



(12) BẢN MÔ TẢ SÁNG CHẾ THUỘC BẢNG ĐỘC QUYỀN SÁNG CHẾ

(19) Cộng hòa xã hội chủ nghĩa Việt Nam (VN)
CỤC SỞ HỮU TRÍ TUỆ

(11)



1-0024062

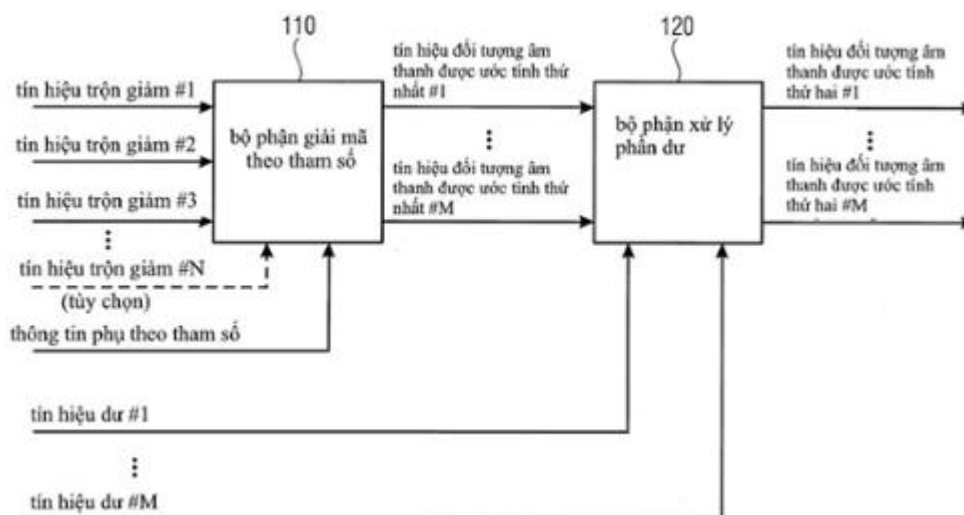
(51)⁷ G10L 19/008

(13) B

- (21) 1-2015-00740 (22) 16/04/2013
(86) PCT/EP2013/057932 16/04/2013 (87) WO2014/023443 13/02/2014
(30) 61/681,730 10/08/2012 US
(45) 25/06/2020 387 (43) 27/07/2015 328A
(73) Fraunhofer-Gesellschaft zur Foerderung der angewandten Forschung e.V. (DE)
Hansastrasse 27c, 80686 Muenchen, Germany
(72) KASTNER, Thorsten (DE); HERRE, Juergen (DE); PAULUS, Jouni (FI);
TERENTIV, Leon (DE); HELLMUTH, Oliver (DE); FUCHS, Harald (DE).
(74) Công ty Luật TNHH AMBYS Hà Nội (AMBYS HANOI)

(54) BỘ MÃ HÓA, BỘ GIẢI MÃ, HỆ THỐNG VÀ PHƯƠNG PHÁP MÃ HÓA ĐỐI TƯỢNG ÂM THANH THEO THAM SỐ SỬ DỤNG KHÁI NIỆM PHẦN DƯ

(57) Sáng chế đề cập đến bộ mã hóa, bộ giải mã, hệ thống và phương pháp mã hóa đối tượng âm thanh theo tham số sử dụng khái niệm phần dư. Bộ giải mã bao gồm bộ phận giải mã theo tham số (110) để tạo ra nhiều tín hiệu đối tượng âm thanh được ước tính thứ nhất bằng cách trộn tăng ba hoặc nhiều hơn ba tín hiệu trộn giảm, trong đó ba hoặc nhiều hơn ba tín hiệu trộn giảm mã hóa nhiều tín hiệu đối tượng âm thanh gốc, trong đó bộ phận giải mã theo tham số (110) được tạo cấu hình để trộn tăng ba hoặc nhiều hơn ba tín hiệu trộn giảm phụ thuộc vào thông tin phụ theo tham số biểu thị thông tin trên nhiều tín hiệu đối tượng âm thanh gốc. Ngoài ra, bộ giải mã bao gồm bộ phận xử lý phần dư (120) để tạo ra nhiều tín hiệu đối tượng âm thanh được ước tính thứ hai bằng cách biến đổi một hoặc nhiều tín hiệu đối tượng âm thanh được ước tính thứ nhất, trong đó bộ phận xử lý phần dư (120) được tạo cấu hình để biến đổi một hoặc nhiều các tín hiệu đối tượng âm thanh được ước tính thứ nhất đã nêu phụ thuộc vào một hoặc nhiều tín hiệu dư.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến mã hóa tín hiệu âm thanh, giải mã và xử lý, và đặc biệt đề cập đến bộ mã hóa, bộ giải mã và phương pháp mà sử dụng các khái niệm phân dư để mã hóa đối tượng âm thanh theo tham số.

Tình trạng kỹ thuật của sáng chế

Gần đây, các tham số kỹ thuật để truyền dẫn/lưu trữ tốc độ truyền dữ liệu hiệu quả các cảnh bao gồm nhiều đối tượng âm thanh đã được đề xuất trong lĩnh vực mã hóa âm thanh (xem ví dụ, [BCC], [JSC], [SAOC], [SAOC1] và [SAOC2]) và được thông báo sự phân tách nguồn (xem ví dụ, [ISS1], [ISS2], [ISS3], [ISS4], [ISS5] và [ISS6]). Những kỹ thuật này nhằm mục đích xây dựng lại cảnh âm thanh đầu ra mong muốn hoặc đối tượng nguồn âm thanh mong muốn trên cơ sở các thông tin phụ bổ sung mô tả cảnh âm thanh được truyền dẫn và/hoặc được lưu trữ và/hoặc các đối tượng nguồn âm thanh trong cảnh âm thanh.

Fig.5 minh họa mã hóa đối tượng âm thanh không gian (Spatial Audio Object Coding - SAOC) hệ thống minh họa khái quát nguyên tắc của các hệ tham số này sử dụng ví dụ nhóm các chuyên gia hình ảnh động SAOC (Moving Picture Experts Group - MPEG) (xem ví dụ, [SAOC], [SAOC1] và [SAOC2]).

Việc xử lý nói chung được thực hiện theo cách lựa chọn thời gian/tần số và có thể được mô tả như sau:

Bộ mã hóa SAOC 510, cụ thể, bộ ước tính thông tin phụ 530 của bộ mã hóa SAOC 510, trích thông tin phụ mô tả các đặc điểm của các tín hiệu đối tượng âm thanh đầu vào tối đa $32 s_1 \dots s_{32}$ (trong các hệ thức dạng đơn giản nhất của nó của các cường độ đối tượng của các tín hiệu đối tượng âm thanh). Bộ trộn 520 của bộ mã hóa SAOC 510 trộn giảm các tín hiệu đối tượng âm thanh $s_1 \dots s_{32}$ để đạt được hỗn hợp tín hiệu kênh đôi hoặc kênh đơn (ví dụ, một hoặc hai tín hiệu trộn giảm) sử dụng các nhân tổ khuếch đại trộn giảm $d_{1,1} \dots d_{32,2}$.

(Các) tín hiệu trộn giảm và thông tin phụ được truyền dẫn hoặc lưu trữ. Để đạt được điều này, (các) tín hiệu âm thanh trộn giảm có thể được mã hóa sử dụng bộ mã hóa âm thanh 540. Bộ mã hóa âm thanh 540 có thể là bộ mã hóa âm thanh thụ cảm nổi tiếng, ví dụ, bộ mã hóa âm thanh MPEG-1 Lớp II hoặc III (aka.mp3), bộ mã hóa âm thanh MPEG mã hóa âm thanh nâng cao (Advanced Audio Coding - AAC), v.v.

Ở phía bộ thu, bộ giải mã âm thanh tương ứng 550, ví dụ, bộ giải mã âm thanh thụ cảm, như bộ giải mã âm thanh MPEG-1 Lớp II hoặc III (aka.mp3), bộ giải mã âm thanh MPEG mã hóa âm thanh nâng cao (AAC), ví dụ, mã hóa (các) tín hiệu âm thanh trộn giảm được mã hóa.

Bộ giải mã SAOC 560 về mặt khái niệm cố gắng khôi phục các tín hiệu đối tượng (âm thanh) gốc ("sự phân tách đối tượng") từ một hoặc hai tín hiệu trộn giảm sử dụng thông tin phụ được truyền dẫn và/hoặc lưu trữ, ví dụ, bằng cách sử dụng bộ tách đối tượng ảo 570. Các tín hiệu đối tượng (âm thanh) gần đúng này $s_{1,est}...s_{32,est}$ sau đó được trộn bằng bộ biểu diễn 580 của bộ giải mã SAOC 560 thành cảnh mục tiêu được biểu diễn bằng các kênh đầu ra 6 âm thanh $y_{1,est}...y_{6,est}$ sử dụng ma trận biểu diễn (được mô tả bởi các hệ số $r_{1,1} ... r_{32,6}$). Đầu ra có thể là kênh đơn, kênh đôi âm lập thể hoặc cảnh mục tiêu đa kênh 5.1 (ví dụ, một, hai hoặc sáu tín hiệu đầu ra âm thanh).

Do những hạn chế cơ bản của sự ước tính tham số của các đối tượng âm thanh tại phía giải mã; trong hầu hết các trường hợp, cảnh đầu ra mục tiêu mong muốn có thể không được tạo ra một cách ưu việt. Tại các điểm hoạt động cực trị (ví dụ, phát lại một mình một đối tượng âm thanh), thường xuyên, việc xử lý không còn có thể đạt được âm thanh chủ đầy đủ. Để đạt được điều này, hệ thống SAOC đã được mở rộng bằng cách đưa vào các đối tượng âm thanh nâng cao (Enhanced Audio Object - EAO) (xem ví dụ, [Dfx], xem ví dụ hơn nữa, [SAOC]). Các đối tượng âm thanh mà được mã hóa dưới dạng các EAO thể hiện khả năng tách tăng từ các đối tượng âm thanh không nâng cao (thông thường) (non-Enhanced Audio Objects – các non-EAO) khác được mã hóa trong tín hiệu trộn giảm tương tự với hao phí tỷ lệ thông tin phụ tăng. Khái niệm EAO xem xét cho từng EAO lỗi dự báo (tín hiệu dư) của mô hình tham số.

Fig.6 mô tả sự ước tính dư tại phía bộ mã hóa, biểu đồ minh họa sự tính toán của các tín hiệu dư cho từng EAO. Trong bộ mã hóa SAOC, các tín hiệu dư (tăng đến 4

EAO) được ước tính sử dụng Thông tin phụ theo tham số (Parametric Side Information - PSI) được trích xuất và các tín hiệu nguồn gốc, dạng sóng được mã hóa và chứa trong dòng dữ liệu như thông tin phụ dư (Residual Side Information - RSI) không tham số. Cụ thể hơn, bộ giải mã PSI SAOC cho các EAO 610 tạo ra các tín hiệu đối tượng âm thanh ước tính $S_{est,EAO}$ từ X trộn giảm. Bộ phận tạo RSI 620 sau đó tạo lên đến bốn tín hiệu dư $S_{res,RSI,\{1,...,4\}}$ dựa trên các tín hiệu đối tượng âm thanh được ước tính đã được tạo $S_{est,EAO}$ và dựa trên các tín hiệu đối tượng âm thanh EAO gốc $s_1, \dots, s_4$.

Fig.7 mô tả cấu trúc cơ bản của bộ giải mã SAOC với hỗ trợ EAO, minh họa tổng quan khái niệm của biểu đồ xử lý EAO được tích hợp vào chuỗi giải mã/chuyển mã SAOC (chuyển mã: chuyển đổi dữ liệu từ một mã hóa thành một mã hóa khác).

Các tham số định hướng tín hiệu trộn giảm, cụ thể, các hệ số dự báo kênh (Channel Prediction Coefficient - CPC) được suy ra từ thông tin phụ theo tham số (Parametric Side Info - PSI) bằng bộ phận ước tính CPC 710.

Các CPC cùng với tín hiệu trộn giảm được cung cấp vào trong Hai-đến N-hộp (Two-to-N-box - TTN-box) 720. TTN-box 720 về mặt khái niệm cố gắng ước tính các EAO ($S_{est,EAO}$) từ tín hiệu trộn giảm được truyền và để cung cấp trộn giảm non-EAO đã ước tính ($X_{est, nonEAO}$) chỉ chứa các non-EAO.

Các tín hiệu dư được truyền dẫn/lưu trữ (và giải mã) ($S_{res, RSI}$) được sử dụng bởi bộ phận xử lý RSI 730 để tăng các ước tính của các EAO ($S_{est, EAO}$) và trộn giảm tương ứng chỉ các đối tượng non-EAO (X_{nonEAO}).

Theo tình trạng kỹ thuật, trong bước tiếp theo, bộ phận xử lý RSI 730 nạp tín hiệu trộn giảm non-EAO (X_{nonEAO}) vào trong bộ xử lý trộn giảm SAOC (bộ phận giải mã PSI) 740 để ước tính các đối tượng non-EAO $S_{est,nonEAO}$. Bộ phận giải mã PSI 740 chuyển các đối tượng âm thanh non-EAO đã ước tính $S_{est,nonEAO}$ đến bộ phận biểu diễn 750. Hơn nữa, bộ phận xử lý RSI nạp trực tiếp các EAO được nâng cao $\hat{s}_{est,EAO}$ vào trong bộ phận biểu diễn 750. Bộ phận biểu diễn 750 sau đó tạo các tín hiệu đầu ra âm lập thể hoặc đơn thể dựa trên các đối tượng âm thanh non-EAO được ước tính $S_{est,nonEAO}$ và dựa trên các EAO được nâng cao $\hat{s}_{est,EAO}$.

Hệ thống thuộc tình trạng kỹ thuật có các nhược điểm sau đây:

Trước khi các tín hiệu dư được đưa vào để tính toán các EAO trong bộ giải mã SAOC, CPCs được định hướng trộn giảm phải được tính từ thông tin phụ theo tham số được truyền dẫn/lưu trữ.

Tất cả các tín hiệu trộn giảm phải được xử lý trong khái niệm phần dư SAOC bất chấp sự hữu dụng của nó đối với việc xử lý EAO.

Khái niệm phần dư SAOC có thể chỉ được sử dụng với các hỗn hợp tín hiệu kênh đôi hoặc kênh đơn do sự hạn chế của TTN-box. Khái niệm phần dư EAO có thể không được sử dụng kết hợp với các hỗn hợp kênh đa (ví dụ hỗn hợp kênh đa 5.1).

Hơn nữa, do sự phức tạp tính toán tương ứng của việc ước tính của chúng, việc xử lý EAO SAOC thiết lập các giới hạn số lượng EAO (tức là, lên đến 4).

Do các giới hạn này, khái niệm xử lý phần dư SAOC EAO có thể không được áp dụng cho các tín hiệu trộn giảm kênh đa (ví dụ, 5.1) hoặc được sử dụng cho nhiều hơn 4 EAO.

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.

[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

[Dfx] C. Falch and L. Terentiev and J. Herre: “Spatial Audio Object Coding with Enhanced Audio Object Separation”, 10th International Conference on Digital Audio Effects, 2010

Bản chất kỹ thuật của sáng chế

Mục đích của sáng chế là đề xuất các khái niệm được cải tiến để mã hóa tín hiệu âm thanh, giải mã tín hiệu âm thanh và xử lý tín hiệu âm thanh. Mục đích của sáng chế được làm sáng tỏ bởi bộ giải mã theo điểm 1, bởi bộ tạo tín hiệu dư theo điểm 11, bởi

bộ mã hóa theo điểm 19, bởi hệ thống theo điểm 21, bởi tín hiệu được mã hóa theo điểm 22, bởi phương pháp theo điểm 23, bởi phương pháp theo điểm 24 và bởi chương trình máy tính theo điểm 25.

Bộ giải mã được đề xuất. Bộ giải mã bao gồm bộ phận giải mã theo tham số để tạo ra nhiều tín hiệu đối tượng âm thanh được ước tính thứ nhất bằng cách trộn tăng ba hoặc nhiều hơn ba tín hiệu trộn giảm, trong đó ba hoặc nhiều hơn ba tín hiệu trộn giảm mã hóa nhiều tín hiệu đối tượng âm thanh gốc, trong đó bộ phận giải mã theo tham số được tạo cấu hình để trộn tăng ba hoặc nhiều hơn ba tín hiệu trộn giảm phụ thuộc vào thông tin phụ theo tham số biểu thị thông tin trên nhiều tín hiệu đối tượng âm thanh gốc. Ngoài ra, bộ giải mã bao gồm bộ phận xử lý phần dư để tạo ra nhiều tín hiệu đối tượng âm thanh được ước tính thứ hai bằng cách biến đổi một hoặc nhiều tín hiệu đối tượng âm thanh được ước tính thứ nhất, trong đó bộ phận xử lý phần dư được tạo cấu hình để biến đổi một hoặc nhiều tín hiệu đối tượng âm thanh được ước tính thứ nhất đã nêu phụ thuộc vào một hoặc nhiều tín hiệu dư.

Phương án thể hiện khái niệm phần dư được định hướng đối tượng trong đó cải thiện chất lượng được thụ cảm của các EAO. Không giống như hệ thống của tình trạng kỹ thuật, khái niệm được đưa ra là không bị giới hạn về số lượng tín hiệu trộn giảm cũng không bị giới hạn về số lượng các EAO. Hai phương pháp để nhận được các tín hiệu dư liên quan đến đối tượng được trình bày. Khái niệm nổi bật trong đó năng lượng của tín hiệu dư được giảm lặp lại với việc tăng số lượng các EAO tại giá trị của sự phức tạp tính toán cao hơn và khái niệm thứ hai với sự phức tạp tính toán ít hơn trong đó tất cả phần dư được ước tính đồng thời.

