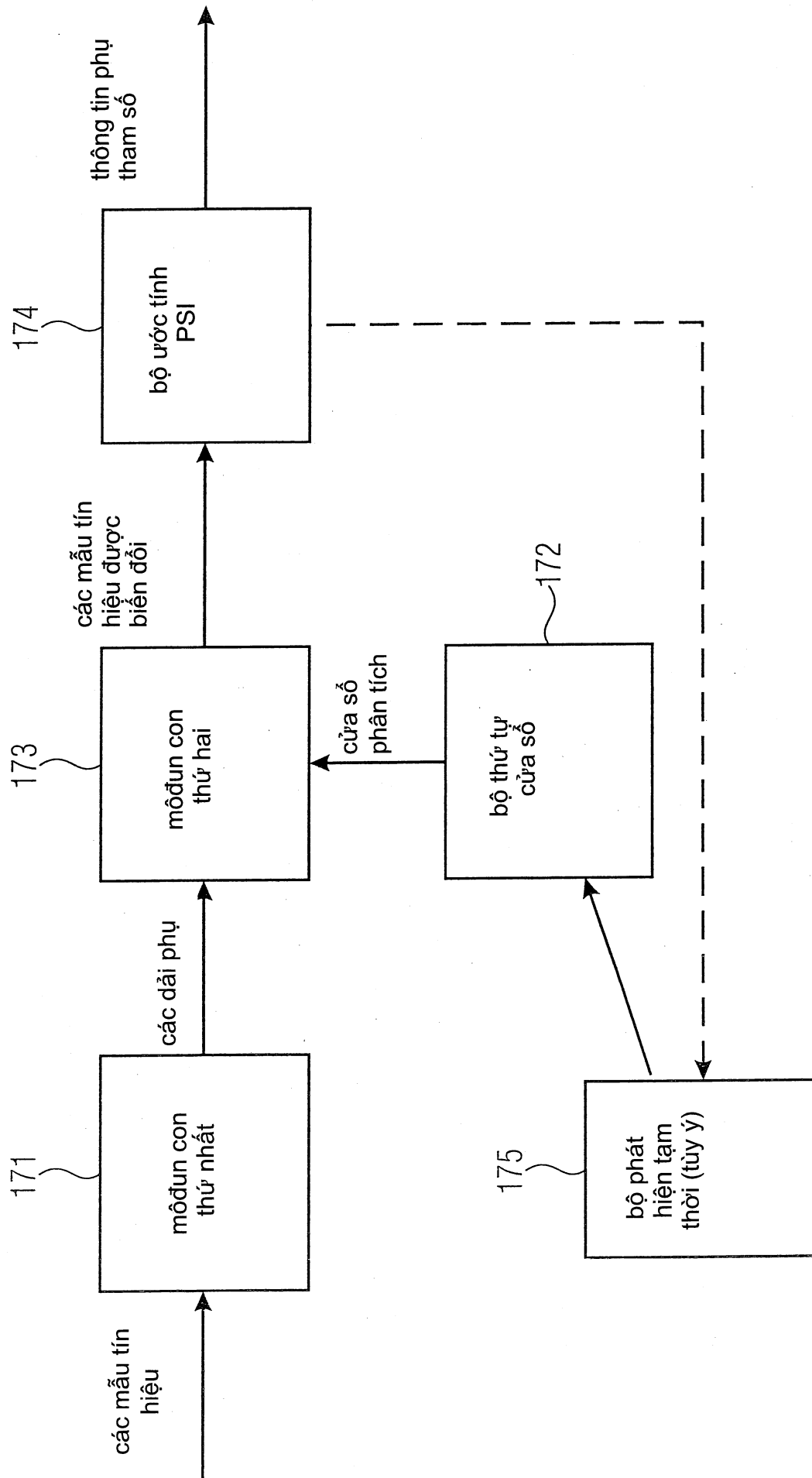


FIG.2B

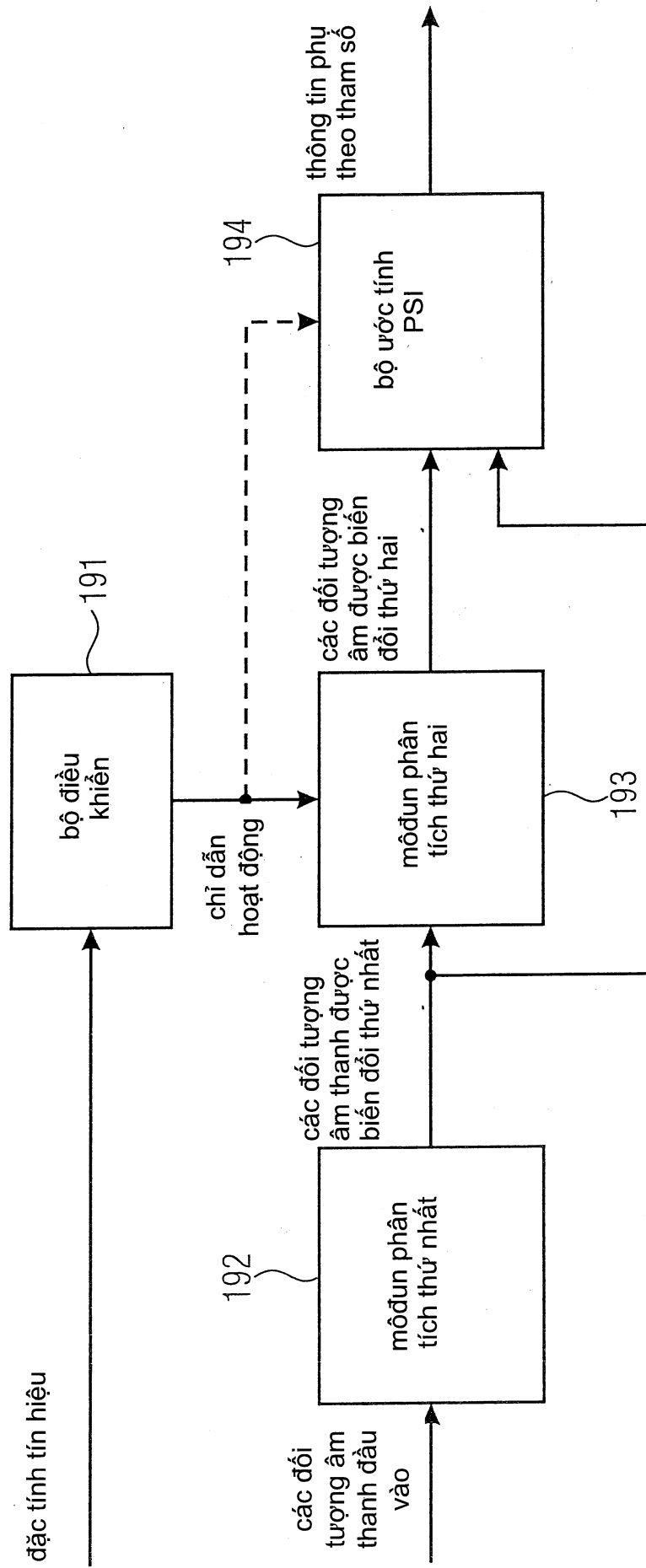


FIG. 2C