Hơn nữa, các phương án đề xuất khái niệm được cải tiến của việc áp dụng các tín hiệu dư được định hướng đối tượng tại phía giải mã, và các khái niệm với sự phức tạp được giảm được thiết kế để các cảnh ứng dụng mà trong đó chỉ các EAO được đánh tín hiệu tại phía giải mã, hoặc việc biến đổi non-EAO được giới hạn tỉ lệ có lợi.

Theo phương án, bộ phận xử lý phần dư có thể được tạo cấu hình để biến đổi một hoặc nhiều tín hiệu đối tượng âm thanh được ước tính thứ nhất đã nêu phụ thuộc vào ít nhất ba tín hiệu dư. Bộ giải mã được làm thích ứng để tạo ít nhất ba kênh đầu ra âm thanh dựa trên nhiều tín hiệu đối tượng âm thanh được ước tính thứ hai.

Theo phương án, bộ giải mã có thể còn bao gồm bộ phận biến đổi trộn giảm. Bộ phận biến đổi trộn giảm có thể xác định một hoặc nhiều tín hiệu đối tượng âm thanh của nhiều tín hiệu đối tượng âm thanh được ước tính thứ hai. Bộ phận biến đổi trộn giảm có thể được làm thích ứng để loại bỏ một hoặc nhiều tín hiệu đối tượng âm thanh được ước tính thứ hai được xác định từ ba hoặc nhiều hơn ba tín hiệu trộn giảm để thu được ba hoặc nhiều hơn ba tín hiệu trộn giảm được biến đổi. Bộ phận giải mã theo tham số có thể được tạo cấu hình để xác định một hoặc nhiều tín hiệu đối tượng âm thanh của các tín hiệu đối tượng âm thanh được ước tính thứ nhất dựa trên ba hoặc nhiều hơn ba tín hiệu trộn giảm được biến đổi.

Trong phương án cụ thể, bộ phận biến đổi trộn giảm có thể, ví dụ, được làm thích ứng để áp dụng công thức $\tilde{\mathbf{X}}_{nonEAO} = \mathbf{X} - \mathbf{D}\mathbf{Z}_{eao}^* \mathbf{S}_{eao}$.

Hơn nữa, bộ giải mã có thể được tạo cấu hình để điều khiển hai hoặc nhiều bước lặp lại. Với mỗi bước lặp lại, bộ phận giải mã theo tham số có thể được làm thích ứng để xác định chính xác một tín hiệu đối tượng âm thanh trong số nhiều tín hiệu đối tượng âm thanh được ước tính thứ nhất. Hơn nữa, với bước lặp lại đã nêu, bộ phận xử lý phân dư có thể được làm thích ứng để xác định chính xác một tín hiệu đối tượng âm thanh của nhiều tín hiệu đối tượng âm thanh được ước tính thứ hai bằng cách biến đổi tín hiệu đối tượng âm thanh đã nêu trong số nhiều tín hiệu đối tượng âm thanh được ước tính thứ nhất. Hơn nữa, với bước lặp lại đã nêu, bộ phận biến đổi trộn giảm có thể được làm thích ứng để loại bỏ tín hiệu đối tượng âm thanh đã nêu trong số nhiều tín hiệu đối tượng âm thanh được ước tính thứ hai từ ba hoặc nhiều hơn ba tín hiệu trộn giảm để biến đổi ba hoặc nhiều hơn tín hiệu trộn giảm. Trong bước lặp lại tiếp theo sau bước lặp lại đã nêu, bộ phận giải mã theo tham số có thể được làm thích ứng để xác định chính xác một tín hiệu đối tượng âm thanh trong số nhiều tín hiệu đối tượng âm thanh được ước tính thứ nhất dựa trên ba hoặc nhiều hơn tín hiệu trộn giảm mà đã được biến đổi.

Theo phương án, mỗi một hoặc nhiều tín hiệu dư có thể biểu thị sự khác biệt giữa một tín hiệu trong số nhiều tín hiệu đối tượng âm thanh gốc và một tín hiệu trong số một hoặc nhiều tín hiệu đối tượng âm thanh được ước tính thứ nhất.

Theo phương án, trong đó bộ phận xử lý phần dư có thể được tạo cấu hình để tạo ra nhiều tín hiệu đối tượng âm thanh được ước tính thứ hai bằng cách biến đổi năm hoặc nhiều hơn năm tín hiệu đối tượng âm thanh được ước tính thứ nhất, trong đó bộ phận xử lý phần dư được tạo cấu hình để biến đổi năm hoặc nhiều hơn năm các tín hiệu đối tượng âm thanh được ước tính thứ nhất phụ thuộc vào năm hoặc nhiều hơn năm tín hiệu dư.

Trong một phương án khác, bộ giải mã có thể được tạo cấu hình để tạo bảy hoặc nhiều hơn bảy kênh đầu ra âm thanh dựa trên nhiều tín hiệu đối tượng âm thanh được ước tính thứ hai.

Theo phương án nữa, bộ giải mã có thể được làm thích ứng để không xác định các hệ số dự báo kênh để xác định nhiều tín hiệu đối tượng âm thanh được ước tính thứ hai. Các phương án đề xuất các khái niệm để việc tính toán các hệ số dự báo kênh mà cho đến nay là cần thiết để giải mã trong tình trạng kỹ thuật SAOC, là không cần thiết nữa để giải mã.

Trong phương án nữa, bộ giải mã có thể là bộ giải mã SAOC.

Ngoài ra, bộ tạo tín hiệu dư được đề xuất. Bộ tạo tín hiệu dư bao gồm bộ phận giải mã theo tham số để tạo ra nhiều tín hiệu đối tượng âm thanh được ước tính thứ nhất bằng cách trộn tăng ba hoặc nhiều hơn ba tín hiệu trộn giảm, trong đó ba hoặc nhiều hơn ba tín hiệu trộn giảm mã hóa nhiều tín hiệu đối tượng âm thanh gốc, trong đó bộ phận giải mã theo tham số được tạo cấu hình để trộn tăng ba hoặc nhiều hơn ba tín hiệu trộn giảm phụ thuộc vào thông tin phụ theo tham số biểu thị thông tin trên nhiều tín hiệu đối tượng âm thanh gốc. Ngoài ra, bộ tạo tín hiệu dư bao gồm bộ phận ước tính phần dư để tạo ra nhiều tín hiệu dư dựa trên nhiều tín hiệu đối tượng âm thanh gốc và dựa trên nhiều tín hiệu đối tượng âm thanh được ước tính, như vậy mà mỗi một tín hiệu trong số nhiều tín hiệu dư là tín hiệu khác nhau biểu thị sự khác biệt giữa một tín hiệu trong số nhiều tín hiệu đối tượng âm thanh gốc và một tín hiệu trong số nhiều tín hiệu đối tượng âm thanh được ước tính.

Trong phương án, bộ phận ước tính phần dư có thể được làm thích ứng để tạo ra ít nhất năm tín hiệu dư dựa trên ít nhất năm tín hiệu đối tượng âm thanh gốc của nhiều

tín hiệu đối tượng âm thanh gốc và dựa trên ít nhất năm tín hiệu đối tượng âm thanh được ước tính của nhiều tín hiệu đối tượng âm thanh được ước tính.

Trong phương án, bộ tạo tín hiệu dư có thể còn bao gồm bộ phận biến đổi trộn giảm đang được làm thích ứng để biến đổi ba hoặc nhiều hơn ba tín hiệu trộn giảm để thu được ba hoặc nhiều hơn ba tín hiệu trộn giảm được biến đổi. Bộ phận giải mã theo tham số có thể được tạo cấu hình để xác định một hoặc nhiều tín hiệu đối tượng âm thanh của các tín hiệu đối tượng âm thanh được ước tính thứ nhất dựa trên ba hoặc nhiều hơn ba tín hiệu trộn giảm được biến đổi.

Trong phương án, bộ phận biến đổi trộn giảm có thể, ví dụ, được tạo cấu hình để biến đổi ba hoặc nhiều hơn ba tín hiệu trộn giảm gốc để thu được ba hoặc nhiều hơn ba tín hiệu trộn giảm được biến đổi, bằng cách loại bỏ một hoặc nhiều tín hiệu trong số nhiều tín hiệu đối tượng âm thanh gốc từ ba hoặc nhiều hơn ba tín hiệu trộn giảm gốc.

Trong một phương án khác, bộ phận biến đổi trộn giảm có thể, ví dụ, được tạo cấu hình để biến đổi ba hoặc nhiều hơn ba tín hiệu trộn giảm gốc để thu được ba hoặc nhiều hơn ba tín hiệu trộn giảm được biến đổi bằng cách tạo ra một hoặc nhiều tín hiệu đối tượng âm thanh được biến đổi dựa trên một hoặc nhiều tín hiệu đối tượng âm thanh được ước tính và dựa trên một hoặc nhiều tín hiệu dư, và bằng cách loại bỏ một hoặc nhiều tín hiệu đối tượng âm thanh được biến đổi từ ba hoặc nhiều hơn ba tín hiệu trộn giảm gốc. Ví dụ, mỗi một hoặc nhiều tín hiệu đối tượng âm thanh được biến đổi có thể được tạo ra bởi bộ phận biến đổi trộn giảm bằng cách biến đổi một tín hiệu đối tượng âm thanh được ước tính, trong đó bộ phận biến đổi trộn giảm có thể được làm thích ứng để biến đổi tín hiệu đối tượng âm thanh được ước tính đã nêu phụ thuộc vào một tín hiệu dư trong số một hoặc nhiều tín hiệu dư.

Trong cả hai phương án được mô tả trên, bộ phận biến đổi trộn giảm có thể, ví dụ, được làm thích ứng để áp dụng công thức $\tilde{\mathbf{X}} = \mathbf{X} - \mathbf{D}\mathbf{Z}_{eao}^* \mathbf{S}_{eao}$, trong đó $\mathbf{X}$ là trộn giảm được biến đổi, trong đó $\mathbf{D}$ biểu thị thông tin trộn giảm, trong đó $\mathbf{S}_{eao}$ bao gồm các tín hiệu đối tượng âm thanh gốc được loại bỏ hoặc các tín hiệu đối tượng âm thanh được biến đổi, trong đó $\mathbf{Z}_{eao}^*$ biểu thị các vị trí của các tín hiệu được loại bỏ, và trong đó $\tilde{\mathbf{X}}$ là tín hiệu trộn giảm được biến đổi. Ví dụ, điểm (vị trí) của tín hiệu đối tượng

âm thanh tương ứng với điểm (vị trí) của đối tượng âm thanh của nó trong danh sách tất cả các đối tượng.

Theo phương án, bộ tạo tín hiệu dư có thể được làm thích ứng để điều khiển hai hoặc nhiều hơn hai bước lặp lại. Với mỗi bước lặp lại, bộ phận giải mã theo tham số có thể được làm thích ứng để xác định chính xác một tín hiệu đối tượng âm thanh trong số nhiều tín hiệu đối tượng âm thanh được ước tính. Hơn nữa, với bước lặp lại đã nêu, bộ phận ước tính phần dư có thể được làm thích ứng để xác định chính xác một tín hiệu dư trong số nhiều tín hiệu dư bằng cách biến đổi tín hiệu đối tượng âm thanh đã nêu của nhiều tín hiệu đối tượng âm thanh được ước tính. Hơn nữa, với bước lặp lại đã nêu, bộ phận biến đổi trộn giảm có thể được làm thích ứng để biến đổi ba hoặc nhiều hơn ba tín hiệu trộn giảm. Trong bước lặp lại tiếp theo sau bước lặp lại đã nêu, bộ phận giải mã theo tham số có thể được làm thích ứng để xác định chính xác một tín hiệu đối tượng âm thanh trong số nhiều tín hiệu đối tượng âm thanh được ước tính dựa trên ba hoặc nhiều hơn ba tín hiệu trộn giảm mà đã được biến đổi.

Trong phương án, bộ mã hóa để mã hóa nhiều tín hiệu đối tượng âm thanh gốc bằng cách tạo ra ba hoặc nhiều hơn ba tín hiệu trộn giảm, bằng cách tạo ra thông tin phụ theo tham số và bằng cách tạo ra nhiều tín hiệu dư được đề xuất. Bộ mã hóa bao gồm bộ tạo trộn giảm để cung cấp ba hoặc nhiều hơn ba tín hiệu trộn giảm biểu thị sự trộn giảm của nhiều tín hiệu đối tượng âm thanh gốc. Ngoài ra, bộ mã hóa bao gồm bộ ước tính thông tin phụ theo tham số để tạo ra thông tin phụ theo tham số biểu thị thông tin trên nhiều tín hiệu đối tượng âm thanh gốc, để thu được thông tin phụ theo tham số. Hơn nữa, bộ mã hóa bao gồm bộ tạo tín hiệu dư theo một trong số các phương án được mô tả ở trên. Bộ phận giải mã theo tham số của bộ tạo tín hiệu dư được làm thích ứng để tạo ra nhiều tín hiệu đối tượng âm thanh được ước tính bằng cách trộn tăng ba hoặc nhiều hơn ba tín hiệu trộn giảm được cung cấp bởi bộ tạo trộn giảm, trong đó các tín hiệu trộn giảm mã hóa nhiều tín hiệu đối tượng âm thanh gốc. Bộ phận giải mã theo tham số được tạo cấu hình để trộn tăng ba hoặc nhiều hơn ba tín hiệu trộn giảm phụ thuộc vào thông tin phụ theo tham số được tạo bởi bộ ước tính thông tin phụ theo tham số. Bộ phận ước tính phần dư của bộ tạo tín hiệu dư được làm thích ứng để tạo ra nhiều tín hiệu dư dựa trên nhiều tín hiệu đối tượng âm thanh gốc và dựa trên nhiều tín hiệu đối tượng âm thanh được ước tính, như vậy mà mỗi một tín hiệu trong số nhiều

tín hiệu dư biểu thị sự khác biệt giữa một tín hiệu trong số nhiều tín hiệu đối tượng âm thanh gốc và một tín hiệu trong số nhiều tín hiệu đối tượng âm thanh được ước tính.

Trong phương án, bộ mã hóa có thể là bộ mã hóa SAOC.

Ngoài ra, hệ thống mã hóa đối tượng âm thanh theo tham số sử dụng khái niệm phân dư được đề xuất. Hệ thống bao gồm bộ mã hóa theo một trong số các phương án được mô tả ở trên để mã hóa nhiều tín hiệu đối tượng âm thanh gốc bằng cách tạo ra ba hoặc nhiều hơn ba tín hiệu trộn giảm, bằng cách tạo ra thông tin phụ theo tham số và bằng cách tạo ra nhiều tín hiệu dư. Ngoài ra, hệ thống bao gồm bộ giải mã theo một trong số các phương án được mô tả ở trên, trong đó bộ giải mã được tạo cấu hình để tạo ra nhiều kênh đầu ra âm thanh dựa trên ba hoặc nhiều hơn ba tín hiệu trộn giảm được tạo ra bởi bộ mã hóa, dựa trên thông tin phụ theo tham số đang được tạo ra bởi bộ mã hóa và dựa trên nhiều tín hiệu dư được tạo bởi bộ mã hóa.

Ngoài ra, tín hiệu âm thanh được mã hóa được đề xuất. Tín hiệu âm thanh được mã hóa bao gồm ba hoặc nhiều hơn ba tín hiệu trộn giảm, thông tin phụ theo tham số và nhiều tín hiệu dư. Ba hoặc nhiều hơn ba tín hiệu trộn giảm là sự trộn giảm của nhiều tín hiệu đối tượng âm thanh gốc. Thông tin phụ theo tham số bao gồm các tham số biểu thị thông tin phụ trên nhiều tín hiệu đối tượng âm thanh gốc. Mỗi một tín hiệu trong số nhiều tín hiệu dư là tín hiệu khác nhau biểu thị sự khác biệt giữa một tín hiệu của nhiều tín hiệu âm thanh gốc và một tín hiệu của nhiều tín hiệu đối tượng âm thanh được ước tính.

Ngoài ra, phương pháp mã hóa đối tượng âm thanh theo tham số sử dụng khái niệm phân dư được đề xuất. Phương pháp bao gồm;

- Tạo ra nhiều tín hiệu đối tượng âm thanh được ước tính thứ nhất bằng cách trộn tăng ba hoặc nhiều hơn ba tín hiệu trộn giảm, trong đó ba hoặc nhiều hơn ba tín hiệu trộn giảm mã hóa nhiều tín hiệu đối tượng âm thanh gốc, trong đó việc tạo ra nhiều tín hiệu đối tượng âm thanh được ước tính thứ nhất bao gồm trộn tăng ba hoặc nhiều hơn ba tín hiệu trộn giảm phụ thuộc vào thông tin phụ theo tham số biểu thị thông tin trên nhiều tín hiệu đối tượng âm thanh gốc. Và:

- Tạo ra nhiều tín hiệu đối tượng âm thanh được ước tính thứ hai bằng cách biến đổi một hoặc nhiều tín hiệu đối tượng âm thanh được ước tính thứ nhất, trong đó tạo ra nhiều tín hiệu đối tượng âm thanh được ước tính thứ hai bao gồm biến đổi một hoặc nhiều tín hiệu đối tượng âm thanh được ước tính thứ nhất đã nêu phụ thuộc vào một hoặc nhiều tín hiệu dư.

Ngoài ra, phương pháp mã hóa đối tượng âm thanh theo tham số sử dụng khái niệm phân dư khác được đề xuất. Phương pháp đã nêu bao gồm:

- Tạo ra nhiều tín hiệu đối tượng âm thanh được ước tính bằng cách trộn tăng ba hoặc nhiều hơn ba tín hiệu trộn giảm, trong đó ba hoặc nhiều hơn ba tín hiệu trộn giảm mã hóa nhiều tín hiệu đối tượng âm thanh gốc, trong đó việc tạo ra nhiều tín hiệu đối tượng âm thanh được ước tính bao gồm trộn tăng ba hoặc nhiều hơn ba tín hiệu trộn giảm phụ thuộc vào thông tin phụ theo tham số biểu thị thông tin trên nhiều tín hiệu đối tượng âm thanh gốc. Và:

- Tạo ra nhiều tín hiệu dư dựa trên nhiều tín hiệu đối tượng âm thanh gốc và dựa trên nhiều tín hiệu đối tượng âm thanh được ước tính, như vậy mà mỗi tín hiệu trong số nhiều tín hiệu dư là tín hiệu khác nhau biểu thị sự khác biệt giữa một tín hiệu của nhiều tín hiệu đối tượng âm thanh gốc và một tín hiệu của nhiều tín hiệu đối tượng âm thanh được ước tính.

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

Mô tả vắn tắt các hình vẽ

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

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

Fig.1b minh họa bộ giải mã theo phương án khác của sáng chế, trong đó bộ giải mã còn bao gồm bộ biểu diễn,

Fig.2a minh họa bộ tạo tín hiệu dư theo phương án của sáng chế,

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

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

Fig.4 minh họa tín hiệu âm thanh được mã hóa theo phương án của sáng chế,

Fig.5 thể hiện tổng quan hệ thống SAOC minh họa nguyên lý hệ thống tham số đã nêu sử dụng ví dụ MPEG SAOC,

Fig.6 thể hiện sự ước tính phần dư tại phía bộ mã hóa, biểu đồ minh họa sự tính toán của các tín hiệu dư cho mỗi EAO.

Fig.7 thể hiện cấu trúc cơ bản của bộ giải mã SAOC với hỗ trợ EAO, minh họa tổng quan khái niệm của biểu đồ xử lý EAO được tích hợp vào chuỗi giải mã/chuyển mã SAOC,

Fig.8 thể hiện tổng quan khái niệm của tham số và phần dư được thể hiện dựa trên biểu đồ mã hóa đối tượng âm thanh theo phương án của sáng chế,

Fig.9 thể hiện khái niệm để ước tính đồng thời tín hiệu dư cho mỗi tín hiệu EAO tại phía mã hóa theo phương án của sáng chế,

Fig.10 minh họa khái niệm của việc giải mã phần dư đồng thời tại phía giải mã theo phương án của sáng chế,

Fig.11 minh họa bộ tạo tín hiệu dư theo phương án của sáng chế, trong đó bộ tạo tín hiệu dư còn bao gồm bộ phận biến đổi trộn giảm,

Fig.12 minh họa bộ giải mã theo phương án khác của sáng chế, trong đó bộ giải mã còn bao gồm bộ phận biến đổi trộn giảm,

Fig.13 minh họa khái niệm của việc tính toán các thành phần dư trong đường nổi tầng tại phía bộ mã hóa theo phương án của sáng chế,

Fig.14 minh họa bộ phận "Giải mã RSI" nổi tầng sử dụng trong tổ hợp với việc tính toán dư nổi tầng tại bên giải mã theo phương án của sáng chế,

Fig.15 minh họa bộ tạo tín hiệu dư theo phương án sử dụng khái niệm nổi tầng, và

Fig.16 minh họa bộ giải mã theo phương án của sáng chế sử dụng khái niệm nổi tầng.

Mô tả chi tiết sáng chế

Bộ tạo tín hiệu dư 200 bao gồm bộ phận giải mã theo tham số 230 để tạo ra nhiều tín hiệu đối tượng âm thanh được ước tính (Tín Hiệu Đối Tượng Âm Thanh Được Ước tính #1, ... Tín Hiệu Đối Tượng Âm Thanh Được Ước tính #M) bằng cách trộn tăng ba hoặc nhiều hơn ba tín hiệu trộn giảm (Tín Hiệu Trộn Giảm #1, Tín Hiệu Trộn Giảm #2, Tín Hiệu Trộn Giảm #3, ..., Tín Hiệu Trộn Giảm #N). Ba hoặc nhiều hơn ba tín hiệu trộn giảm (Tín Hiệu Trộn Giảm #1, Tín Hiệu Trộn Giảm #2, Tín Hiệu Trộn Giảm #3, ..., Tín Hiệu Trộn Giảm #N) mã hóa nhiều tín hiệu đối tượng âm thanh gốc (Tín Hiệu Đối Tượng Âm Thanh Gốc #1, ..., Tín Hiệu Đối Tượng Âm Thanh Gốc #M). Bộ phận giải mã theo tham số 230 được tạo cấu hình để trộn tăng ba hoặc nhiều hơn ba tín hiệu trộn giảm (Tín Hiệu Trộn Giảm #1, Tín Hiệu Trộn Giảm #2, Tín Hiệu Trộn Giảm #3, ..., Tín Hiệu Trộn Giảm #N) phụ thuộc vào thông tin phụ theo tham số biểu thị thông tin trên nhiều tín hiệu đối tượng âm thanh gốc (Tín Hiệu Đối Tượng Âm Thanh Gốc #1, ..., Tín Hiệu Đối Tượng Âm Thanh Gốc #M).

Ngoài ra, bộ tạo tín hiệu dư 200 bao gồm bộ phận ước tính phần dư 240 để tạo ra nhiều tín hiệu dư (Tín Hiệu Dư #1, ..., Tín Hiệu Dư #M) dựa trên nhiều tín hiệu đối tượng âm thanh gốc (Tín Hiệu Đối Tượng Âm Thanh Gốc #1, ..., Tín Hiệu Đối Tượng Âm Thanh Gốc #M) và dựa trên nhiều tín hiệu đối tượng âm thanh được ước tính (Tín Hiệu Đối Tượng Âm Thanh Được Ước tính #1, ... Tín Hiệu Đối Tượng Âm Thanh Được Ước tính #M), như vậy mà mỗi tín hiệu của nhiều tín hiệu dư (Tín Hiệu Dư #1, ..., Tín Hiệu Dư #M) là tín hiệu khác nhau biểu thị sự khác nhau giữa một tín hiệu của nhiều tín hiệu đối tượng âm thanh gốc (Tín Hiệu Đối Tượng Âm Thanh Gốc #1, Tín Hiệu Đối Tượng Âm Thanh Gốc #M) và một tín hiệu trong số nhiều tín hiệu đối tượng

âm thanh được ước tính (Tín Hiệu Đối Tượng Âm Thanh Được Ước tính #1, ... Tín Hiệu Đối Tượng Âm Thanh Được Ước tính #M).

Bộ mã hóa theo phương án được mô tả ở trên khắc phục những hạn chế SAOC (xem tài liệu [SAOC]) của tình trạng kỹ thuật.

Hệ thống SAOC hiện tại điều khiển việc trộn giảm bởi việc sử dụng một hoặc nhiều các hộp-từ-hai-đến-một hoặc một hoặc nhiều các hộp-từ-ba-đến-hai. Ngoài ra, vì những hạn chế cơ bản này, các hệ thống SAOC hiện tại có thể trộn giảm các tín hiệu đối tượng âm thanh đến nhiều nhất hai kênh trộn giảm/hai tín hiệu trộn giảm.

Các khái niệm dùng cho các bộ tạo tín hiệu dư và dùng cho các bộ mã hóa được đề xuất, mà cho phép khắc phục những hạn chế của SAOC để việc mã hóa đối tượng âm thanh đến nay là thuận lợi cho các hệ thống truyền dẫn mà sử dụng nhiều hơn hai kênh truyền dẫn.

Trong phương án, bộ phận ước tính phần dư 240 được làm thích ứng để tạo ra ít nhất năm tín hiệu dư dựa trên ít nhất năm tín hiệu đối tượng âm thanh gốc trong số nhiều tín hiệu đối tượng âm thanh gốc và dựa trên ít nhất năm tín hiệu đối tượng âm thanh được ước tính trong số nhiều tín hiệu đối tượng âm thanh được ước tính.

Fig.2b minh họa bộ mã hóa theo phương án của sáng chế. Bộ mã hóa theo Fig.2b bao gồm bộ tạo tín hiệu dư 200.

Ngoài ra, bộ mã hóa bao gồm bộ tạo trộn giảm 210 để cung cấp ba hoặc nhiều hơn ba tín hiệu trộn giảm (Tín Hiệu Trộn Giảm #1, Tín Hiệu Trộn Giảm #2, Tín Hiệu Trộn Giảm #3, ..., Tín Hiệu Trộn Giảm #N) biểu thị sự trộn giảm của nhiều tín hiệu đối tượng âm thanh gốc (Tín Hiệu Đối Tượng Âm Thanh Gốc #1, ..., Tín Hiệu Đối Tượng Âm Thanh Gốc #M, (các) Tín Hiệu Đối Tượng Âm Thanh Gốc hơn nữa).

Về Tín Hiệu Đối Tượng Âm Thanh Gốc #1, ..., Tín Hiệu Đối Tượng Âm Thanh Gốc #M, bộ phận ước tính phần dư 240 tạo ra tín hiệu dư (Tín Hiệu Dư #1, ..., Tín Hiệu Dư #M). Do đó, Tín Hiệu Đối Tượng Âm Thanh Gốc #1, ..., Tín Hiệu Đối Tượng Âm Thanh Gốc #M đề cập đến các đối tượng âm thanh được nâng cao (EAO).

Tuy nhiên, như có thể thấy trên Fig.2b, (các) tín hiệu đối tượng âm thanh gốc hơn nữa có thể tồn tại tùy ý, mà được trộn giảm, nhưng sẽ không tạo ra tín hiệu dư. Do đó (các) tín hiệu đối tượng âm thanh gốc hơn nữa này đề cập đến các đối tượng âm thanh không được nâng cao (Non-Enhanced Audio Object – Non-EAO).

Bộ mã hóa theo Fig.2b còn bao gồm bộ ước tính thông tin phụ theo tham số 220 để tạo ra thông tin phụ theo tham số biểu thị thông tin trên nhiều tín hiệu đối tượng âm thanh gốc (Tín Hiệu Đối Tượng Âm Thanh Gốc #1, ..., Tín Hiệu Đối Tượng Âm Thanh Gốc #M, (các) Tín Hiệu Đối Tượng Âm Thanh Gốc hơn nữa), để thu được thông tin phụ theo tham số. Trong phương án theo Fig.2b, bộ ước tính thông tin phụ theo tham số cũng đưa các tín hiệu đối tượng âm thanh gốc ((các) tín hiệu đối tượng âm thanh gốc hơn nữa) đề cập đến các non-EAO vào trong tài khoản.

Trong phương án, số tín hiệu đối tượng âm thanh gốc có thể bằng số tín hiệu dư, ví dụ, khi tất cả các tín hiệu đối tượng âm thanh gốc đề cập đến các EAO.

Tuy nhiên, trong các phương án khác, số tín hiệu dư có thể khác số tín hiệu đối tượng âm thanh gốc và/hoặc có thể khác số tín hiệu đối tượng âm thanh được ước tính, ví dụ, khi các tín hiệu đối tượng âm thanh gốc đề cập đến các Non-EAO.

Theo một số phương án, bộ mã hóa là bộ mã hóa SAOC.

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

Bộ giải mã bao gồm bộ phận giải mã theo tham số 110 để tạo ra nhiều tín hiệu đối tượng âm thanh được ước tính thứ nhất (Tín Hiệu Đối Tượng Âm Thanh Được Ước tính Thứ Nhất #1, ..., Tín Hiệu Đối Tượng Âm Thanh Được Ước tính Thứ Nhất #M) bằng cách trộn tăng ba hoặc nhiều hơn ba tín hiệu trộn giảm (Tín Hiệu Trộn Giảm #1, Tín Hiệu Trộn Giảm #2, Tín Hiệu Trộn Giảm #3, ..., Tín Hiệu Trộn Giảm #N), trong đó ba hoặc nhiều hơn ba tín hiệu trộn giảm (Tín Hiệu Trộn Giảm #1, Tín Hiệu Trộn Giảm #2, Tín Hiệu Trộn Giảm #3, ..., Tín Hiệu Trộn Giảm #N) mã hóa nhiều tín hiệu đối tượng âm thanh gốc, trong đó bộ giải mã tham số 110 được tạo cấu hình để trộn tăng ba hoặc nhiều hơn ba tín hiệu trộn giảm (Tín Hiệu Trộn Giảm #1, Tín Hiệu Trộn Giảm #2, Tín Hiệu Trộn Giảm #3, ..., Tín Hiệu Trộn Giảm #N) phụ thuộc vào

thông tin phụ theo tham số biểu thị thông tin trên nhiều tín hiệu đối tượng âm thanh gốc.

Ngoài ra, bộ giải mã bao gồm bộ phận xử lý phần dư 120 để tạo ra nhiều tín hiệu đối tượng âm thanh được ước tính thứ hai (Tín Hiệu Đối Tượng Âm Thanh Được Ước tính Thứ Hai #1, ...Tín Hiệu Đối Tượng Âm Thanh Được Ước tính Thứ Hai #M) bằng cách biến đổi một hoặc nhiều tín hiệu đối tượng âm thanh được ước tính thứ nhất (Tín Hiệu Đối Tượng Âm Thanh Được Ước tính Thứ Nhất #1, ...Tín Hiệu Đối Tượng Âm Thanh Được Ước tính Thứ Nhất #M), trong đó bộ phận xử lý phần dư 120 được tạo cấu hình để biến đổi một hoặc nhiều tín hiệu đối tượng âm thanh được ước tính thứ nhất đã nêu (Tín Hiệu Đối Tượng Âm Thanh Được Ước tính Thứ Nhất #1, ...Tín Hiệu Đối Tượng Âm Thanh Được Ước tính Thứ Nhất #M) phụ thuộc vào một hoặc nhiều tín hiệu dư (Tín Hiệu Dư #1, ...,Tín Hiệu Dư #M).

Bộ giải mã theo phương án được mô tả ở trên khắc phục những hạn chế SAOC (xem tài liệu [SAOC]) của tình trạng kỹ thuật.

Hơn nữa, các hệ thống SAOC hiện tại điều khiển việc trộn tăng bằng cách sử dụng một hoặc nhiều hộp-từ-một-đến-hai (các hộp OTT) hoặc một hoặc nhiều hộp-từ-hai-đến-ba (các hộp TTT). Ngoài ra, vì những hạn chế này, các tín hiệu đối tượng âm thanh được mã hóa với hơn hai tín hiệu trộn giảm/kênh trộn giảm có thể không được trộn tăng bởi các bộ giải mã SAOC thuộc tình trạng kỹ thuật.

Các khái niệm dùng cho các bộ giải mã được đề xuất, mà cho phép khắc phục những hạn chế của SAOC để việc mã hóa đối tượng âm thanh đến nay là thuận lợi cho các hệ thống truyền dẫn mà sử dụng nhiều hơn hai kênh truyền dẫn.

Fig.1b minh họa bộ giải mã theo phương án khác, trong đó bộ giải mã còn bao gồm bộ phận biểu diễn 130 để tạo ra nhiều kênh đầu ra âm thanh (Kênh Đầu Ra Âm Thanh #1, ..., Kênh Đầu Ra Âm Thanh #R) từ các tín hiệu đối tượng âm thanh được ước tính thứ hai (Tín Hiệu Đối Tượng Âm Thanh Được Ước tính thứ hai #1, ... Tín Hiệu Đối Tượng Âm Thanh Được Ước tính thứ hai #M) phụ thuộc vào thông tin biểu diễn. Ví dụ, thông tin biểu diễn có thể là ma trận biểu diễn và/hoặc các hệ số của ma trận biểu diễn và bộ phận biểu diễn 130 có thể được tạo cấu hình để áp dụng ma trận

biểu diễn trên các tín hiệu đối tượng âm thanh được ước tính thứ hai (Tín Hiệu Đối Tượng Âm Thanh Được Ước tính thứ hai #1, ... Tín Hiệu Đối Tượng Âm Thanh Được Ước tính thứ hai #M) để thu được nhiều kênh đầu ra âm thanh (Kênh Đầu Ra Âm Thanh #1, ..., Kênh Đầu Ra Âm Thanh #R).

Theo phương án, bộ phận xử lý phần dư 120 được tạo cấu hình để biến đổi một hoặc nhiều tín hiệu đối tượng âm thanh được ước tính thứ nhất đã nêu phụ thuộc vào ít nhất ba tín hiệu dư. Bộ giải mã được làm thích ứng để tạo ra ít nhất ba kênh đầu ra âm thanh dựa trên nhiều tín hiệu đối tượng âm thanh được ước tính thứ hai.

Trong một phương án khác, mỗi một hoặc nhiều tín hiệu dư có thể biểu thị sự khác biệt giữa một tín hiệu của nhiều tín hiệu đối tượng âm thanh gốc và một của một hoặc nhiều tín hiệu đối tượng âm thanh được ước tính thứ nhất.

Theo phương án, bộ phận xử lý phần dư 120 được làm thích ứng để tạo ra nhiều tín hiệu đối tượng âm thanh được ước tính thứ hai bằng cách biến đổi năm hoặc nhiều hơn năm tín hiệu đối tượng âm thanh được ước tính thứ nhất. Bộ phận xử lý phần dư 120 được làm thích ứng để biến đổi năm hoặc nhiều hơn năm tín hiệu đối tượng âm thanh được ước tính thứ nhất đã nêu phụ thuộc vào năm hoặc nhiều hơn năm tín hiệu dư.

Trong phương án khác, bộ giải mã được tạo cấu hình để tạo ra bảy hoặc nhiều hơn bảy kênh đầu ra âm thanh dựa trên nhiều tín hiệu đối tượng âm thanh được ước tính thứ hai.

Theo phương án nữa, bộ giải mã được làm thích ứng để không xác định các hệ số dự báo kênh để xác định nhiều tín hiệu đối tượng âm thanh được ước tính thứ hai.

Trong phương án nữa, bộ giải mã là bộ giải mã SAOC.