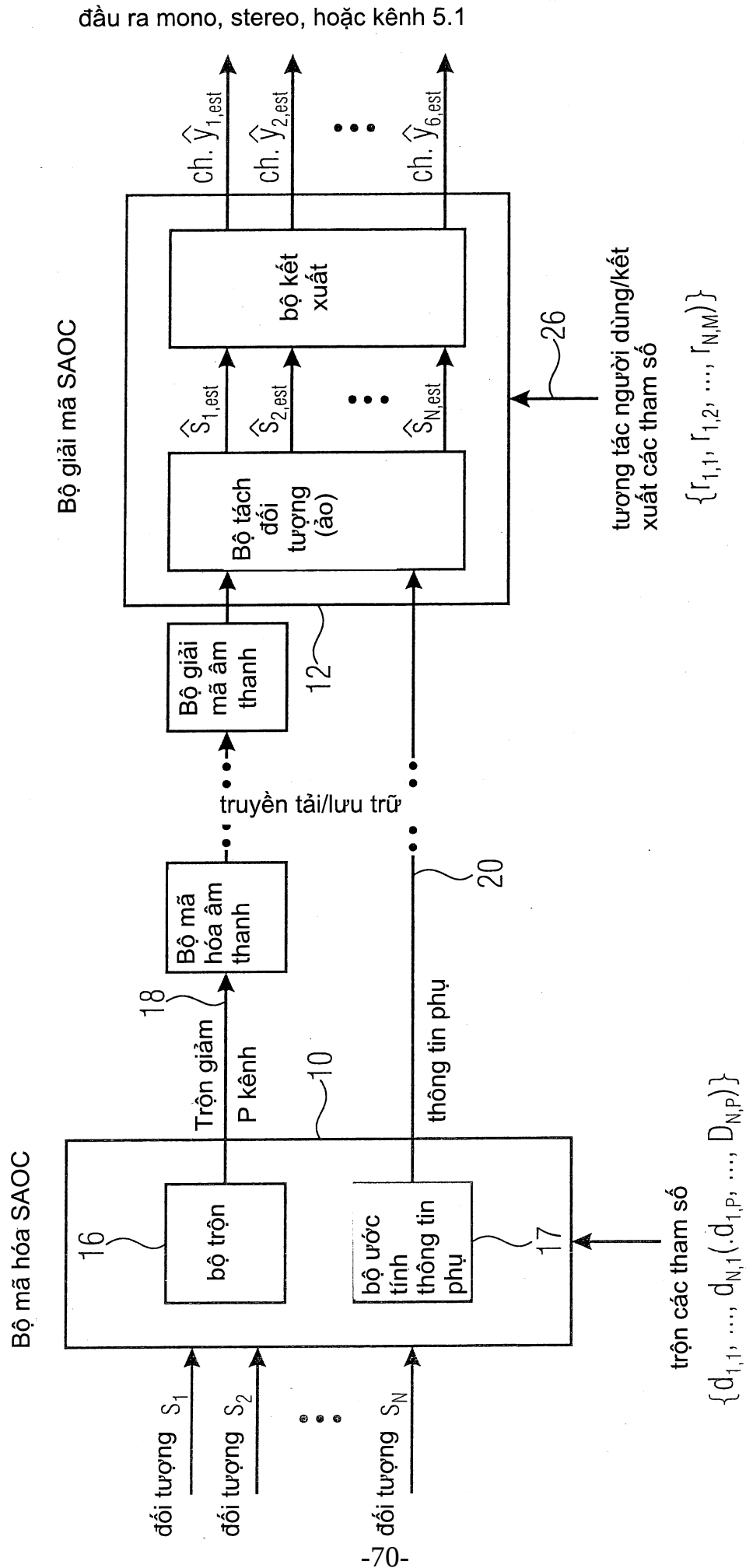


FIG.3

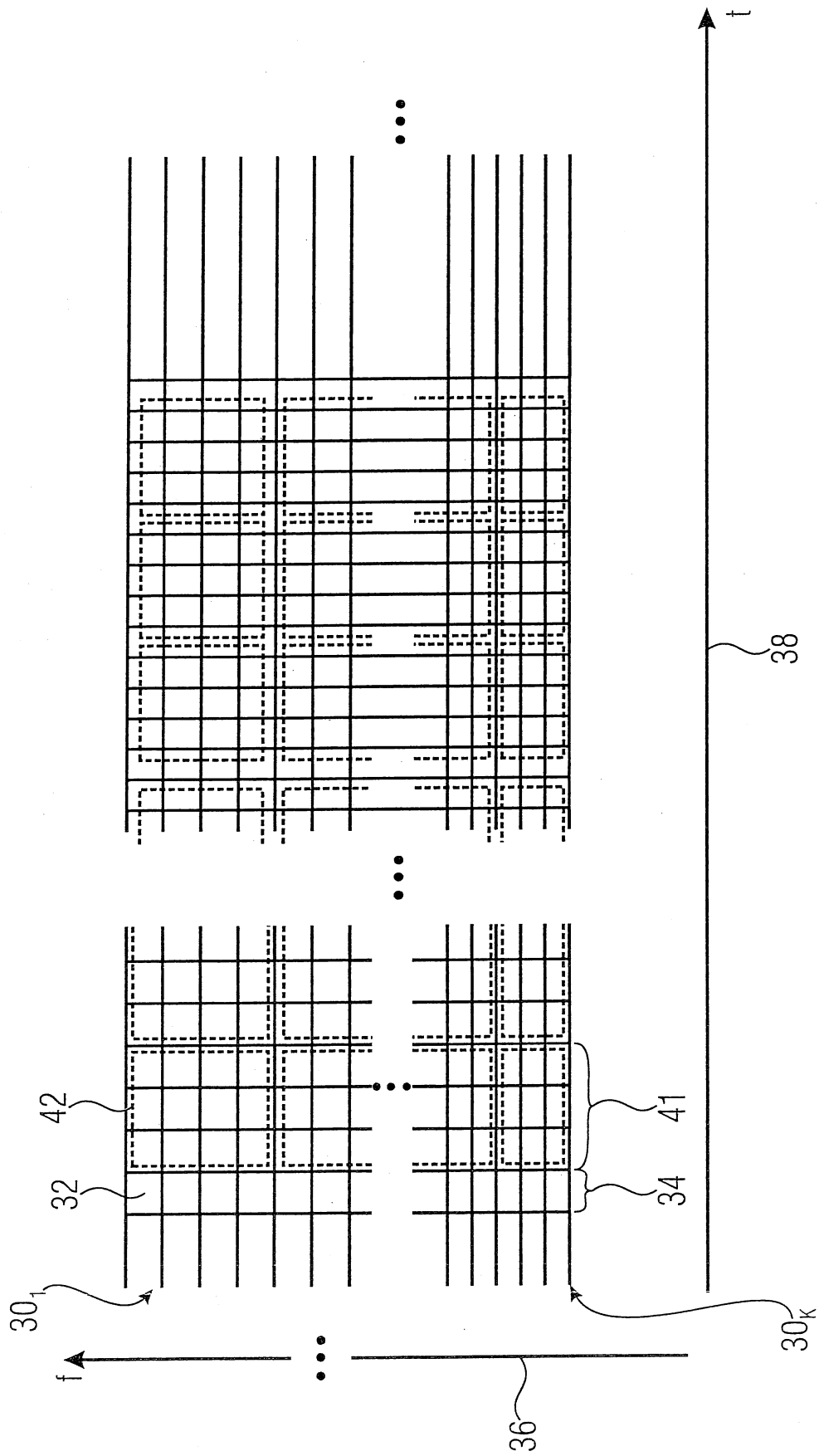


FIG.4