Fig.3 minh họa hệ thống theo phương án của sáng chế. Hệ thống bao gồm bộ mã hóa 310 theo một trong các phương án được mô tả ở trên để mã hóa nhiều tín hiệu đối tượng âm thanh gốc (Tín Hiệu Đối Tượng Âm Thanh Gốc #1, ..., Tín Hiệu Đối Tượng Âm Thanh Gốc #M) bằng cách tạo ra ba hoặc nhiều hơn ba tín hiệu trộn giảm, bằng cách tạo ra thông tin phụ theo tham số và bằng cách tạo ra nhiều tín hiệu dư. Ngoài ra,

hệ thống bao gồm bộ giải mã 320 theo một trong số các phương án được mô tả ở trên, trong đó bộ giải mã 320 được tạo cấu hình để tạo ra nhiều tín hiệu đối tượng âm thanh được ước tính thứ hai dựa trên ba hoặc nhiều hơn ba tín hiệu trộn giảm được tạo ra bởi bộ mã hóa 310, dựa trên thông tin phụ theo tham số được tạo ra bởi bộ mã hóa 310 và dựa trên nhiều tín hiệu dư được tạo ra bởi bộ mã hóa 310.

Fig.4 minh họa tín hiệu âm thanh được mã hóa theo phương án của sáng chế. Tín hiệu âm thanh được mã hóa bao gồm ba hoặc nhiều hơn ba tín hiệu trộn giảm 410, thông tin phụ theo tham số 420 và nhiều tín hiệu dư 430. Ba hoặc nhiều hơn ba tín hiệu trộn giảm 410 là sự trộn giảm của nhiều tín hiệu đối tượng âm thanh gốc. Thông tin phụ theo tham số 420 bao gồm các tham số biểu thị thông tin phụ trên nhiều tín hiệu đối tượng âm thanh gốc. Mỗi tín hiệu của nhiều tín hiệu dư 430 là các tín hiệu khác nhau biểu thị sự khác biệt giữa một tín hiệu của nhiều tín hiệu âm thanh gốc và một tín hiệu của nhiều tín hiệu đối tượng âm thanh được ước tính.

Sau đây, tổng quan khái niệm theo phương án của sáng chế được đề xuất.

Fig.8 mô tả tổng quan khái niệm của tham số được thể hiện và đối tượng âm thanh trên cơ sở dư mã hóa biểu đồ theo phương án, trong đó biểu đồ mã hóa thể hiện tín hiệu trộn giảm cấp cao và hỗ trợ EAO cấp cao.

Ở phía bộ mã hóa, bộ ước tính thông tin phụ theo tham số ("bộ phận tạo PSI") 220 tính toán PSI để ước tính tín hiệu đối tượng tại bộ giải mã khai thác nguồn và các đặc điểm liên quan trộn giảm. Bộ phận tạo RSI 245 tính toán mỗi tín hiệu đối tượng được nâng cao thông tin dư bằng cách phân tích sự khác nhau giữa các tín hiệu đối tượng được ước tính và tín hiệu đối tượng gốc. Bộ phận tạo RSI có thể, ví dụ, bao gồm bộ phận giải mã theo tham số 230 và bộ phận ước tính phần dư 240.

Tại phía giải mã, bộ phận giải mã theo tham số (bộ phận "giải mã PSI") 110 ước tính các tín hiệu đối tượng từ các tín hiệu trộn giảm với PSI đã xác định. Trong bước thứ hai, bộ phận xử lý phần dư (bộ phận "giải mã RSI") 120 sử dụng RSI để cải thiện chất lượng của các tín hiệu đối tượng được ước tính đã được nâng cao. Tất cả các tín hiệu đối tượng (các đối tượng âm thanh không nâng cao và nâng cao) có thể, ví dụ, được chuyển đến bộ phận biểu diễn 130 để tạo ra cảnh đầu ra mục tiêu.

Cần lưu ý rằng không cần thiết phải xem xét tất cả các tín hiệu trộn giảm. Các tín hiệu trộn giảm có thể bị bỏ qua khỏi việc tính toán nếu sự đóng góp của nó trong việc ước tính hoặc/và ước tính và nâng cao các tín hiệu đối tượng có thể bị bỏ qua.

Để dễ hiểu, các bước xử lý trên Fig.8 và trong các hình dưới đây được hình dung như các bộ phận xử lý riêng biệt. Trong thực tế, chúng có thể được kết hợp một cách hiệu quả để giảm bớt sự phức tạp trong tính toán.

Sau đây, khái niệm mã hóa/giải mã dư chung được đề xuất.

Fig.9 mô tả khái niệm để ước tính đồng thời tín hiệu dư cho mỗi tín hiệu EAO tại phía mã hóa theo phương án của sáng chế.

Bộ phận giải mã theo tham số (bộ phận "giải mã PSI") 230 cho ra sự ước tính của các tín hiệu đối tượng âm thanh (các tín hiệu đối tượng âm thanh được ước tính $S_{est,PSI,\{1,...,M\}}$ xác định PSI được ước tính và (các) tín hiệu trộn giảm như đầu vào. Các tín hiệu đối tượng âm thanh được ước tính $S_{est,PSI,\{1,...,M\}}$ được so sánh với các tín hiệu nguồn không bị biến đổi gốc $s_1,...,s_M$ trong các bộ phận ước tính phần dư (bộ phận "Ước tính RSI") 240. Bộ phận ước tính phần dư 240 cung cấp số hạng tín hiệu dư/lỗi $S_{res,RSI,\{1,...,M\}}$ cho mỗi đối tượng âm thanh được nâng cao.

Fig.10 mô tả bộ phận "giải mã RSI" được sử dụng trong sự kết hợp với sự ước tính dư trong bộ giải mã. Cụ thể, Fig.10 minh họa khái niệm của việc giải mã dư chung tại phía giải mã theo phương án của sáng chế.

Các tín hiệu đối tượng âm thanh được ước tính (thứ nhất) $S_{est,PSI,\{1,...,M\}}$ từ bộ phận giải mã theo tham số (bộ phận "giải mã PSI") 110 được nạp cùng với thông tin dư ("thông tin phụ dư") vào trong bộ phận xử lý phần dư ("giải mã RSI") 120. Bộ phận xử lý phần dư 120 tính toán từ thông tin (phụ) dư và các tín hiệu đối tượng âm thanh được ước tính $S_{est,PSI,\{1,...,M\}}$ các tín hiệu đối tượng âm thanh được ước tính thứ hai $S_{est,RSI,\{1,...,M\}}$, ví dụ, các tín hiệu đối tượng âm thanh được nâng cao và không được nâng cao, cho ra các tín hiệu đối tượng âm thanh được ước tính thứ hai $S_{est,RSI,\{1,...,M\}}$, ví dụ, các tín hiệu đối tượng âm thanh được nâng cao và không được nâng cao, như đầu ra của bộ phận xử lý phần dư 120.

Thêm vào đó, việc ước tính lại các non-EAO có thể được thực hiện (không được minh họa trên Fig.10). Các EAO được loại bỏ từ hỗn hợp tín hiệu và việc giữ lại các non-EAO được ước tính lại từ hỗn hợp này. Điều này cho ra sự ước tính được cải thiện của các đối tượng được so sánh với sự ước tính từ hỗn hợp tín hiệu mà chứa tất cả các tín hiệu đối tượng. Việc ước tính lại này có thể được bỏ qua, nếu mục tiêu là chỉ điều khiển các tín hiệu đối tượng được nâng cao trong hỗn hợp.

Fig.11 minh họa bộ tạo tín hiệu dư theo phương án của sáng chế, trong đó.

Trên Fig.11, bộ tạo tín hiệu dư 200 còn bao gồm bộ phận biến đổi trộn giảm 250 được làm thích ứng để biến đổi ba hoặc nhiều hơn ba tín hiệu trộn giảm để thu được ba hoặc nhiều hơn ba tín hiệu trộn giảm được biến đổi.

Bộ phận giải mã theo tham số 230 được tạo cấu hình để xác định một hoặc nhiều tín hiệu đối tượng âm thanh trong số các tín hiệu đối tượng âm thanh được ước tính thứ nhất dựa trên ba hoặc nhiều hơn ba tín hiệu trộn giảm được biến đổi.

Sau đó, bộ phận ước tính phần dư 240 có thể, ví dụ, xác định một hoặc nhiều tín hiệu dư dựa trên một hoặc nhiều tín hiệu đối tượng âm thanh đã nêu của các tín hiệu đối tượng âm thanh được ước tính thứ nhất.

Trong phương án, bộ phận biến đổi trộn giảm 250 có thể, ví dụ, được tạo cấu hình để biến đổi ba hoặc nhiều hơn ba tín hiệu trộn giảm gốc để thu được ba hoặc nhiều hơn ba tín hiệu trộn giảm được biến đổi, bằng cách loại bỏ một hoặc nhiều tín hiệu trong số nhiều tín hiệu đối tượng âm thanh gốc từ ba hoặc nhiều hơn ba tín hiệu trộn giảm gốc.

Trong phương án khác, bộ phận biến đổi trộn giảm 250 có thể, ví dụ, được tạo cấu hình để biến đổi ba hoặc nhiều hơn ba tín hiệu trộn giảm gốc để thu được ba hoặc nhiều hơn ba tín hiệu trộn giảm được biến đổi bằng cách tạo ra một hoặc nhiều tín hiệu đối tượng âm thanh được biến đổi dựa trên một hoặc nhiều tín hiệu đối tượng âm thanh được ước tính và dựa trên một hoặc nhiều tín hiệu dư, và bằng cách loại bỏ một hoặc nhiều tín hiệu đối tượng âm thanh được biến đổi từ ba hoặc nhiều hơn ba tín hiệu trộn giảm gốc. Ví dụ, mỗi một hoặc nhiều tín hiệu đối tượng âm thanh được biến đổi có thể

được tạo ra từ bộ phận biến đổi trộn giảm bằng cách biến đổi một tín hiệu đối tượng âm thanh được ước tính, trong đó bộ phận biến đổi trộn giảm có thể được làm thích ứng để biến đổi tín hiệu đối tượng âm thanh được ước tính đã nêu phụ thuộc vào một tín hiệu của một hoặc nhiều tín hiệu dư.

Trong cả hai phương án được mô tả ở trên, bộ phận biến đổi trộn giảm có thể, ví dụ, được làm thích ứng để áp dụng công thức

$$\tilde{\mathbf{X}} = \mathbf{X} - \mathbf{D}\mathbf{Z}_{eao}^* \mathbf{S}_{eao},$$

trong đó $\mathbf{X}$ là sự trộn giảm được biến đổi,

trong đó $\mathbf{D}$ biểu thị thông tin trộn giảm tương ứng,

trong đó $\mathbf{S}_{eao}$ chứa các tín hiệu đối tượng âm thanh gốc được loại bỏ hoặc các tín hiệu đối tượng âm thanh được biến đổi được loại bỏ,

trong đó $\mathbf{Z}_{eao}^*$ biểu thị các vị trí của các tín hiệu được loại bỏ, và

trong đó $\tilde{\mathbf{X}}$ là tín hiệu trộn giảm được biến đổi.

Ví dụ, điểm (vị trí) của tín hiệu đối tượng âm thanh tương ứng với điểm (vị trí) của đối tượng âm thanh của nó trong danh sách tất cả các đối tượng.

Fig.12 minh họa bộ giải mã theo phương án của sáng chế.

Trong phương án theo Fig.12, bộ giải mã còn bao gồm bộ phận biến đổi trộn giảm 140.

Bộ phận xử lý phần dư 120 xác định một hoặc nhiều tín hiệu đối tượng âm thanh trong số nhiều tín hiệu đối tượng âm thanh được ước tính thứ hai.

Bộ phận biến đổi trộn giảm 140 được làm thích ứng để loại bỏ một hoặc nhiều tín hiệu đối tượng âm thanh được ước tính thứ hai được xác định từ ba hoặc nhiều hơn ba tín hiệu trộn giảm để thu được ba hoặc nhiều hơn ba tín hiệu trộn giảm được biến đổi.

Bộ phận giải mã theo tham số 110 được tạo cấu hình để xác định một hoặc nhiều tín hiệu đối tượng âm thanh của các tín hiệu đối tượng âm thanh được ước tính thứ nhất dựa trên ba hoặc nhiều hơn ba tín hiệu trộn giảm được biến đổi.

Sau đó bộ phận xử lý phần dư 120 có thể, ví dụ, xác định một hoặc nhiều tín hiệu đối tượng âm thanh được ước tính dựa trên một hoặc nhiều tín hiệu đối tượng âm thanh được xác định của các tín hiệu đối tượng âm thanh được ước tính thứ nhất.

Trong phương án cụ thể, bộ phận biến đổi trộn giảm 130 có thể, ví dụ, được làm thích ứng để áp dụng công thức:

$$\tilde{\mathbf{X}}_{nonEAO} = \mathbf{X} - \mathbf{D}\mathbf{Z}_{eao}^* \mathbf{S}_{eao}.$$

để loại bỏ một hoặc nhiều tín hiệu đối tượng âm thanh trong số nhiều tín hiệu đối tượng âm thanh được ước tính thứ hai bằng bộ phận xử lý phần dư 120 từ ba hoặc nhiều hơn ba tín hiệu trộn giảm để thu được ba hoặc nhiều hơn ba tín hiệu trộn giảm được biến đổi, trong đó

$\mathbf{X}$ biểu thị ba hoặc nhiều hơn ba tín hiệu trộn giảm trước khi được biến đổi

$\tilde{\mathbf{X}}_{nonEAO}$ biểu thị ba hoặc nhiều hơn ba tín hiệu trộn giảm được biến đổi

$\mathbf{D}$ biểu thị ma trận trộn giảm

$\mathbf{Z}_{eao}$ biểu thị bản đồ ma trận con thể hiện các vị trí (điểm) của các EAO

(Để chi tiết hơn về các biến thể cụ thể của phương án này, xem phần mô tả dưới đây).

Sau đây, khái niệm mã hóa/giải mã dư nổi tăng được đề xuất.

Fig.13 minh họa khái niệm của việc tính toán các thành phần dư trong cách nổi tăng tại phía bộ mã hóa theo phương án của sáng chế. So với khái niệm tính toán dư chung, sự tiếp cận nổi tăng giảm trong mỗi bước lặp lại năng lượng của năng lượng dư tại giá trị của độ phức tạp tính toán cao hơn. Trong mỗi bước, một tín hiệu của các tín hiệu đối tượng âm thanh gốc (s_m) (hoặc, trong phương án thay thế, tín hiệu đối tượng

âm thanh được ước tính; xem các mũi tên đứt đoạn 2461, 2462) của đối tượng âm thanh được nâng cao được loại bỏ khỏi hỗn hợp tín hiệu (trộn giảm) trước khi hỗn hợp tín hiệu (trộn giảm) được chuyển đến bộ phận xử lý tiếp theo 2452. Theo cách này, số tín hiệu đối tượng trong hỗn hợp tín hiệu (trộn giảm) giảm với mỗi bước xử lý. Sự ước tính của tín hiệu đối tượng âm thanh được nâng cao (tín hiệu đối tượng âm thanh được ước tính thứ hai) trong bước tiếp theo do đó cải thiện, do việc giảm liên tục năng lượng của các tín hiệu dư.

(Cần lưu ý rằng, trong phương án thay thế đó, trong mỗi bước lặp lại, tín hiệu đối tượng âm thanh được ước tính được loại bỏ khỏi hỗn hợp tín hiệu, các bộ phận con biến đổi trộn giảm 2501, 2502 không cần phải nhận các tín hiệu đối tượng âm thanh gốc S_M).

Ngược lại, trong phương án, trong mỗi bước lặp lại, tín hiệu đối tượng âm thanh gốc được loại bỏ khỏi hỗn hợp tín hiệu, các bộ phận con biến đổi trộn giảm 2501, 2502 không cần phải nhận các tín hiệu đối tượng âm thanh được ước tính.)

Cụ thể hơn, Fig.13 minh họa nhiều bộ phận con tạo RSI 2451, 2452. Nhiều bộ phận con tạo RSI 2451, 2452 cùng nhau tạo thành bộ phận tạo RSI.

Mỗi bộ phận của nhiều bộ phận con tạo RSI 2451, 2452 bao gồm bộ phận con giải mã tham số 2301. Nhiều bộ phận con giải mã tham số 2301 cùng nhau tạo thành bộ phận giải mã theo tham số. Các bộ phận con giải mã tham số 2301 tạo ra các tín hiệu đối tượng âm thanh được ước tính thứ nhất $s_{est,PSI,\{1,...,M\}}$.

Mỗi một bộ phận trong số nhiều bộ phận con tạo RSI 2451, 2452 bao gồm bộ phận con ước tính dư 2401. Nhiều bộ phận con ước tính dư 2401 cùng nhau tạo thành bộ phận ước tính phần dư. Các bộ phận con ước tính dư 2401 tạo ra các tín hiệu đối tượng âm thanh được ước tính thứ hai $s_{est,RSI,M}$, $s_{est,RSI,M-1}$.

Ngoài ra, Fig.13 minh họa nhiều bộ phận con biến đổi trộn giảm RSI 2501, 2502. Mỗi bộ phận trong số các bộ phận con biến đổi trộn giảm 2501, 2502 cùng nhau tạo thành bộ phận biến đổi trộn giảm.

Fig.14 bộc lộ bộ phận " giải mã RSI" nổi tăng sử dụng trong tổ hợp với tính toán dư nổi tăng tại phía giải mã theo phương án của sáng chế.

Trong mỗi bước, một trong những tín hiệu đối tượng được nâng cao là được ước tính bởi bộ phận con giải mã tham số ("giải mã PSI) 1101 (để thu được một trong những tín hiệu đối tượng âm thanh được ước tính thứ nhất $S_{est,PSI,M}$), và một trong những tín hiệu đối tượng âm thanh được ước tính thứ nhất $S_{est,PSI,M}$ sau đó được xử lý cùng với tín hiệu dư tương ứng $S_{res,RSI,M}$ bởi bộ phận con xử lý dư ("xử lý RSI") 1201, để cho ra bản được nâng cao của tín hiệu đối tượng (một trong những tín hiệu đối tượng âm thanh được ước tính thứ hai) $S_{est,RSI,M}$. Tín hiệu đối tượng được nâng cao $S_{est,RSI,M}$ được loại bỏ khỏi tín hiệu trộn giảm bởi bộ phận con biến đổi trộn giảm ("biến đổi trộn giảm") 1401 trước khi các tín hiệu trộn giảm được biến đổi được nạp vào trong bộ phận con giải mã dư tiếp theo ("giải mã dư") 1252 .