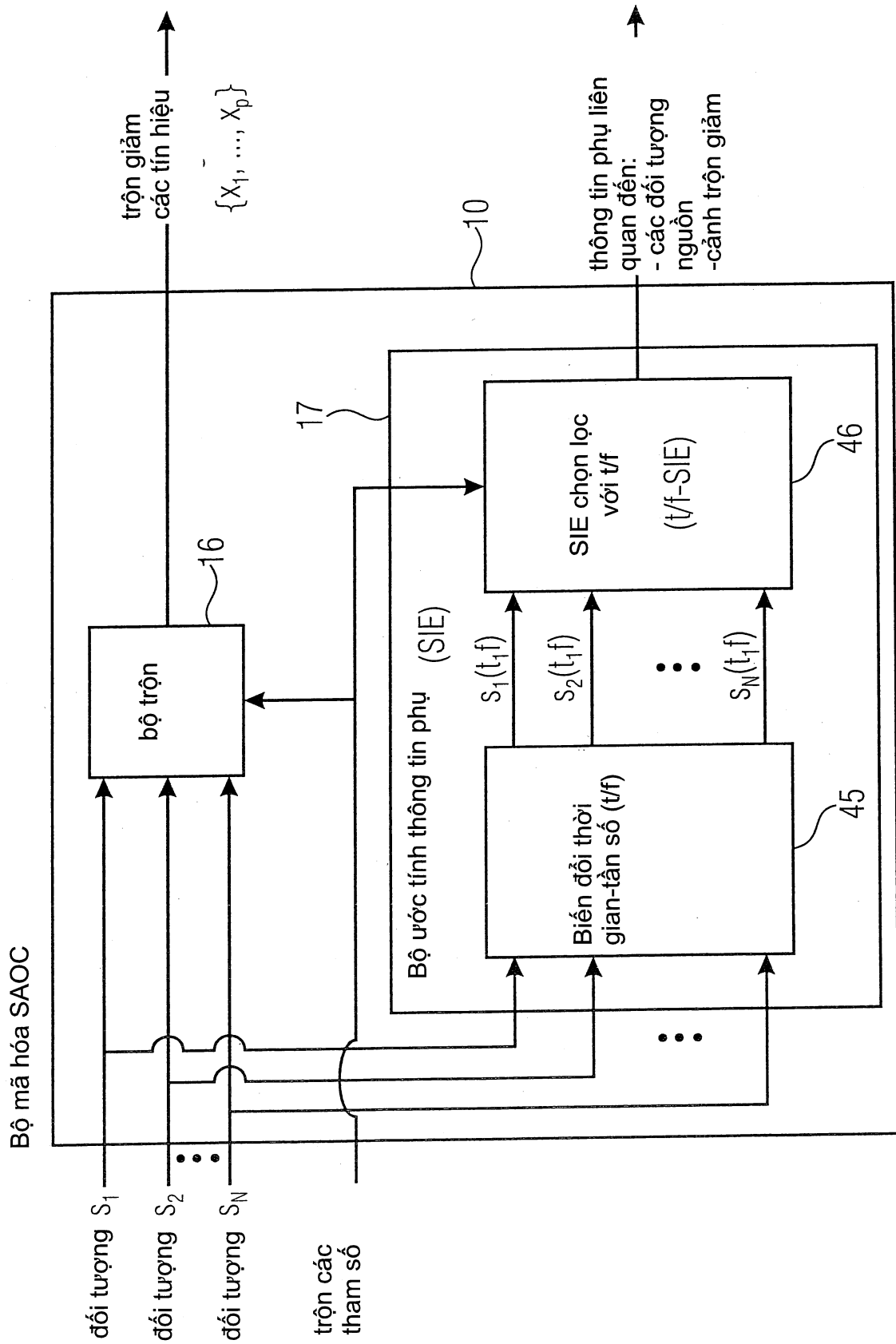


FIG.5

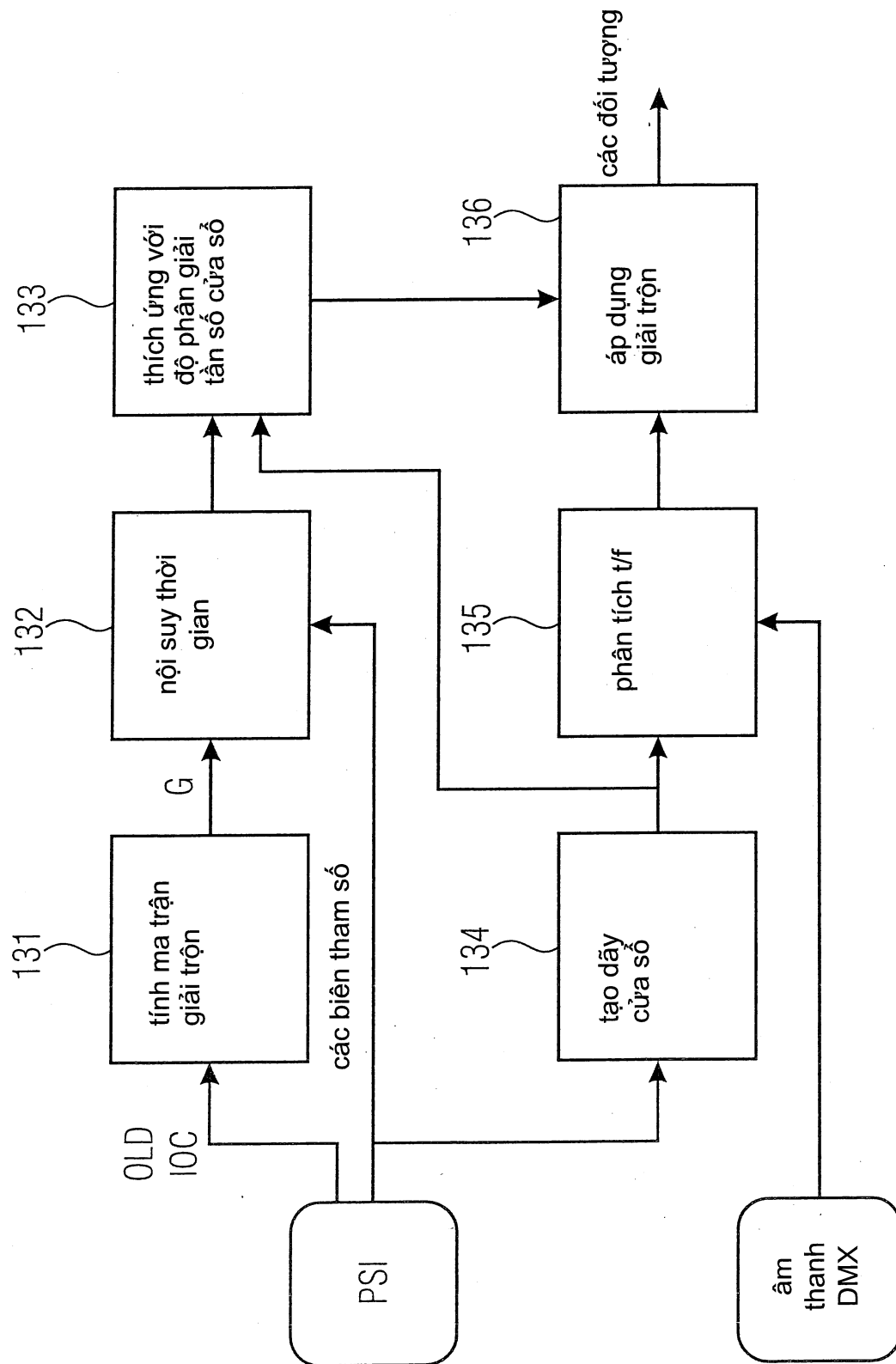


FIG. 6