Giống như khái niệm mã hóa/giải mã dư chung, các non-EAO có thể được ước tính lại bổ sung.

Cụ thể hơn, Fig.14 minh họa nhiều bộ phận con giải mã dư 1251, 1252. Nhiều bộ phận con giải mã dư 1251, 1252 cùng nhau tạo thành bộ phận giải mã dư.

Mỗi bộ phận trong số nhiều bộ phận con giải mã dư 1251, 1252 bao gồm bộ phận con giải mã tham số 1101. Nhiều bộ phận con giải mã tham số 1101 cùng nhau tạo thành bộ phận giải mã theo tham số. Các bộ phận con giải mã tham số 1101 tạo ra các tín hiệu đối tượng âm thanh được ước tính thứ nhất $S_{est,PSI,\{1,...,M\}}$.

Mỗi bộ phận trong số nhiều bộ phận con giải mã dư 1251, 1252 bao gồm bộ phận con xử lý dư 1201. Nhiều bộ phận con xử lý dư 1201 cùng nhau tạo thành bộ phận xử lý phần dư. Các bộ phận con xử lý dư 1201 tạo ra các tín hiệu đối tượng âm thanh được ước tính thứ hai $S_{est,RSI,M}$, $S_{est,RSI,M-1}$.

Ngoài ra, Fig.14 minh họa nhiều bộ phận con biến đổi trộn giảm 1401, 1402. Mỗi bộ phận trong số các bộ phận con biến đổi trộn giảm 1401, 1402 cùng nhau tạo thành bộ phận biến đổi trộn giảm.

Fig.15 minh họa bộ tạo tín hiệu dư theo phương án sử dụng khái niệm nổi tăng.

Trên Fig.15 bộ tạo tín hiệu dư bao gồm bộ phận biến đổi trộn giảm 250.

Bộ tạo tín hiệu dư 200 được làm thích ứng để điều khiển hai hoặc nhiều hơn hai bước lặp lại:

Với mỗi bước lặp lại, bộ phận giải mã theo tham số 230 được làm thích ứng để xác định chính xác một tín hiệu đối tượng âm thanh trong số nhiều tín hiệu đối tượng âm thanh được ước tính.

Ngoài ra, với bước lặp lại đã nêu, bộ phận ước tính phần dư 240 được làm thích ứng để xác định chính xác một tín hiệu dư trong số nhiều tín hiệu dư bằng cách biến đổi tín hiệu đối tượng âm thanh đã nêu của nhiều tín hiệu đối tượng âm thanh được ước tính.

Hơn nữa, với bước lặp lại đã nêu, bộ phận biến đổi trộn giảm 250 được làm thích ứng để biến đổi ba hoặc nhiều hơn ba tín hiệu trộn giảm.

Trong bước lặp lại tiếp theo sau bước lặp lại đã nêu, bộ phận giải mã theo tham số 230 được làm thích ứng để xác định chính xác một tín hiệu đối tượng âm thanh trong số nhiều tín hiệu đối tượng âm thanh được ước tính dựa trên ba hoặc nhiều hơn ba tín hiệu trộn giảm mà đã được biến đổi.

Fig.16 minh họa bộ giải mã theo phương án của sáng chế sử dụng khái niệm nổi tảng. Trên Fig.16 bộ giải mã lại bao gồm bộ phận biến đổi trộn giảm 140.

Bộ giải mã theo Fig.16 được làm thích ứng để điều khiển hai hoặc nhiều hơn hai bước lặp lại:

Với mỗi bước lặp lại, bộ phận giải mã theo tham số 110 được làm thích ứng để xác định chính xác một tín hiệu đối tượng âm thanh trong số nhiều tín hiệu đối tượng âm thanh được ước tính thứ nhất.

Ngoài ra, với bước lặp lại đã nêu, bộ phận xử lý phần dư 120 có thể được làm thích ứng để xác định chính xác một tín hiệu đối tượng âm thanh trong số nhiều tín hiệu đối tượng âm thanh được ước tính thứ hai bằng cách biến đổi tín hiệu đối tượng âm thanh đã nêu của nhiều tín hiệu đối tượng âm thanh được ước tính thứ nhất.

Hơn nữa, với bước lặp lại đã nêu, bộ phận biến đổi trộn giảm 140 được làm thích ứng để loại bỏ tín hiệu đối tượng âm thanh đã nêu trong số nhiều tín hiệu đối tượng âm thanh được ước tính thứ hai từ ba hoặc nhiều hơn ba tín hiệu trộn giảm để biến đổi ba hoặc nhiều hơn ba tín hiệu trộn giảm.

Trong bước lặp lại tiếp theo sau bước lặp lại đã nêu, bộ phận giải mã theo tham số 110 được làm thích ứng để xác định chính xác một tín hiệu đối tượng âm thanh trong số nhiều tín hiệu đối tượng âm thanh được ước tính thứ nhất dựa trên ba hoặc nhiều hơn ba tín hiệu trộn giảm mà đã được biến đổi.

Sau đây, kênh dẫn toán học trên ví dụ của khái niệm mã hóa/giải mã dư chung được mô tả:

Kí hiệu sau được sử dụng sau đây:

Các kích thước:

$N_{Objects}$ - số tín hiệu đối tượng âm thanh

N_{DmxCh} - số tín hiệu trộn giảm

$N_{UpmixCh}$ - số kênh trộn tăng

$N_{Samples}$ - số dữ liệu được xử lý

N_{EAO} - số EAO

Các thuật ngữ

$\mathbf{Z}^*$ - bộ điều khiển sao (*) ký hiệu chuyển vị liên hợp của ma trận đã cho.

$\mathbf{S}$ - tín hiệu đối tượng âm thanh gốc được cung cấp cho bộ mã hóa (kích thước $N_{Objects} \times N_{Samples}$)

$\mathbf{D}$ - ma trận trộn giảm (kích thước $N_{DmxCh} \times N_{Objects}$)

$\mathbf{R}$ - ma trận biểu diễn (kích thước $N_{UpmixCh} \times N_{Objects}$)

$\mathbf{X}$ - tín hiệu âm thanh trộn giảm $\mathbf{X} = \mathbf{DS}$ (kích thước $N_{DmxCh} \times N_{Samples}$)

$\mathbf{Y}$ - tín hiệu đầu ra âm thanh lý tưởng $\mathbf{Y} = \mathbf{RS}$ (kích thước $N_{UpmixCh} \times N_{Samples}$)

$\mathbf{S}_{est}$ - lấy tròn tín hiệu đối tượng được khôi phục theo tham số $\mathbf{S}_{est} \square \mathbf{S}$ được định nghĩa là $\mathbf{S}_{est} = \mathbf{GX}$ (kích thước $N_{Objects} \times N_{Samples}$)

$\hat{\mathbf{S}}_{est}$ - đầu ra bộ giải mã bao gồm tất cả tín hiệu non-EAO (được ước tính theo tham số) và EAO (dự cộng theo tham số) ước tính kích thước $N_{Objects} \times N_{Samples}$

$\hat{\mathbf{Y}}_{est}$ - lấy tròn tín hiệu đầu ra âm thanh trộn tăng $\hat{\mathbf{Y}}_{est} \square \mathbf{Y}$ được định nghĩa là $\hat{\mathbf{Y}}_{est} = \mathbf{R}\hat{\mathbf{S}}_{est}$ (kích thước $N_{UpmixCh} \times N_{Samples}$)

$\mathbf{Z}_{nonEao}$; $\mathbf{Z}_{eao}$ - ma trận con bản đồ biểu thị các vị trí của các non-EAO và các EAO trong danh sách tất cả các đối tượng. Lưu ý $\mathbf{Z}_{nonEao}\mathbf{Z}_{eao}^* = [0]$ (kích thước $(N_{Objects} - N_{EAO}) \times N_{Objects}$; $N_{EAO} \times N_{Objects}$). Non-EAO $\mathbf{Z}_{nonEao}$ và các ma trận bản đồ $\mathbf{Z}_{eao}$ tương ứng được định nghĩa là

$$\mathbf{Z}_{nonEao}(i,j) = \begin{cases} 1, & \text{nếu đối tượng } j \text{ là non-EAO thứ } i, \\ 0, & \text{ngược lại} \end{cases}$$

$$\mathbf{Z}_{eao}(i,j) = \begin{cases} 1, & \text{nếu đối tượng } j \text{ là EAO thứ } i, \\ 0, & \text{ngược lại} \end{cases}$$

Ví dụ, với $N_{Objects} = 5$ và các đối tượng số 2 và số 4 là các EAO, các ma trận này là

$$\mathbf{Z}_{nonEao} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{Z}_{eao} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

$\mathbf{D}_{nonEao}$ - ma trận con trộn giảm tương ứng với các non-EAO, được định nghĩa là

$$\mathbf{D}_{nonEao} = \mathbf{D}\mathbf{Z}_{nonEao}^* \text{ (kích thước } N_{DmxCh} \times (N_{Objects} - N_{EAO}))$$

$\mathbf{D}_{eao}$ - ma trận con trộn giảm tương ứng với Các EAO, được định nghĩa là
 $\mathbf{D}_{eao} = \mathbf{D}\mathbf{Z}_{eao}^*$ (kích thước $N_{DmxCh} \times N_{EAO}$)

$\mathbf{G}$ - ma trận ước tính nguồn theo tham số (kích thước $N_{Objects} \times N_{DmxCh}$)

$\mathbf{E}$ - ma trận hiệp phương sai theo đối tượng (kích thước $N_{Objects} \times N_{Objects}$)

$\mathbf{E}_{nonEao}$ - ma trận con hiệp phương sai tương ứng với các non-EAO, được định nghĩa là $\mathbf{E}_{nonEao} = \mathbf{Z}_{nonEao}\mathbf{E}\mathbf{Z}_{nonEao}^*$ (kích thước $(N_{Objects} - N_{EAO}) \times (N_{Objects} - N_{EAO})$)

$\mathbf{S}_{eao}$ - tín hiệu EAO bao gồm các sự khôi phục của các EAO (kích thước $N_{EAO} \times N_{Samples}$)

$\mathbf{S}_{nonEao}$ - tín hiệu non-EAO bao gồm các sự khôi phục của các non-EAO (kích thước $(N_{Objects} - N_{EAO}) \times N_{Samples}$)

$\mathbf{S}_{res}$ - các tín hiệu dư cho các EAO (kích thước $N_{EAO} \times N_{Samples}$)

$\tilde{\mathbf{X}}_{nonEao}$ - tín hiệu trộn giảm được biến đổi chỉ bao gồm các tín hiệu non-EAO; được tính toán như sự khác biệt giữa trộn giảm SAOC và trộn giảm của các EAO được khôi phục (kích thước $N_{DmxCh} \times N_{Samples}$)

Tất cả các ma trận được đưa vào (nói chung) là biến thể thời gian và tần số.

Bây giờ, phương pháp chung với sự ước tính lại tín hiệu non-EAO tại phía bộ giải mã được xem xét:

Phương pháp chung có thể được mô tả như phép tính gần đúng hai bước với việc chiết xuất thứ nhất tất cả các tín hiệu EAO từ tín hiệu trộn giảm tương ứng và sau đó việc khôi phục tất cả các tín hiệu non-EAO đánh giá các EAO. Các tín hiệu đối tượng được thu hồi lại từ các tín hiệu trộn giảm ($\mathbf{X}$) sử dụng PSI ($\mathbf{E}$, $\mathbf{D}$) và tín hiệu dư được sáp nhập ($\mathbf{S}_{res}$).

Đây được xem là tín hiệu đầu ra được biểu diễn cuối cùng $\hat{\mathbf{Y}}_{est}$ được đưa ra như:

$$\hat{\mathbf{Y}}_{est} = \mathbf{R}\hat{\mathbf{S}}_{est}.$$

Tín hiệu đối tượng đầu ra theo bộ giải mã $\hat{\mathbf{S}}_{est}$ có thể được biểu diễn lại như tổng sau:

$$\hat{\mathbf{S}}_{est} = \mathbf{Z}_{eao}^* \mathbf{S}_{eao} + \mathbf{Z}_{nonEao}^* \mathbf{S}_{nonEao}.$$

Tín hiệu EAO $\mathbf{S}_{eao}$ được tính từ việc trộn giảm $\mathbf{X}$ với sự trợ giúp của ma trận khôi phục EAO theo tham số $\mathbf{G}_{eao}$ và các phần dư EAO tương ứng $\mathbf{S}_{res}$ như sau đây:

$$\mathbf{S}_{eao} = \mathbf{G}_{eao} \mathbf{X} + \mathbf{S}_{res}.$$

Tín hiệu non-EAO $\mathbf{S}_{nonEao}$ được tính toán từ trộn giảm được biến đổi $\tilde{\mathbf{X}}_{nonEao}$ với sự trợ giúp của ma trận khôi phục non-EAO theo tham số $\tilde{\mathbf{G}}_{nonEao}$ như sau đây:

$$\mathbf{S}_{nonEao} = \tilde{\mathbf{G}}_{nonEao} \tilde{\mathbf{X}}_{nonEao}.$$

Tín hiệu $\tilde{\mathbf{X}}_{nonEao}$ trộn giảm được biến đổi được xác định như sự khác biệt giữa sự trộn giảm $\mathbf{X}$ và sự trộn giảm tương ứng của các EAO được khôi phục như sau, do việc xóa bỏ các EAO từ tín hiệu trộn giảm $\mathbf{X}$:

$$\tilde{\mathbf{X}}_{nonEao} = \mathbf{X} - \mathbf{D}\mathbf{Z}_{eao}^* \mathbf{S}_{eao}.$$

Ở đây các ma trận khôi phục đối tượng theo tham số cho các EAO $\mathbf{G}_{eao}$ và các non-EAO $\tilde{\mathbf{G}}_{nonEao}$ được xác định sử dụng PSI ($\mathbf{E}$, $\mathbf{D}$) như sau:

$$\mathbf{G}_{eao} = \mathbf{Z}_{eao} \mathbf{E} \mathbf{D}^* \mathbf{J}, \quad \mathbf{J} \approx (\mathbf{D} \mathbf{E} \mathbf{D}^*)^{-1}.$$

$$\tilde{\mathbf{G}}_{nonEao} = \mathbf{E}_{nonEao} \mathbf{D}_{nonEao}^* \mathbf{J}_{nonEao}, \quad \mathbf{J}_{nonEao} \approx (\mathbf{D}_{nonEao} \mathbf{E}_{nonEao} \mathbf{D}_{nonEao}^*)^{-1}.$$

Sau đây, phương pháp đơn giản hóa “A” không bao gồm việc ước tính lại tín hiệu non-EAO tại phía bộ giải mã được mô tả:

Nếu chỉ các EAO trong hỗn hợp tín hiệu được điều khiển, cảnh mục tiêu có thể diễn giải như tổ hợp tuyến tính của các tín hiệu trộn giảm và các tín hiệu EAO. Việc ước tính lại bổ sung các tín hiệu non-EAO do đó có thể được bỏ qua. Phương pháp chung với việc ước tính lại tín hiệu non-EAO có thể được đơn giản hóa đến phương pháp một bước:

$$\hat{\mathbf{S}}_{est} = \mathbf{S}_{est} + \mathbf{X}_{dif}.$$

Tín hiệu $\mathbf{X}_{dif} = f(\mathbf{S}_{res}, \mathbf{D})$ bao gồm các tín hiệu dư được truyền dẫn của các EAO và các số hạng bù dư để định nghĩa sau được giữ:

$$\mathbf{D}\hat{\mathbf{S}}_{est} = \mathbf{X}.$$

Đây là điều kiện đủ để biểu diễn bất kỳ cảnh âm thanh nào mà được hạn chế để chỉ điều khiển các EAO.

Với $\mathbf{D}\hat{\mathbf{S}}_{est} = \mathbf{D}(\mathbf{S}_{est} + \mathbf{X}_{dif}) = \mathbf{X}$ và $\mathbf{D}\mathbf{S}_{est} = \mathbf{X}$, điều kiện ràng buộc sau đây cho số hạng $\mathbf{X}_{dif}$ phải được thỏa mãn:

$$\mathbf{D}\mathbf{X}_{dif} = 0.$$

Số hạng $\mathbf{X}_{dif}$ bao gồm các thành phần mà được xác định bởi bộ mã hóa (và được truyền dẫn hoặc lưu trữ) $\mathbf{S}_{res}$ và các thành phần $\mathbf{X}_{nonEao}$ được xác định sử dụng phương trình này.

Sử dụng các định nghĩa của ma trận trộn giảm ($\mathbf{D} = \mathbf{D}_{eao}\mathbf{Z}_{eao} + \mathbf{D}_{nonEao}\mathbf{Z}_{nonEao}$) và số hạng bù ($\mathbf{X}_{dif} = \mathbf{Z}_{eao}^*\mathbf{S}_{res} + \mathbf{Z}_{nonEao}^*\mathbf{X}_{nonEao}$) có thể nhận được phương trình sau:

$$\mathbf{D}\mathbf{X}_{dif} = \mathbf{D}_{eao}\mathbf{Z}_{eao}\mathbf{Z}_{eao}^*\mathbf{S}_{res} + \mathbf{D}_{nonEao}\mathbf{Z}_{nonEao}\mathbf{Z}_{nonEao}^*\mathbf{X}_{nonEao} + \mathbf{D}_{eao}\mathbf{Z}_{eao}\mathbf{Z}_{nonEao}^*\mathbf{X}_{nonEao} + \mathbf{D}_{nonEao}\mathbf{Z}_{nonEao}\mathbf{Z}_{eao}^*\mathbf{S}_{res} = 0$$

Với $\mathbf{Z}_{eao}\mathbf{Z}_{eao}^* = \mathbf{I}$, $\mathbf{Z}_{nonEao}\mathbf{Z}_{nonEao}^* = \mathbf{I}$ và $\mathbf{Z}_{nonEao}\mathbf{Z}_{eao}^* = [0]$, $\mathbf{Z}_{eao}\mathbf{Z}_{nonEao}^* = [0]$, phương trình có thể được đơn giản hóa thành:

$$\mathbf{D}_{eao}\mathbf{S}_{res} + \mathbf{D}_{nonEao}\mathbf{X}_{nonEao} = 0.$$

Giải phương trình tuyến tính cho $\mathbf{X}_{nonEao}$ đưa ra:

$$\mathbf{X}_{nonEao} = -\left(\mathbf{D}_{nonEao}^* \mathbf{D}_{nonEao}\right)^{-1} \mathbf{D}_{nonEao}^* \mathbf{D}_{eao} \mathbf{S}_{res}.$$

Sau khi giải hệ phương trình tuyến tính cảnh mục tiêu được mong muốn có thể được tính toán như tổng sau của số hạng dự đoán tham số và số hạng nâng cao dư như sau:

$$\hat{\mathbf{Y}}_{est} = \mathbf{R} \hat{\mathbf{S}}_{est}, \quad \hat{\mathbf{S}}_{est} = \mathbf{S}_{est} + \mathbf{X}_{dif}, \quad \mathbf{X}_{dif} = \mathbf{Z}_{eao}^* \mathbf{S}_{res} - \mathbf{Z}_{nonEao}^* \left(\mathbf{D}_{nonEao}^* \mathbf{D}_{nonEao}\right)^{-1} \mathbf{D}_{nonEao}^* \mathbf{D}_{eao} \mathbf{S}_{res}.$$

Sau đây, phương pháp đơn giản hóa “B” không bao gồm việc ước tính lại tín hiệu non-EAO tại phía bộ giải mã được đề xuất:

Xem xét số hạng bù $\mathbf{X}_{dif}$ như ở trên ($\hat{\mathbf{S}}_{est} = \mathbf{S}_{est} + \mathbf{X}_{dif}$) cho dự đoán tín hiệu tham số $\mathbf{S}_{est}$ và thể hiện nó như công thức sau $\mathbf{X}_{dif} = \mathbf{H}_{enh} \mathbf{Z}_{eao}^* \mathbf{S}_{res}$ của các tín hiệu dư $\mathbf{S}_{res}$ dẫn vào:

$$\hat{\mathbf{S}}_{est} = \mathbf{S}_{est} + \mathbf{H}_{enh} \mathbf{Z}_{eao}^* \mathbf{S}_{res}$$

Công thức thay thế bao gồm ba phần sau gồm tổ hợp tuyến tính thích hợp của các tín hiệu trộn giảm ($\mathbf{H}_{dmx} \mathbf{X}$), các đối tượng được tăng cường ($\mathbf{H}_{enh} \mathbf{Z}_{eao}^* \mathbf{Z}_{eao} \mathbf{S}_{enh}$), và các đối tượng không được tăng cường ($\mathbf{H}_{est} \mathbf{S}_{est}$) như sau:

$$\hat{\mathbf{S}}_{est} = \mathbf{H}_{dmx} \mathbf{X} + \mathbf{H}_{enh} \mathbf{Z}_{eao}^* \mathbf{Z}_{eao} \mathbf{S}_{enh} + \mathbf{H}_{est} \mathbf{S}_{est}.$$

Các ma trận có các kích thước $\mathbf{H}_{dmx} : N_{Objects} \times N_{DmxCh}$, $\mathbf{H}_{enh} : N_{Objects} \times N_{Objects}$, $\mathbf{S}_{enh} : N_{Objects} \times N_{Samples}$, và $\mathbf{H}_{est} : N_{Objects} \times N_{Objects}$.

Giả sử $\mathbf{D} \mathbf{S}_{est} = \mathbf{X}$ và định nghĩa của $\mathbf{S}_{enh} = \mathbf{S}_{est} + \mathbf{Z}_{eao}^* \mathbf{S}_{res}$, điều này có thể được viết như sau:

$$\hat{\mathbf{S}}_{est} = \left(\mathbf{H}_{dmx} \mathbf{D} + \mathbf{H}_{enh} \mathbf{Z}_{eao}^* \mathbf{Z}_{eao} + \mathbf{H}_{est}\right) \mathbf{S}_{est} + \mathbf{H}_{enh} \mathbf{Z}_{eao}^* \mathbf{S}_{res}.$$

So sánh với điều này, và định nghĩa trước đó của các tín hiệu được khôi phục $\hat{\mathbf{S}}_{est} = \mathbf{S}_{est} + \mathbf{H}_{enh} \mathbf{Z}_{eao}^* \mathbf{S}_{res}$, suy ra rằng:

$$\mathbf{H}_{dmx} \mathbf{D} + \mathbf{H}_{enh} \mathbf{Z}_{eao}^* \mathbf{Z}_{eao} + \mathbf{H}_{est} = \mathbf{I}.$$

Có thể nhận được số hạng $\mathbf{H}_{est}$ là:

$$\mathbf{H}_{est} = \mathbf{I} - \mathbf{H}_{ext} \mathbf{D}_{ext}.$$

Lỗi trong việc khôi phục cuối cùng sẽ được giảm thiểu, trong khi việc tạo các tín hiệu không nâng cao được giảm thiểu. Do đó, nhắm mục tiêu cho $\mathbf{H}_{est} \square 0$ cho phép giải các số hạng $\mathbf{H}_{ext}$ từ hệ phương trình tuyến tính:

$$\mathbf{H}_{ext} = \mathbf{D}_{ext}^* (\mathbf{D}_{ext} \mathbf{D}_{ext}^*)^{-1},$$

tại đó ma trận trộn giảm được mở rộng $\mathbf{D}_{ext}$ và ma trận trộn tăng $\mathbf{H}_{ext}$ được định nghĩa như các ma trận liên kết:

$$\mathbf{D}_{ext} = \begin{bmatrix} \mathbf{D} \\ \mathbf{Z}_{eao}^* \mathbf{Z}_{eao} \end{bmatrix} \text{ và } \mathbf{H}_{ext} = [\mathbf{H}_{dmx} \quad \mathbf{H}_{enh}], \text{ và do đó } \mathbf{H}_{enh} = \mathbf{H}_{ext} \begin{bmatrix} \mathbf{0}^{N_{DmxCh} \times N_{Objects}} \\ \mathbf{I}^{N_{Objects} \times N_{Objects}} \end{bmatrix}$$

Sau khi giải hệ phương trình tuyến tính số hạng điều chỉnh mong muốn $\mathbf{X}_{dif}$ có thể thu được là:

$$\mathbf{X}_{dif} = \mathbf{D}_{ext}^* (\mathbf{D}_{ext} \mathbf{D}_{ext}^*)^{-1} \begin{bmatrix} \mathbf{0}^{N_{DmxCh} \times N_{Objects}} \\ \mathbf{I}^{N_{Objects} \times N_{Objects}} \end{bmatrix} \mathbf{Z}_{eao}^* \mathbf{S}_{res},$$

Dẫn đến kết quả đầu ra cuối cùng là $\hat{\mathbf{Y}}_{est} = \mathbf{R} \hat{\mathbf{S}}_{est}$, $\hat{\mathbf{S}}_{est} = \mathbf{S}_{est} + \mathbf{X}_{dif}$.

Sau đây, phương pháp được đơn giản hóa “C” được xem xét:

Nếu chỉ các EAO được xử lý một cách tùy tiện, cảnh mục tiêu bất kỳ có thể được tạo ra bởi tổ hợp tuyến tính của các tín hiệu trộn giảm và các EAO. Lưu ý rằng thay vì trộn giảm, trộn giảm với các EAO được loại bỏ có thể được sử dụng. Cảnh mục tiêu có thể được tạo ra hoàn toàn nếu việc xử lý phần dư trả lại hoàn toàn các EAO. Việc biểu

diễn cảnh mục tiêu bất kỳ có thể được thực hiện bằng cách tìm các ma trận biểu diễn hai thành phần $\mathbf{R}_D$ và $\mathbf{R}_{eao}$ cho việc trộn giảm và khôi phục EAO. Các ma trận có các kích thước

$\mathbf{R}_D : N_{UpmixCh} \times N_{DmxCh}$ và $\mathbf{R}_{eao} : N_{UpmixCh} \times N_{EAO}$. Ma trận biểu diễn mục tiêu $\mathbf{R}$ có thể được biểu diễn như sản phẩm của các ma trận biểu diễn được tổ hợp và ma trận trộn giảm như

$$\mathbf{R} = \begin{bmatrix} \mathbf{R}_D & \mathbf{R}_{eao} \end{bmatrix} \begin{bmatrix} \mathbf{D} \\ \mathbf{Z}_{eao}^* \mathbf{Z}_{eao} \end{bmatrix} = \mathbf{R}_{ext} \mathbf{D}_{ext}$$

Từ đây, $\mathbf{R}_{ext}$ có thể được giải với

$$\mathbf{R}_{ext} = \mathbf{R} \mathbf{D}_{ext}^* \left(\mathbf{D}_{ext} \mathbf{D}_{ext}^* \right)^{-1}$$

và các ma trận con $\mathbf{R}_D$ và $\mathbf{R}_{eao}$ có thể được trích xuất từ kết quả với

$$\mathbf{R}_D = \mathbf{R}_{ext} \begin{bmatrix} \mathbf{I}^{N_{DmxCh} \times N_{DmxCh}} \\ \mathbf{0}^{N_{Objects} \times N_{DmxCh}} \end{bmatrix} \quad \text{và} \quad \mathbf{R}_{eao} = \mathbf{R}_{ext} \begin{bmatrix} \mathbf{0}^{(N_{Objects} + N_{DmxCh} - N_{EAO}) \times N_{EAO}} \\ \mathbf{I}^{N_{EAO} \times N_{EAO}} \end{bmatrix}$$

Cảnh mục tiêu bây giờ có thể được tính là:

$$\hat{\mathbf{Y}}_{est} = \mathbf{R}_D \mathbf{X} + \mathbf{R}_{eao} \mathbf{S}_{eao},$$

trong đó $\mathbf{S}_{eao}$ chứa các khôi phục đầy đủ của các EAO và được định nghĩa (như trước đây) $\mathbf{S}_{eao} = \mathbf{G}_{eao} \mathbf{X} + \mathbf{S}_{res}$.

Phương trình đơn giản có thể được tạo công thức để biểu diễn mục tiêu sử dụng trộn giảm cùng với xóa bỏ các EAO từ hỗn hợp bằng việc trừ $\mathbf{D}_{eao} \mathbf{S}_{eao}$ từ trộn giảm.

Sau đây, dẫn xuất toán học khác và chi tiết hơn về khái niệm mã hóa/giải mã dư chung được mô tả và sự thống nhất giữa phương pháp tạo ra và sự đơn giản hóa "A" được đề xuất.

Từ bây giờ trong bản mô tả, các ký hiệu sau đây được áp dụng. Nếu vì một số yếu tố, các ký hiệu sau đây là không phù hợp với các ký hiệu cung cấp ở trên, từ bây giờ trong bản mô tả, chỉ có các ký hiệu sau đây được áp dụng cho các yếu tố này.

Các định nghĩa:

S là các tín hiệu đối tượng có kích thước $N_{Objects} \times N_{Samples}$

E = SS* là ma trận hiệp phương sai đối tượng có kích thước $N_{Objects} \times N_{Objects}$

D là ma trận trộn giảm có kích thước $N_{DmxCh} \times N_{Objects}$

X = DS là tín hiệu trộn giảm có kích thước $N_{DmxCh} \times N_{Samples}$

G = ED*J là ma trận trộn tăng có kích thước $N_{Objects} \times N_{DmxCh}$

M_{ren} là ma trận biểu diễn có kích thước $N_{UpmixCh} \times N_{Objects}$

X_{res} là các tín hiệu dư có kích thước $N_{EAO} \times N_{Samples}$

R_{eao} là ma trận có kích thước $N_{EAO} \times N_{Objects}$ biểu thị các vị trí (điểm) của các EAO được định nghĩa là

$$\mathbf{R}_{eao}(i, j) = \begin{cases} 1 \\ 0 \end{cases},$$

R_{nonEao} là ma trận có kích thước $(N_{Objects} - N_{EAO}) \times N_{Objects}$ biểu thị các vị trí (điểm) của các non-EAO được định nghĩa là

$$\mathbf{R}_{nonEao}(i, j) = \begin{cases} 1, & \text{nếu đối tượng } j \text{ là non-EAO thứ } i, \\ 0 & \text{ngược lại} \end{cases}$$

Các ma trận con của một số tương ứng với các non-EAO ở trên có thể được chỉ rõ với hỗ trợ của các ma trận chọn lọc **R_{nonEao}** như:

$$\mathbf{E}_{nonEao} = \mathbf{R}_{nonEao} \mathbf{E} \mathbf{R}_{nonEao}^*$$

$$\mathbf{D}_{nonEao} = \mathbf{D}\mathbf{R}_{nonEao}^*$$

$$\begin{aligned}\mathbf{G}_{nonEao} &= \mathbf{E}_{nonEao} \mathbf{D}_{nonEao}^* \mathbf{J}_{nonEao} = \mathbf{E}_{nonEao} \mathbf{D}_{nonEao}^* (\mathbf{D}_{nonEao} \mathbf{E}_{nonEao} \mathbf{D}_{nonEao}^*)^{-1} \\ &= \mathbf{R}_{nonEao} \mathbf{E} \mathbf{R}_{nonEao}^* \mathbf{R}_{nonEao} \mathbf{D}^* (\mathbf{D} \mathbf{R}_{nonEao}^* \mathbf{R}_{nonEao} \mathbf{E} \mathbf{R}_{nonEao}^* \mathbf{R}_{nonEao} \mathbf{D}^*)^{-1}\end{aligned}$$

Sau đây, mô tả toán học cụ thể khác về phương pháp chung (với việc ước tính lại tín hiệu non-EAO tại bộ giải mã) được đề xuất:

Các tín hiệu đối tượng được thu hồi từ việc trộn giảm sử dụng thông tin phụ và các tín hiệu dư được sáp nhập. Đầu ra từ bộ giải mã $\hat{\mathbf{X}}$ được đưa ra như sau:

$$\hat{\mathbf{X}} = \mathbf{M}_{ren} \mathbf{R}_{eao}^* \mathbf{X}_{eao} + \mathbf{M}_{ren} \mathbf{R}_{nonEao}^* \mathbf{X}_{nonEao}.$$

Số hạng EAO $\mathbf{X}_{eao}$ có kích thước N_{EAO} với các EAO được tính toán như sau

$$\mathbf{X}_{eao} = \mathbf{R}_{eao} \mathbf{E} \mathbf{D}^* \mathbf{J} \mathbf{X} + \mathbf{X}_{res},$$

trong đó số hạng tín hiệu dư $\mathbf{X}_{res}$ có kích thước N_{EAO} chứa các tín hiệu dư các EAO.

Số hạng non-EAO $\mathbf{X}_{nonEao}$ có kích thước $N_{Objects} - N_{EAO}$ chứa các non-EAO được tính toán như sau

$$\mathbf{X}_{nonEao} = \mathbf{E}_{nonEao} \mathbf{D}_{nonEao}^* \mathbf{J}_{nonEao} \tilde{\mathbf{X}}_{nonEao}, \quad \mathbf{J}_{nonEao} \approx (\mathbf{D}_{nonEao} \mathbf{E}_{nonEao} \mathbf{D}_{nonEao}^*)^{-1}.$$

trong đó tín hiệu trộn giảm được biến đổi $\tilde{\mathbf{X}}_{nonEao}$ chỉ bao gồm các tín hiệu non-EAO được tính như sự khác biệt giữa trộn giảm SAOC và trộn giảm các EAO được khôi phục

$$\tilde{\mathbf{X}}_{nonEao} = \mathbf{X} - \mathbf{D} \mathbf{R}_{eao}^* \mathbf{X}_{eao}.$$

Ma trận con hiệp phương sai $\mathbf{E}_{nonEao}$ có kích thước $(N_{Objects} - N_{EAO}) \times (N_{Objects} - N_{EAO})$ tương ứng với các non-EAO được tính như sau

$$\mathbf{E}_{nonEao} = \mathbf{R}_{nonEao} \mathbf{E} \mathbf{R}_{nonEao}^*.$$

Ma trận con hiệp phương sai $\mathbf{D}_{nonEao}$ có kích thước $N_{DmxCh} \times (N_{Objects} - N_{EAO})$ tương ứng với các non-EAO được tính như sau

$$\mathbf{D}_{nonEao} = \mathbf{D}\mathbf{R}_{nonEao}^*.$$

Sau đây, mô tả toán học cụ thể khác về phương pháp đơn giản hóa "A" (không gồm ước tính lại tín hiệu non-EAO tại bộ giải mã) được đề xuất:

Các tín hiệu đối tượng được thu hồi từ việc trộn giảm sử dụng thông tin phụ và các tín hiệu dư được sáp nhập. Đầu ra cuối cùng từ bộ giải mã $\hat{\mathbf{X}}$ được đưa ra như sau:

$$\hat{\mathbf{X}} = \mathbf{M}_{ren}(\mathbf{E}\mathbf{D}^*\mathbf{J}\mathbf{X} + \mathbf{X}_{dif}).$$

Số hạng $\mathbf{X}_{dif}$ có kích thước $N_{Objects}$ sáp nhập N_{EAO} tín hiệu dư $\mathbf{X}_{res}$ cho các EAO và số hạng dự đoán $\mathbf{X}_{nonEao}$ cho các non-EAO như sau

$$\mathbf{X}_{dif} = \mathbf{R}_{eao}^* \mathbf{X}_{res} + \mathbf{R}_{nonEao}^* \mathbf{X}_{nonEao}.$$

Số hạng dự đoán $\mathbf{X}_{nonEao}$ được ước tính như sau

$$\mathbf{X}_{nonEao} = -(\mathbf{D}_{nonEao}^* \mathbf{D}_{nonEao})^{-1} \mathbf{D}_{nonEao}^* \mathbf{D}_{eao} \mathbf{X}_{res}$$