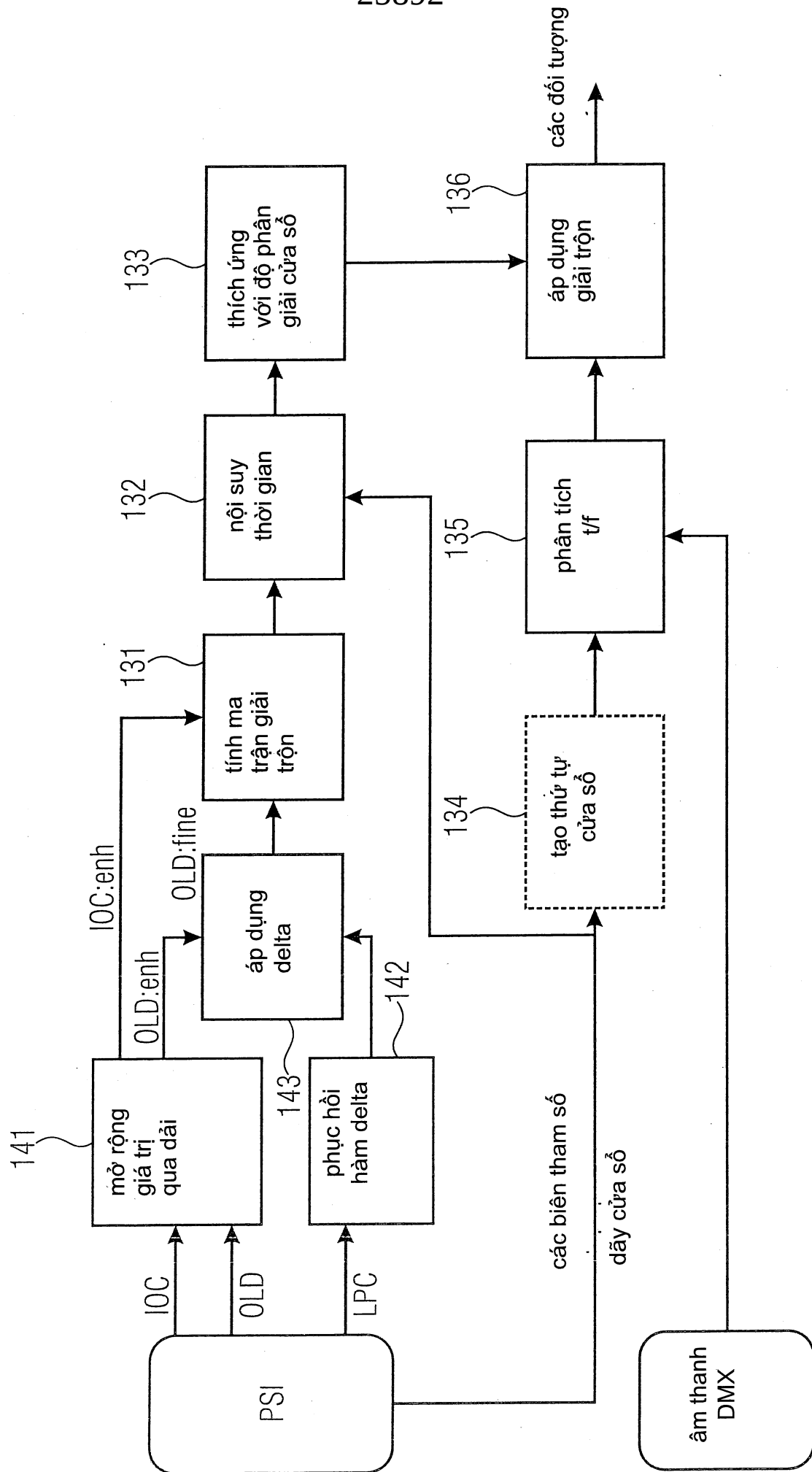


FIG. 7

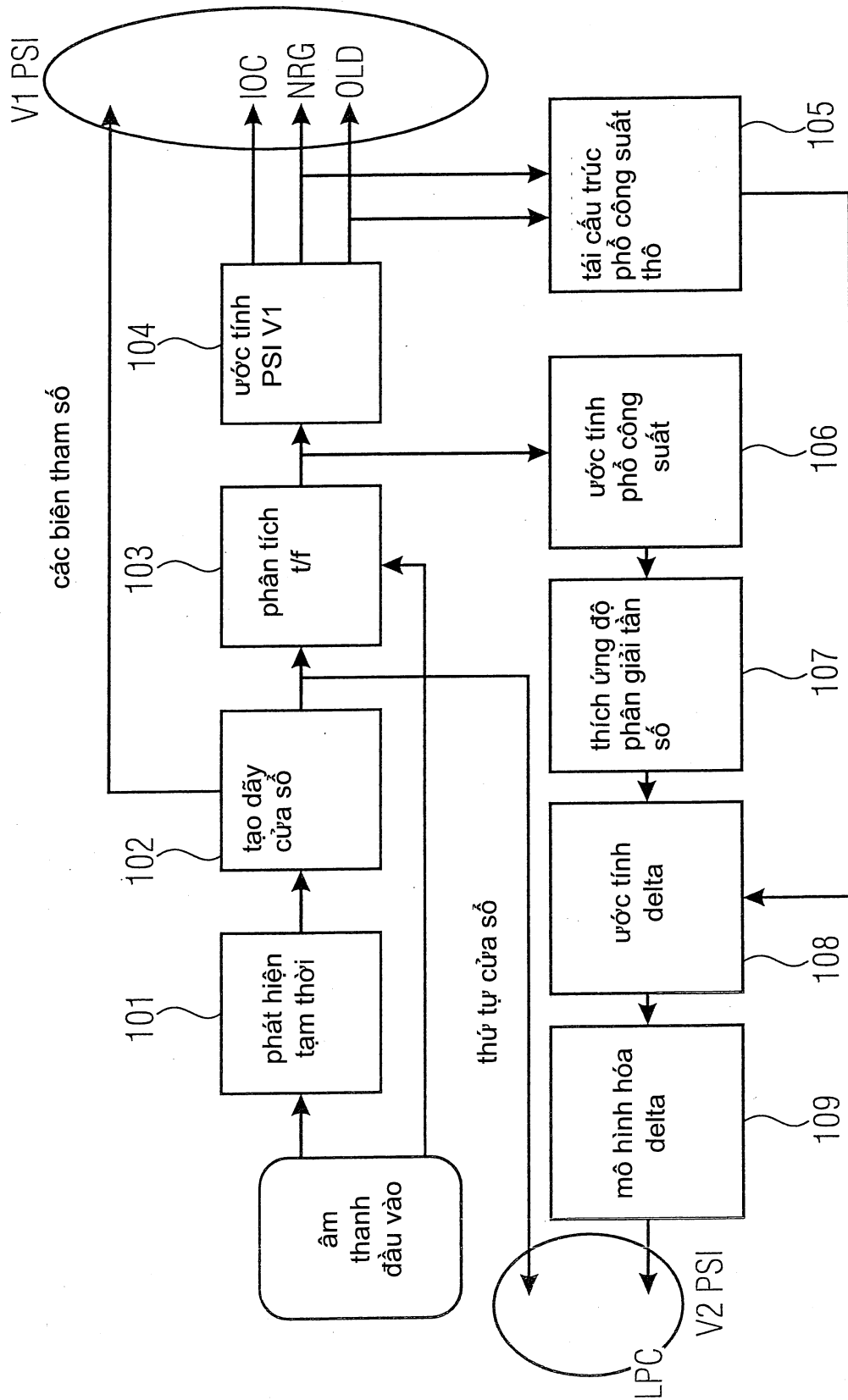


FIG. 8

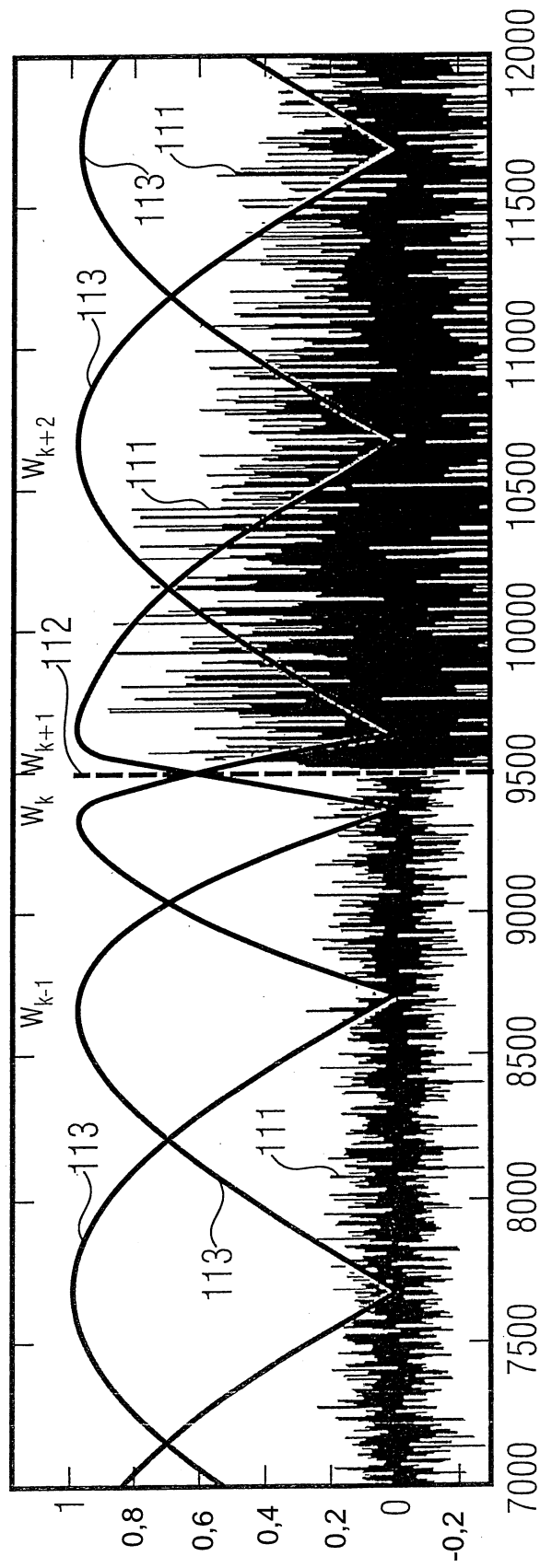