Ma trận con trộn giảm $\mathbf{D}_{eao}$ tương ứng với các EAO và $\mathbf{D}_{nonEao}$ tương ứng với các đối tượng chính quy được định nghĩa như sau

$$\mathbf{D} = \mathbf{D}_{eao} \mathbf{R}_{eao} + \mathbf{R}_{nonEao} \mathbf{D}_{nonEao}.$$

Sau đây, trường hợp đặc biệt của ma trận biểu diễn 1 được xem xét:

Xem xét trường hợp đặc biệt sau đây của ma trận biểu diễn trộn giảm đơn giản $\mathbf{M}_D$ có kích thước $N_{DmxCh} \times N_{Objects}$ với biến đổi tùy ý các EAO và chỉ có tỉ lệ thống nhất (được so sánh với trộn giảm) của các non-EAO

$$\mathbf{M}_D = \mathbf{M}\mathbf{R}_{eao}^* \mathbf{R}_{eao} + a\mathbf{D}\mathbf{R}_{nonEao}^* \mathbf{R}_{nonEao}.$$

Bây giờ, mô tả toán học chi tiết của phương pháp chung được đề xuất:

$$\begin{aligned}
\hat{X} &= M_D (R_{eao}^* X_{eao} + R_{nonEao}^* X_{nonEao}) \\
&= M_D R_{eao}^* (R_{eao} ED^* JX + X_{res}) + M_D R_{nonEao}^* G_{nonEao} (X - DR_{eao}^* X_{eao}) \\
&= M_D R_{eao}^* (R_{eao} ED^* JX + X_{res}) + M_D R_{nonEao}^* G_{nonEao} (X - DR_{eao}^* (R_{eao} ED^* JX + X_{res})) \\
&= MR_{eao}^* (R_{eao} ED^* JX + X_{res}) + a DR_{nonEao}^* G_{nonEao} (X - DR_{eao}^* (R_{eao} ED^* JX + X_{res})) \\
&= MR_{eao}^* (R_{eao} ED^* JX + X_{res}) + \\
&\quad a DR_{nonEao}^* R_{nonEao} ER_{nonEao}^* R_{nonEao} D^* (D R_{nonEao}^* R_{nonEao} ER_{nonEao}^* R_{nonEao} D^*)^{-1} (X - DR_{eao}^* (R_{eao} ED^* JX + X_{res})) \\
&= MR_{eao}^* (R_{eao} ED^* JX + X_{res}) + a (X - DR_{eao}^* (R_{eao} ED^* JX + X_{res})) \\
&= MR_{eao}^* X_{eao} + a (X - DR_{eao}^* X_{eao})
\end{aligned}$$

Bây giờ, mô tả toán học chi tiết của phương pháp được đơn giản hóa "A" được đề xuất:

$$\begin{aligned}
\hat{X} &= M_D (GX + X_{dif}) \\
&= M_D (GX + R_{eao}^* X_{res} + R_{nonEao}^* X_{nonRes}) \\
&= M_D (GX + R_{eao}^* X_{res} - R_{nonEao}^* (D_{nonEao}^* D_{nonEao})^{-1} D_{nonEao}^* D_{eao} X_{res}) \\
&= M_D (GX + R_{eao}^* X_{res} - R_{nonEao}^* D_{nonEao}^* (D_{nonEao} D_{nonEao}^*)^{-1} D_{eao} X_{res}) \\
&= M_D (R_{eao}^* R_{eao} GX + R_{eao}^* X_{res} + R_{nonEao}^* R_{nonEao} GX - R_{nonEao}^* D_{nonEao}^* (D_{nonEao} D_{nonEao}^*)^{-1} D_{eao} X_{res}) \\
&= M_D (R_{eao}^* X_{eao} + R_{nonEao}^* (R_{nonEao} GX - D_{nonEao}^* (D_{nonEao} D_{nonEao}^*)^{-1} D_{eao} X_{res})) \\
&= MR_{eao}^* X_{eao} + a DR_{nonEao}^* R_{nonEao} R_{nonEao}^* (R_{nonEao} GX - D_{nonEao}^* (D_{nonEao} D_{nonEao}^*)^{-1} D_{eao} X_{res}) \\
&= MR_{eao}^* X_{eao} + a DR_{nonEao}^* R_{nonEao} GX - a D_{nonEao}^* D_{nonEao}^* (D_{nonEao} D_{nonEao}^*)^{-1} D_{eao} X_{res} \\
&= MR_{eao}^* X_{eao} + a DR_{nonEao}^* R_{nonEao} GX - a D_{eao} X_{res} \\
&= MR_{eao}^* X_{eao} + a (X - DR_{eao}^* R_{eao} GX) - a D_{eao} X_{res} \\
&= MR_{eao}^* X_{eao} + a (X - DR_{eao}^* X_{eao})
\end{aligned}$$

Có thể thấy rằng hai kết quả là giống nhau khi giả thiết của ma trận biểu diễn giữ nguyên.

Bây giờ, trường hợp đặc biệt của ma trận biểu diễn 2 được xem xét:

Bao gồm điều kiện bổ sung về cấu trúc của ma trận biểu diễn M_S có kích thước $N_{DmxCh} \times N_{Objects}$: tất cả các non-EAO được biến đổi chỉ bởi hệ số tỷ lệ thông thường a được so với trộn giảm, và tất cả các EAO cũng được biến đổi chỉ bởi hệ số tỷ lệ thông thường b được so với trộn giảm.

$$\mathbf{M}_D = b\mathbf{D}\mathbf{R}_{eao}^* \mathbf{R}_{eao} + a\mathbf{D}\mathbf{R}_{nonEao}^* \mathbf{R}_{nonEao} = \mathbf{D}(b\mathbf{R}_{eao}^* \mathbf{R}_{eao} + a\mathbf{R}_{nonEao}^* \mathbf{R}_{nonEao}).$$

Tiếp tục từ các kết quả trước đó, đầu ra của hệ thống sẽ là

$$\begin{aligned}\hat{\mathbf{X}} &= b\mathbf{D}\mathbf{R}_{eao}^* \mathbf{X}_{eao} + a(\mathbf{X} - \mathbf{D}\mathbf{R}_{eao}^* \mathbf{X}_{eao}) \\ &= a\mathbf{X} + (b-a)\mathbf{D}\mathbf{R}_{eao}^* \mathbf{X}_{eao} \\ &= a\mathbf{X} + (b-a)\mathbf{D}\mathbf{R}_{eao}^* (\mathbf{R}_{eao} \mathbf{E} \mathbf{D}^* \mathbf{J} \mathbf{X} + \mathbf{X}_{res})\end{aligned}$$

Mặc dù một vài khía cạnh đã được mô tả trong ngữ cảnh của thiết bị, rõ ràng là 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 phương pháp hoặc dấu hiệu của bước phương pháp. Tương tự, các khía cạnh được mô tả trong ngữ cảnh của bước phương pháp cũng thể hiện phần mô tả của khối tương ứng hoặc mục hoặc dấu hiệu của thiết bị tương ứng.

Tín hiệu đã được khai triển theo sáng chế có thể được lưu trữ trên môi trường lưu trữ số hoặc có thể được truyền trên môi trường truyền dẫn như môi trường truyền dẫn không dây hoặc môi trường truyền dẫn có dây như Liên mạng.

Phụ thuộc vào một số yêu cầu thực thi, các phương án theo sáng chế có thể được thực thi trong phần cứng hoặc phần mềm. Việc thực thi này 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 các tín hiệu điều khiển có thể đọc kiểu điện tử được lưu trữ trên đó, mà liên kết (hoặc có khả năng liên kết) 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 chuyển tiếp có các tín hiệu điều khiển có thể đọc được kiểu điện tử, mà có khả năng kết hợp với hệ thống máy tính lập trình được, do đó một trong các phương pháp được mô tả ở đây được thực hiện.

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

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

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

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

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

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 làm thích ứng để thực hiện một trong các phương pháp được mô tả ở đây.

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

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

Các phương án nêu trên chỉ đơn thuần minh họa các nguyên lý của sáng chế. Cần hiểu rằng các sự biến thể và biến đổi của các phương án và các chi tiết được mô tả ở đây sẽ là rõ ràng đối với người có trình độ trong lĩnh vực kỹ thuật tương ứng. Do đó

sáng chế chỉ bị giới hạn bởi phạm vi của các điểm yêu cầu bảo hộ và mà không bị giới hạn bởi các chi tiết cụ thể được thể hiện bởi bản mô tả và giải thích các phương án ở đây.

Sự tham chiếu trong phần mô tả này tới bất kỳ công bố nào trước đó (hoặc thông tin có nguồn gốc từ nó), hoặc cho bất kỳ vấn đề nào được biết, là không phải và không nên được coi là sự xác nhận hoặc thừa nhận hoặc bất kỳ hình thức gợi ý nào mà công bố trước đó (hoặc thông tin có nguồn gốc từ nó) hoặc dạng đã biết tạo thành một phần của kiến thức chung phổ biến trong lĩnh vực mà phần mô tả này liên quan.

Xuyên suốt bản mô tả này và các yêu cầu bảo hộ sau đây, trừ khi có yêu cầu khác, từ "bao gồm" sẽ được hiểu là bao gồm một số nguyên hoặc bước hoặc nhóm các số nguyên đã nêu hoặc các bước nhưng không loại trừ bất kỳ số nguyên hoặc bước hoặc nhóm các số nguyên hoặc bước nào khác.

Các số tham chiếu trong các yêu cầu bảo hộ sau đây không giới hạn phạm vi của các yêu cầu bảo hộ tương ứng.

YÊU CẦU BẢO HỘ

1. Bộ giải mã, bao gồm:

bộ phận giải mã theo tham số (110) để tạo ra nhiều tín hiệu đối tượng âm thanh được ước tính thứ nhất bằng cách trộn tăng ba hoặc nhiều hơn ba tín hiệu trộn giảm, trong đó ba hoặc nhiều hơn ba tín hiệu trộn giảm mã hóa nhiều tín hiệu đối tượng âm thanh gốc, trong đó bộ phận giải mã theo tham số (110) được tạo cấu hình để trộn tăng ba hoặc nhiều hơn ba tín hiệu trộn giảm phụ thuộc vào thông tin phụ theo tham số biểu thị thông tin trên nhiều tín hiệu đối tượng âm thanh gốc, và

bộ phận xử lý phần dư (120) để tạo ra nhiều tín hiệu đối tượng âm thanh được ước tính thứ hai bằng cách biến đổi một hoặc nhiều tín hiệu đối tượng âm thanh được ước tính thứ nhất, trong đó bộ phận xử lý phần dư (120) được tạo cấu hình để biến đổi một hoặc nhiều tín hiệu đối tượng âm thanh được ước tính thứ nhất đã nêu phụ thuộc vào một hoặc nhiều tín hiệu dư.

2. Bộ giải mã theo điểm 1,

trong đó bộ phận xử lý phần dư (120) được tạo cấu hình để biến đổi một hoặc nhiều tín hiệu đối tượng âm thanh được ước tính thứ nhất đã nêu phụ thuộc vào ít nhất ba tín hiệu dư, và

trong đó bộ giải mã được làm thích ứng để tạo ra ít nhất ba kênh đầu ra âm thanh dựa trên nhiều tín hiệu đối tượng âm thanh được ước tính thứ hai.

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

trong đó bộ giải mã còn bao gồm bộ phận biến đổi trộn giảm (140) được làm thích ứng để loại bỏ một hoặc nhiều tín hiệu đối tượng âm thanh của nhiều tín hiệu đối tượng âm thanh được ước tính thứ hai được xác định bằng bộ phận xử lý phần dư (120) từ ba hoặc nhiều hơn ba tín hiệu trộn giảm để thu được ba hoặc nhiều hơn ba tín hiệu trộn giảm được biến đổi, và

trong đó bộ phận giải mã theo tham số (110) được tạo cấu hình để xác định một hoặc nhiều tín hiệu đối tượng âm thanh của các tín hiệu đối tượng âm thanh được ước tính thứ nhất dựa trên ba hoặc nhiều hơn ba tín hiệu trộn giảm được biến đổi.

4. Bộ giải mã theo điểm 3,

trong đó bộ phận biến đổi trộn giảm (140) được làm thích ứng để áp dụng công thức:

$$\tilde{\mathbf{X}}_{nonEAO} = \mathbf{X} - \mathbf{D}\mathbf{Z}_{eao}^* \mathbf{S}_{eao}.$$

để loại bỏ một hoặc nhiều tín hiệu đối tượng âm thanh của nhiều tín hiệu đối tượng âm thanh được ước tính thứ hai được xác định bằng bộ phận xử lý phần dư (120) từ ba hoặc nhiều hơn ba tín hiệu trộn giảm để thu được ba hoặc nhiều hơn ba tín hiệu trộn giảm được biến đổi,

trong đó:

$\mathbf{X}$ biểu thị ba hoặc nhiều hơn ba tín hiệu trộn giảm trước khi được biến đổi

$\tilde{\mathbf{X}}_{nonEAO}$ biểu thị ba hoặc nhiều hơn ba tín hiệu trộn giảm được biến đổi

$\mathbf{D}$ biểu thị thông tin trộn giảm

$\mathbf{S}_{eao}$ bao gồm một hoặc nhiều tín hiệu đối tượng âm thanh đã nêu của nhiều tín hiệu đối tượng âm thanh được ước tính thứ hai, và

$\mathbf{Z}_{eao}^*$ biểu thị các vị trí của một hoặc nhiều tín hiệu đối tượng âm thanh đã nêu của nhiều tín hiệu đối tượng âm thanh được ước tính thứ hai.

5. Bộ giải mã theo điểm 3 hoặc điểm 4,

trong đó bộ giải mã được làm thích ứng để điều khiển hai hoặc nhiều hơn hai bước lặp lại,

trong đó, với mỗi bước lặp lại, bộ phận giải mã theo tham số (110) được làm thích ứng để xác định chính xác một tín hiệu đối tượng âm thanh trong số nhiều tín hiệu đối tượng âm thanh được ước tính thứ nhất,

trong đó với bước lặp lại đã nêu, bộ phận xử lý phần dư (120) được làm thích ứng để xác định chính xác một tín hiệu đối tượng âm thanh trong số nhiều tín hiệu đối tượng âm thanh được ước tính thứ hai bằng cách biến đổi tín hiệu đối tượng âm thanh đã nêu của nhiều tín hiệu đối tượng âm thanh được ước tính thứ nhất,

trong đó, với bước lặp lại đã nêu, bộ phận biến đổi trộn giảm (140) được làm thích ứng để loại bỏ tín hiệu đối tượng âm thanh đã nêu của nhiều tín hiệu đối tượng âm thanh được ước tính thứ hai từ ba hoặc nhiều hơn ba tín hiệu trộn giảm để biến đổi ba hoặc nhiều hơn ba tín hiệu trộn giảm, và

trong đó, với bước lặp lại tiếp theo sau bước lặp lại đã nêu, bộ phận giải mã theo tham số (110) được làm thích ứng để xác định chính xác một tín hiệu đối tượng âm thanh trong số nhiều tín hiệu đối tượng âm thanh được ước tính thứ nhất dựa trên ba hoặc nhiều hơn ba tín hiệu trộn giảm mà đã được biến đổi.

6. Bộ giải mã theo một điểm bất kỳ trong số các điểm trên, trong đó mỗi tín hiệu dư trong số một hoặc nhiều tín hiệu dư biểu thị sự khác biệt giữa một tín hiệu trong số nhiều tín hiệu đối tượng âm thanh gốc và một tín hiệu trong số một hoặc nhiều tín hiệu đối tượng âm thanh được ước tính thứ nhất.

7. Bộ giải mã theo điểm 1 hoặc điểm 2,

trong đó bộ phận xử lý phần dư (120) được làm thích ứng để tạo ra nhiều tín hiệu đối tượng âm thanh được ước tính thứ hai bằng cách biến đổi năm hoặc nhiều hơn năm tín hiệu đối tượng âm thanh được ước tính thứ nhất,

trong đó bộ phận xử lý phần dư (120) được tạo cấu hình để biến đổi năm hoặc nhiều hơn năm tín hiệu đối tượng âm thanh được ước tính thứ nhất đã nêu phụ thuộc vào năm hoặc nhiều hơn năm tín hiệu dư.

8. Bộ giải mã theo điểm 1 hoặc điểm 2, trong đó bộ giải mã được tạo cấu hình để tạo ra bảy hoặc nhiều hơn bảy kênh đầu ra âm thanh dựa trên nhiều tín hiệu đối tượng âm thanh được ước tính thứ hai.

9. Bộ giải mã theo một điểm bất kỳ trong số các điểm trên, trong đó bộ giải mã được làm thích ứng để không xác định Các Hệ Số Dự Đoán Kênh để xác định nhiều tín hiệu đối tượng âm thanh được ước tính thứ hai.

10. Bộ giải mã theo một điểm bất kỳ trong số các điểm trên, trong đó bộ giải mã là bộ giải mã mã hóa đối tượng âm thanh không gian (SAOC - Spatial Audio Object Coding).

11. Bộ tạo tín hiệu dư (200), bao gồm:

bộ phận giải mã theo tham số (230) để tạo ra nhiều tín hiệu đối tượng âm thanh được ước tính bằng cách trộn tăng ba hoặc nhiều hơn ba tín hiệu trộn giảm, trong đó ba hoặc nhiều hơn ba tín hiệu trộn giảm mã hóa nhiều tín hiệu đối tượng âm thanh gốc, trong đó bộ phận giải mã theo tham số (230) được tạo cấu hình để trộn tăng ba hoặc nhiều hơn ba tín hiệu trộn giảm phụ thuộc vào thông tin phụ theo tham số biểu thị thông tin trên nhiều tín hiệu đối tượng âm thanh gốc, và

bộ phận ước tính phần dư (240) để tạo ra nhiều tín hiệu dư dựa trên nhiều tín hiệu đối tượng âm thanh gốc và dựa trên nhiều tín hiệu đối tượng âm thanh được ước tính, sao cho mỗi tín hiệu trong số nhiều tín hiệu dư là tín hiệu khác nhau biểu thị sự khác biệt giữa một tín hiệu trong số nhiều tín hiệu đối tượng âm thanh gốc và một tín hiệu trong số nhiều tín hiệu đối tượng âm thanh được ước tính.