FIG.9

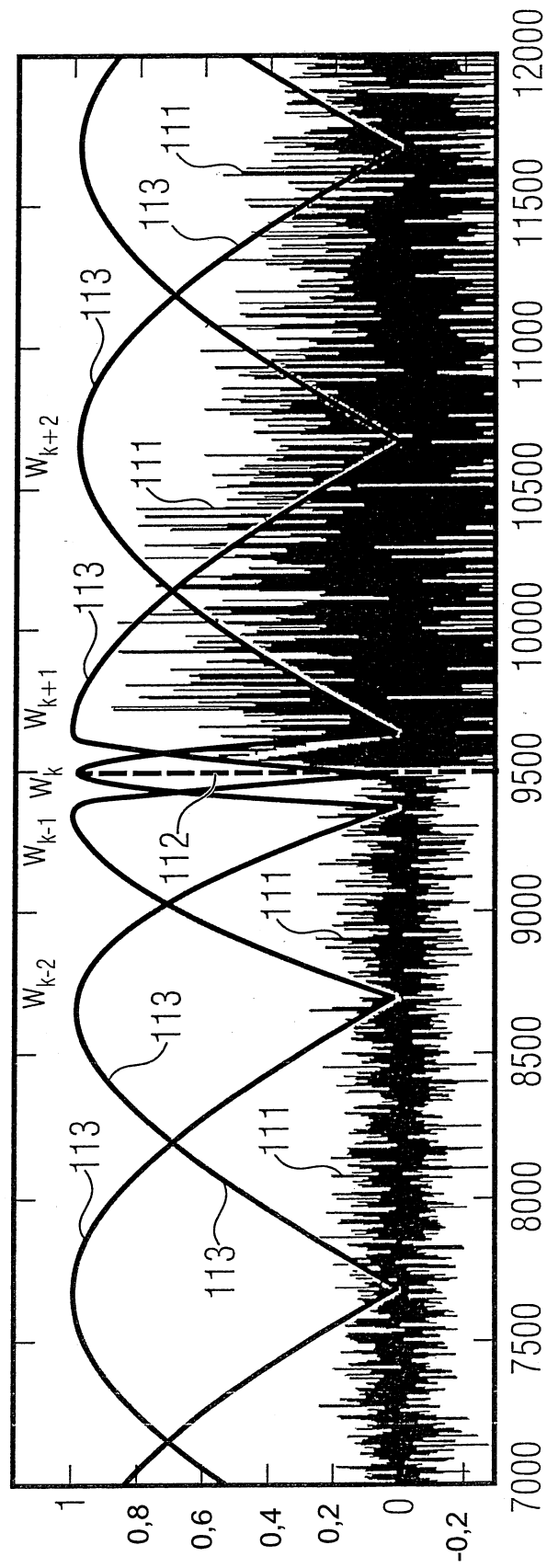


FIG.10

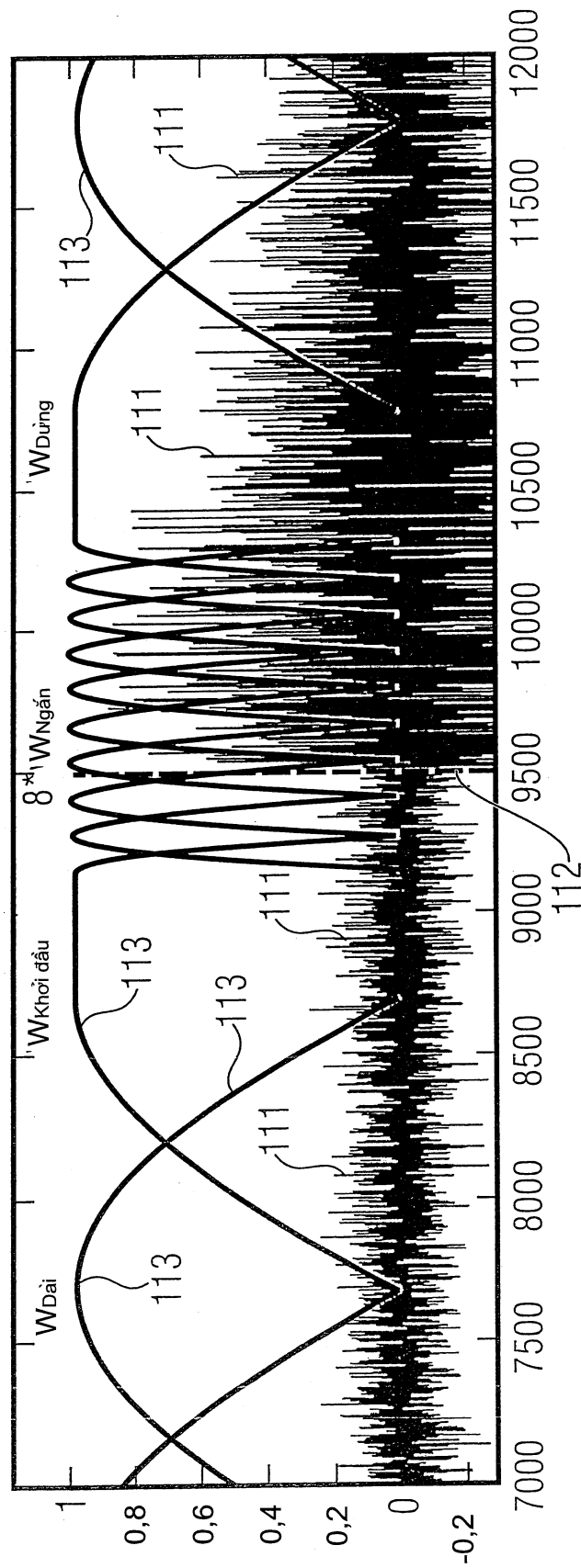


FIG.11

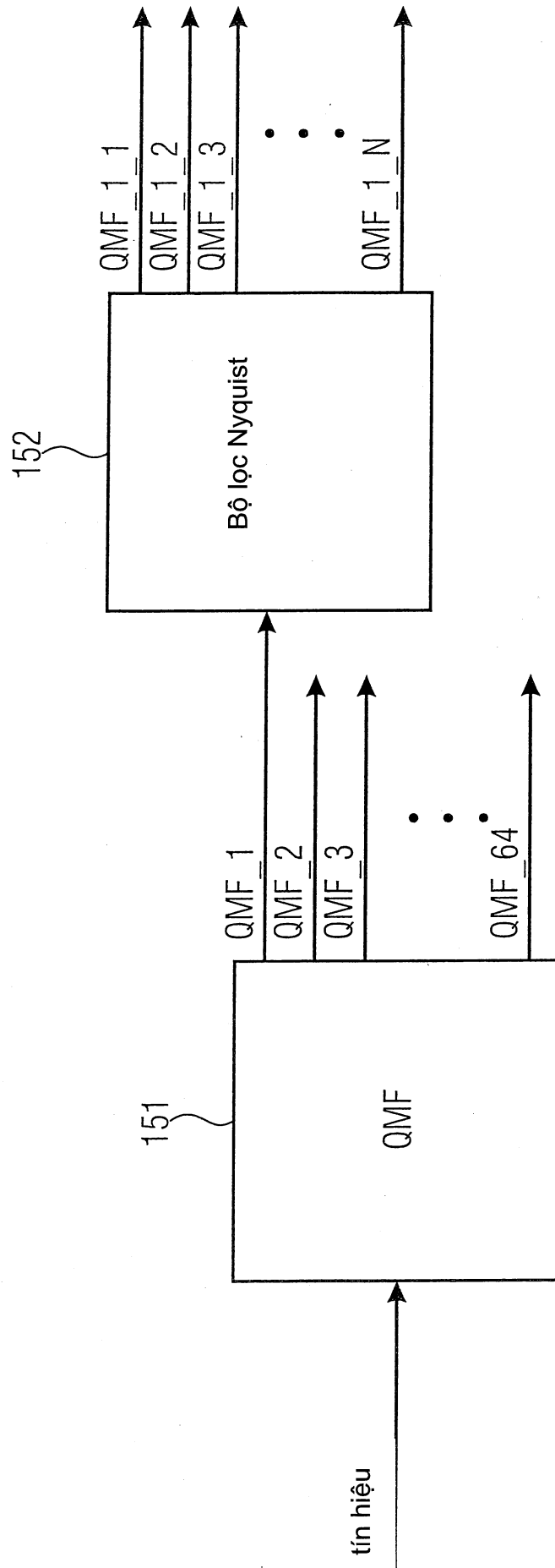


FIG.12

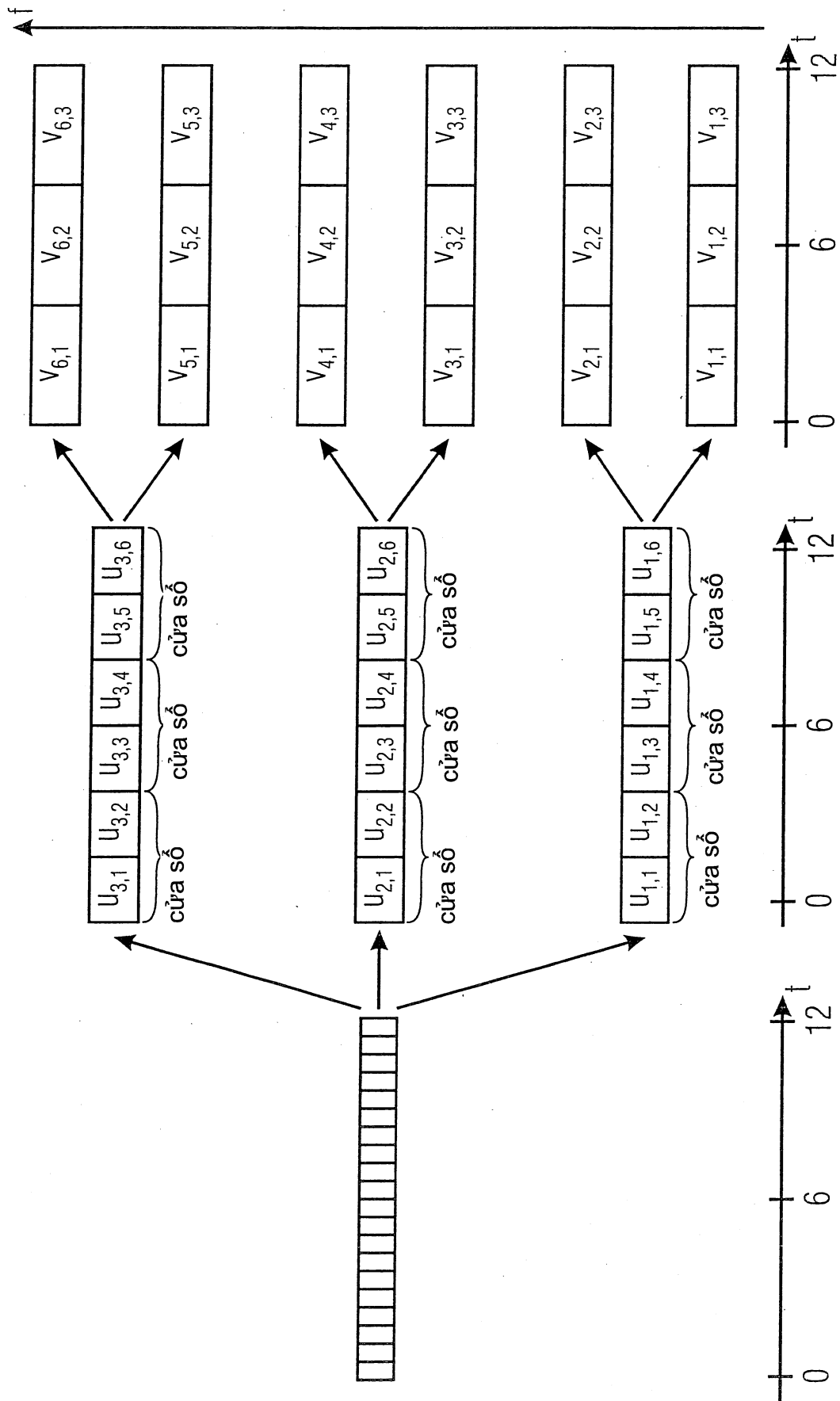


FIG.13

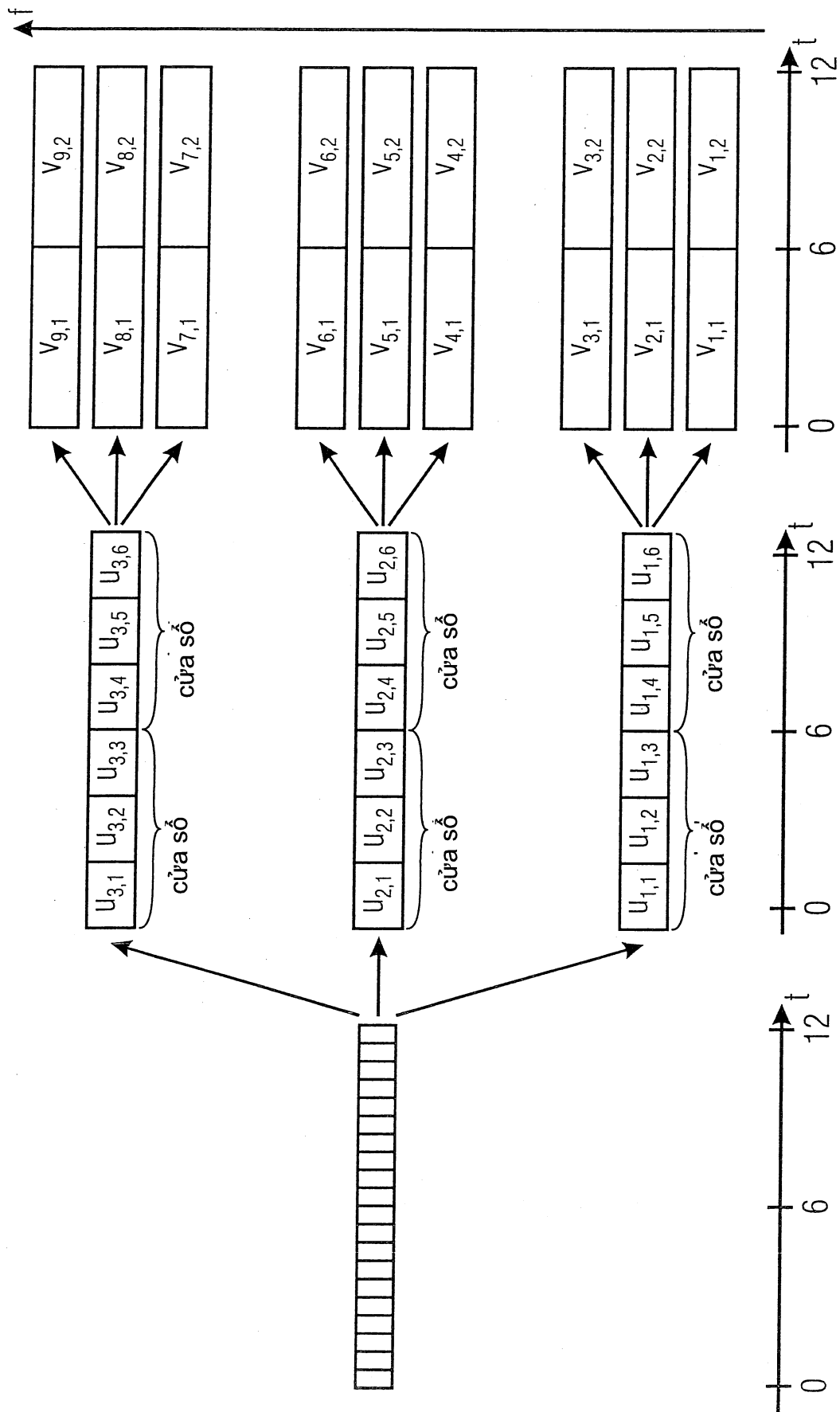


FIG.14

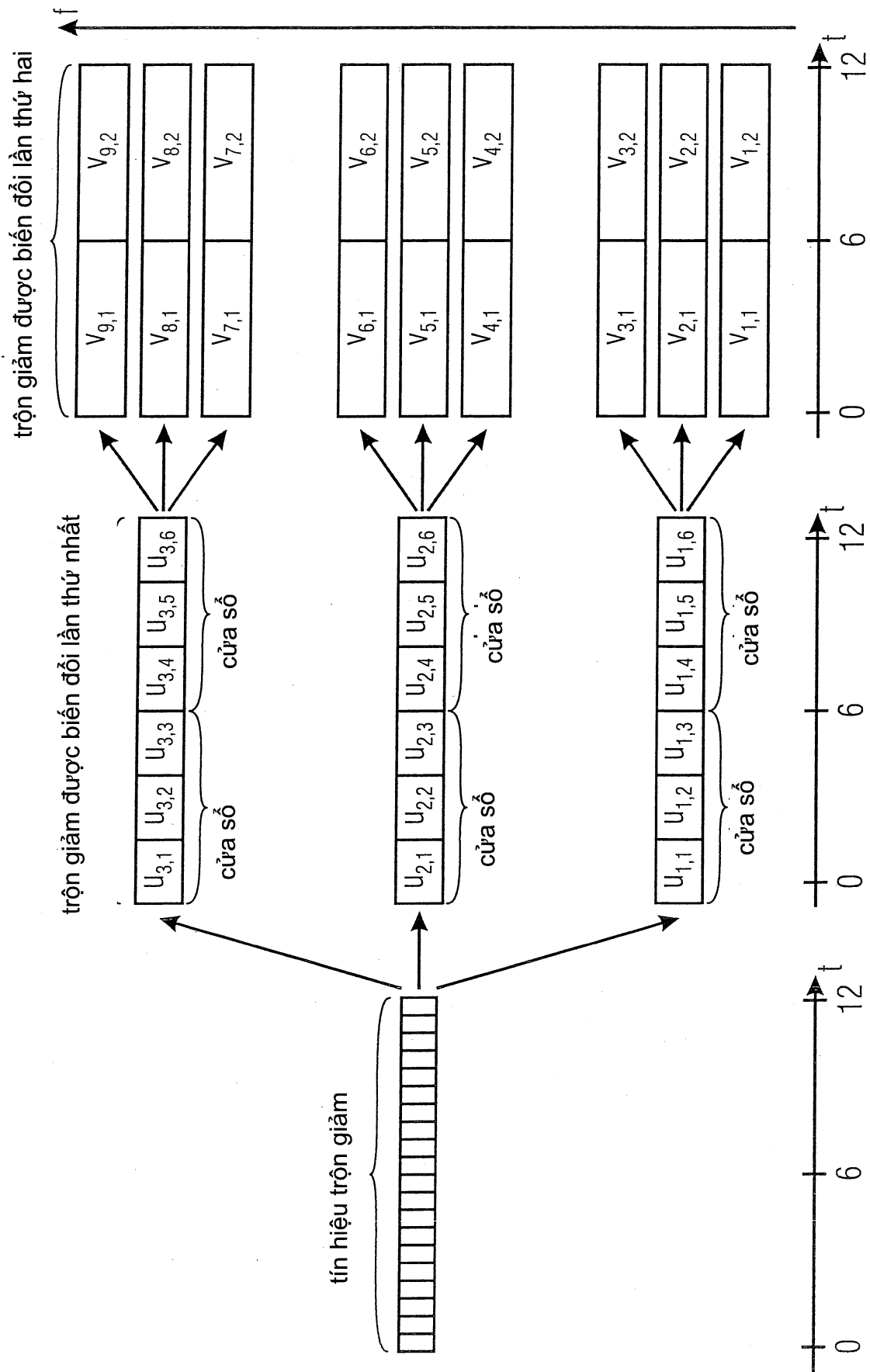


FIG.15

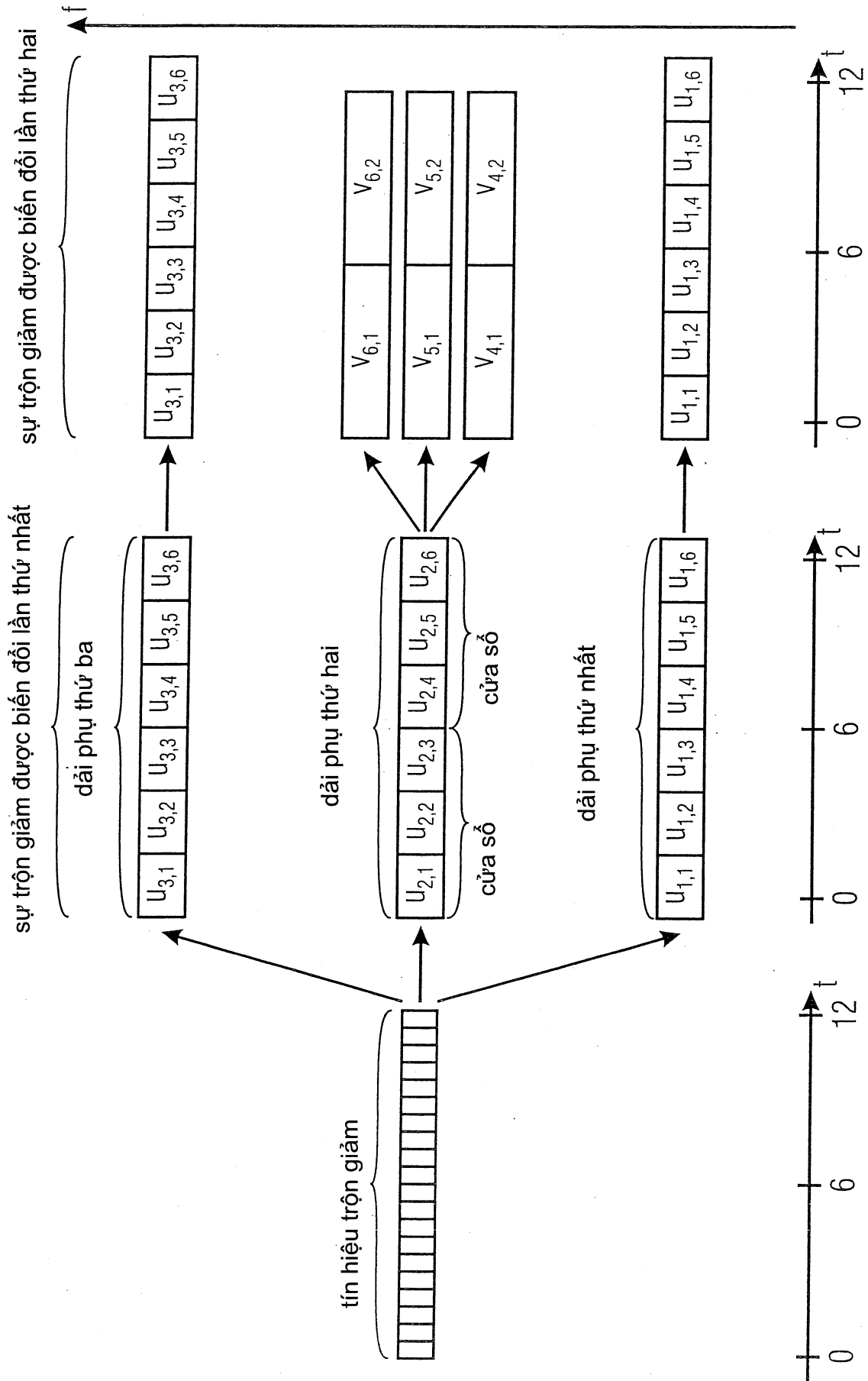


FIG.17

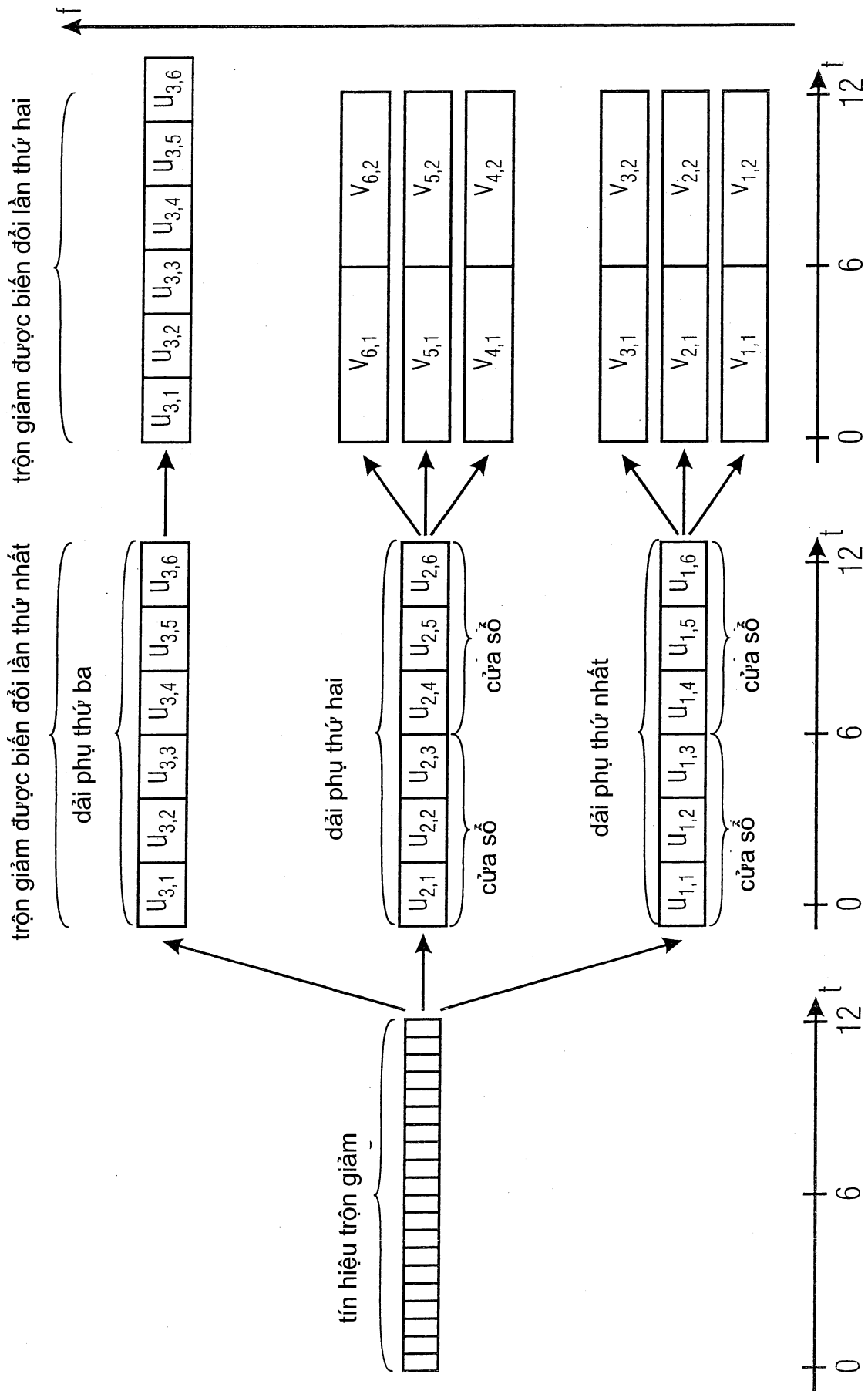


FIG.18