12. Bộ tạo tín hiệu dư (200) theo điểm 11,

trong đó bộ tạo tín hiệu dư (200) còn bao gồm bộ phận biến đổi trộn giảm (250) được làm thích ứng để biến đổi ba hoặc nhiều hơn ba tín hiệu trộn giảm để thu được ba hoặc nhiều hơn ba tín hiệu trộn giảm được biến đổi, và

trong đó bộ phận giải mã theo tham số (230) được tạo cấu hình để xác định một hoặc nhiều tín hiệu đối tượng âm thanh trong số các tín hiệu đối tượng âm thanh được ước tính thứ nhất dựa trên ba hoặc nhiều hơn ba tín hiệu trộn giảm được biến đổi.

13. Bộ tạo tín hiệu dư (200) theo điểm 12, trong đó bộ phận biến đổi trộn giảm (250) được tạo cấu hình để biến đổi ba hoặc nhiều hơn ba tín hiệu trộn giảm gốc để thu được ba hoặc nhiều hơn ba tín hiệu trộn giảm được biến đổi, bằng cách loại bỏ một hoặc nhiều tín hiệu đối tượng âm thanh gốc từ ba hoặc nhiều hơn ba tín hiệu trộn giảm gốc.

14. Bộ tạo tín hiệu dư theo điểm 13,

trong đó bộ phận biến đổi trộn giảm (250) được làm thích ứng để áp dụng công thức:

$$\tilde{\mathbf{X}}_{nonEAO} = \mathbf{X} - \mathbf{D}\mathbf{Z}_{eao}^* \mathbf{S}_{eao}.$$

để loại bỏ một hoặc nhiều tín hiệu trong số nhiều tín hiệu đối tượng âm thanh gốc từ ba hoặc nhiều hơn ba tín hiệu trộn giảm để thu được ba hoặc nhiều hơn ba tín hiệu trộn giảm được biến đổi,

trong đó:

$\mathbf{X}$ biểu thị ba hoặc nhiều hơn ba tín hiệu trộn giảm trước khi được biến đổi

$\tilde{\mathbf{X}}_{nonEAO}$ biểu thị ba hoặc nhiều hơn ba tín hiệu trộn giảm được biến đổi

$\mathbf{D}$ biểu thị thông tin trộn giảm

$\mathbf{S}_{eao}$ bao gồm một hoặc nhiều tín hiệu trong số nhiều tín hiệu đối tượng âm thanh gốc đã nêu, và

$\mathbf{Z}_{eao}^*$ biểu thị các vị trí của một hoặc nhiều tín hiệu trong số nhiều tín hiệu đối tượng âm thanh gốc đã nêu.

15. Bộ tạo tín hiệu dư (200) theo điểm 12, trong đó bộ phận biến đổi trộn giảm (250) được tạo cấu hình để biến đổi ba hoặc nhiều hơn ba tín hiệu trộn giảm gốc để thu được ba hoặc nhiều hơn ba tín hiệu trộn giảm được biến đổi bằng cách tạo ra một hoặc nhiều tín hiệu đối tượng âm thanh được biến đổi dựa trên một hoặc nhiều tín hiệu đối tượng âm thanh được ước tính và dựa trên một hoặc nhiều tín hiệu dư, và bằng cách loại bỏ một hoặc nhiều tín hiệu đối tượng âm thanh được biến đổi từ ba hoặc nhiều hơn ba tín hiệu trộn giảm gốc.

16. Bộ tạo tín hiệu dư theo điểm 15,

trong đó bộ phận biến đổi trộn giảm (250) được làm thích ứng để áp dụng công thức:

$$\tilde{\mathbf{X}}_{nonEAO} = \mathbf{X} - \mathbf{D}\mathbf{Z}_{eao}^* \mathbf{S}_{eao}.$$

để loại bỏ một hoặc nhiều tín hiệu đối tượng âm thanh được biến đổi từ ba hoặc nhiều hơn ba tín hiệu trộn giảm để thu được ba hoặc nhiều hơn ba tín hiệu trộn giảm được biến đổi,

trong đó:

$\mathbf{X}$ biểu thị ba hoặc nhiều hơn ba tín hiệu trộn giảm trước khi được biến đổi

$\tilde{\mathbf{X}}_{nonEAO}$ biểu thị ba hoặc nhiều hơn ba tín hiệu trộn giảm được biến đổi

$\mathbf{D}$ biểu thị thông tin trộn giảm

$\mathbf{S}_{eao}$ bao gồm một hoặc nhiều tín hiệu đối tượng âm thanh được biến đổi đã nêu, và

$\mathbf{Z}_{eao}^*$ biểu thị các vị trí của một hoặc nhiều tín hiệu đối tượng âm thanh được biến đổi đã nêu.

17. Bộ tạo tín hiệu dư (200) theo một điểm trong số các điểm từ 12 đến 16,

trong đó, bộ tạo tín hiệu dư (200) được làm thích ứng để điều khiển hai hoặc nhiều hơn hai bước lặp lại,

trong đó, với mỗi bước lặp lại, bộ phận giải mã theo tham số (230) được làm thích ứng để xác định chính xác một tín hiệu đối tượng âm thanh trong số nhiều tín hiệu đối tượng âm thanh được ước tính,

trong đó với bước lặp lại đã nêu, bộ phận ước tính phần dư (240) được làm thích ứng để xác định chính xác một tín hiệu dư trong số nhiều tín hiệu dư bằng cách biến đổi tín hiệu đối tượng âm thanh đã nêu trong số nhiều tín hiệu đối tượng âm thanh được ước tính,

trong đó, với bước lặp lại đã nêu, bộ phận biến đổi trộn giảm (250) được làm thích ứng để biến đổi ba hoặc nhiều hơn ba tín hiệu trộn giảm, và

trong đó, với bước lặp lại tiếp theo sau bước lặp lại đã nêu, bộ phận giải mã theo tham số (230) được làm thích ứng để xác định chính xác một tín hiệu đối tượng âm thanh trong số nhiều tín hiệu đối tượng âm thanh được ước tính dựa trên ba hoặc nhiều hơn ba tín hiệu trộn giảm mà đã được biến đổi.

18. Bộ tạo tín hiệu dư (200) theo một điểm bất kỳ trong các điểm từ 11 đến 17, trong đó bộ phận ước tính phần dư (240) được làm thích ứng để tạo ra ít nhất năm tín hiệu dư dựa trên ít nhất năm tín hiệu đối tượng âm thanh gốc trong số nhiều tín hiệu đối tượng âm thanh gốc và dựa trên ít nhất năm tín hiệu đối tượng âm thanh được ước tính trong số nhiều tín hiệu đối tượng âm thanh được ước tính.

19. Bộ mã hóa để mã hóa nhiều tín hiệu đối tượng âm thanh gốc bằng cách tạo ra ba hoặc nhiều hơn ba tín hiệu trộn giảm, bằng cách tạo ra thông tin phụ theo tham số và bằng cách tạo ra nhiều tín hiệu dư, trong đó bộ mã hóa bao gồm:

bộ tạo trộn giảm (210) để cung cấp ba hoặc nhiều hơn ba tín hiệu trộn giảm biểu thị sự trộn giảm trong số nhiều tín hiệu đối tượng âm thanh gốc,

bộ ước tính thông tin phụ theo tham số (220) để tạo ra thông tin phụ theo tham số biểu thị thông tin trên nhiều tín hiệu đối tượng âm thanh gốc, để thu được thông tin phụ theo tham số, và

bộ tạo tín hiệu dư (200) theo một điểm bất kỳ trong số các điểm từ 11 đến 18,

trong đó bộ phận giải mã theo tham số (230) của bộ tạo tín hiệu dư (200) được làm thích ứng để tạo ra nhiều tín hiệu đối tượng được ước tính bằng cách trộn tăng ba hoặc nhiều hơn ba tín hiệu trộn giảm được cung cấp bởi bộ tạo trộn giảm (210), trong đó các tín hiệu trộn giảm mã hóa nhiều tín hiệu đối tượng âm thanh gốc, trong đó bộ phận giải mã theo tham số (230) được tạo cấu hình để trộn tăng ba hoặc nhiều hơn ba tín hiệu trộn giảm phụ thuộc vào thông tin phụ theo tham số được tạo ra bởi bộ ước tính thông tin phụ theo tham số (220) và

trong đó bộ phận ước tính phần dư (240) của bộ tạo tín hiệu dư (200) được làm thích ứng để tạo ra nhiều tín hiệu dư dựa trên nhiều tín hiệu đối tượng âm thanh gốc và dựa trên nhiều tín hiệu đối tượng âm thanh được ước tính, sao cho mỗi một tín hiệu trong số nhiều tín hiệu dư biểu thị sự khác biệt giữa một tín hiệu trong số nhiều tín hiệu đối tượng âm thanh gốc và một tín hiệu trong số nhiều tín hiệu đối tượng âm thanh được ước tính.

20. Bộ mã hóa theo điểm 19, trong đó bộ mã hóa là bộ mã hóa SAOC.

21. Hệ thống mã hóa đối tượng âm thanh theo tham số sử dụng khái niệm phần dư, bao gồm:

bộ mã hóa (310) theo điểm 19 hoặc điểm 20 để mã hóa nhiều tín hiệu đối tượng âm thanh gốc bằng cách tạo ra ba hoặc nhiều hơn ba tín hiệu trộn giảm, bằng cách tạo ra thông tin phụ theo tham số và bằng cách tạo ra nhiều tín hiệu dư, và

bộ giải mã (320) theo một điểm bất kỳ trong số các điểm từ 1 đến 10, trong đó bộ giải mã (320) được tạo cấu hình để tạo ra nhiều tín hiệu đối tượng âm thanh được ước tính thứ hai dựa trên ba hoặc nhiều hơn ba tín hiệu trộn giảm được tạo ra bởi bộ mã hóa (310), dựa trên thông tin phụ theo tham số được tạo ra bởi bộ mã hóa (310) và dựa trên nhiều tín hiệu dư được tạo ra bởi bộ mã hóa (310).

22. Bộ mã hóa để mã hóa nhiều tín hiệu đối tượng âm thanh gốc theo điểm 19, trong đó tín hiệu âm thanh được mã hóa bao gồm: ba hoặc nhiều hơn ba tín hiệu trộn giảm (410), thông tin phụ theo tham số (420) và nhiều tín hiệu dư (430),

trong đó ba hoặc nhiều hơn ba tín hiệu trộn giảm (410) là sự trộn giảm của nhiều tín hiệu đối tượng âm thanh gốc,

trong đó thông tin phụ theo tham số (420) bao gồm các tham số biểu thị thông tin phụ trên nhiều tín hiệu đối tượng âm thanh gốc,

trong đó mỗi tín hiệu trong số nhiều tín hiệu dư (430) là các tín hiệu khác nhau biểu thị sự khác biệt giữa một tín hiệu trong số nhiều tín hiệu âm thanh gốc và một tín hiệu trong số nhiều tín hiệu đối tượng âm thanh được ước tính.

23. Phương pháp mã hóa đối tượng âm thanh theo tham số sử dụng khái niệm phân dư, phương pháp bao gồm các bước:

tạo ra nhiều tín hiệu đối tượng âm thanh được ước tính thứ nhất bằng cách trộn tăng ba hoặc nhiều hơn ba tín hiệu trộn giảm, trong đó ba hoặc nhiều hơn ba tín hiệu trộn giảm mã hóa nhiều tín hiệu đối tượng âm thanh gốc, trong đó tạo ra nhiều tín hiệu đối tượng âm thanh được ước tính thứ nhất bao gồm trộn tăng ba hoặc nhiều hơn các tín hiệu trộn giảm phụ thuộc vào thông tin phụ theo tham số biểu thị thông tin trên nhiều tín hiệu đối tượng âm thanh gốc, và

tạo ra nhiều tín hiệu đối tượng âm thanh được ước tính thứ hai bằng cách biến đổi một hoặc nhiều tín hiệu đối tượng âm thanh được ước tính thứ nhất, trong đó tạo ra nhiều tín hiệu đối tượng âm thanh được ước tính thứ hai bao gồm biến đổi một hoặc nhiều tín hiệu đối tượng âm thanh được ước tính thứ nhất đã nêu phụ thuộc vào một hoặc nhiều tín hiệu dư.

24. Phương pháp mã hóa đối tượng âm thanh theo tham số sử dụng khái niệm phân dư, phương pháp bao gồm các bước:

tạo ra nhiều tín hiệu đối tượng âm thanh được ước tính bằng cách trộn tăng ba hoặc nhiều hơn ba tín hiệu trộn giảm, trong đó ba hoặc nhiều hơn ba tín hiệu trộn giảm

mã hóa nhiều tín hiệu đối tượng âm thanh gốc, trong đó tạo ra nhiều tín hiệu đối tượng âm thanh được ước tính bao gồm trộn tăng ba hoặc nhiều hơn ba tín hiệu trộn giảm phụ thuộc vào thông tin phụ theo tham số biểu thị thông tin trên nhiều tín hiệu đối tượng âm thanh gốc, và

tạo ra nhiều tín hiệu dư dựa trên nhiều tín hiệu đối tượng âm thanh gốc và dựa trên nhiều tín hiệu đối tượng âm thanh được ước tính, sao cho mỗi tín hiệu trong số nhiều tín hiệu dư là tín hiệu khác nhau biểu thị sự khác biệt giữa một tín hiệu trong số nhiều tín hiệu đối tượng âm thanh gốc và một tín hiệu trong số nhiều tín hiệu đối tượng âm thanh được ước tính.

25. Vật ghi có thể đọc được bằng máy tính chứa chương trình máy tính để thực hiện phương pháp theo điểm 23 khi chạy trên máy tính hoặc bộ xử lý tín hiệu.

26. Vật ghi có thể đọc được bằng máy tính chứa chương trình máy tính để thực hiện phương pháp theo điểm 24 khi chạy trên máy tính hoặc bộ xử lý tín hiệu.

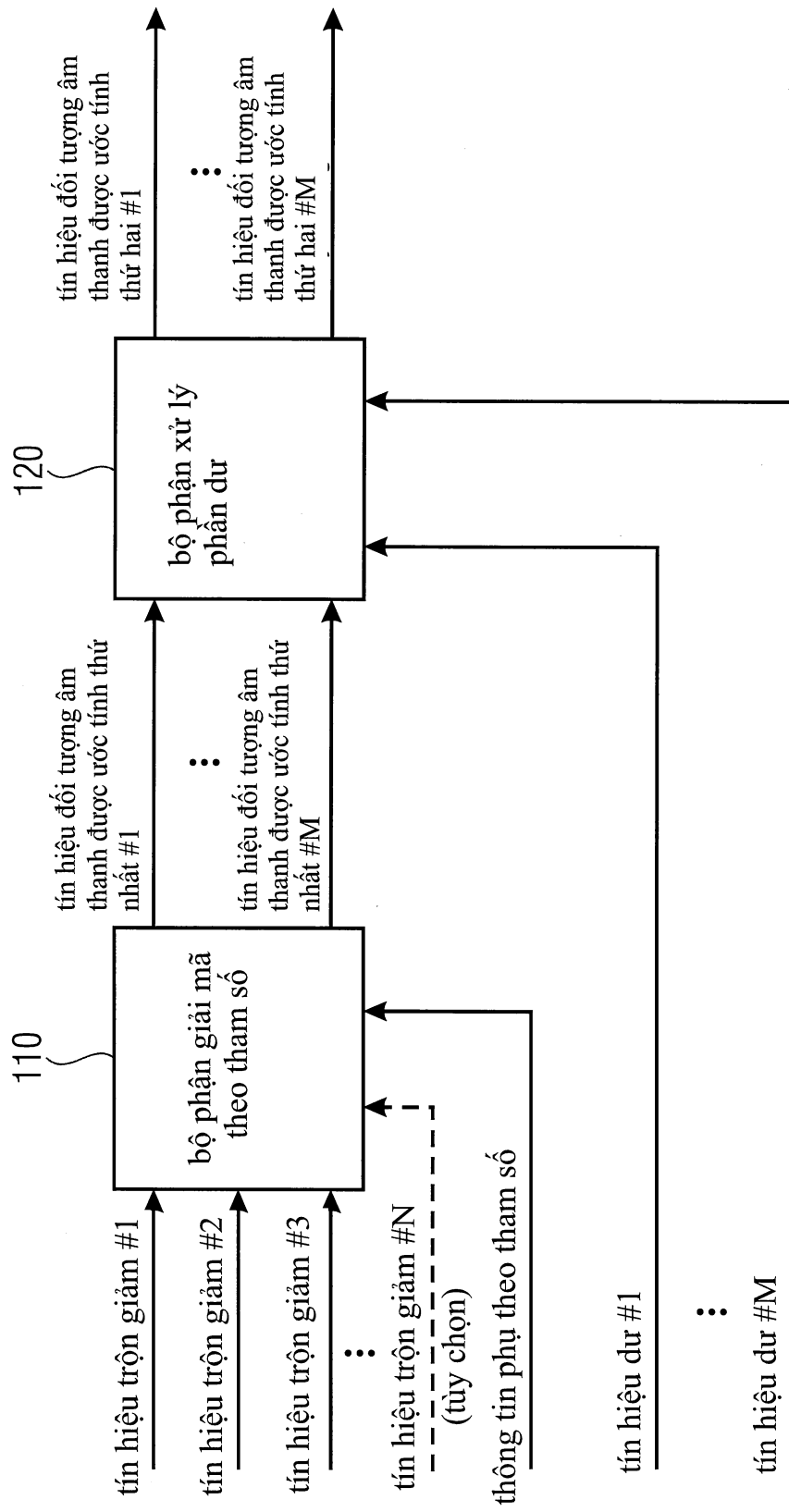


FIG 1A

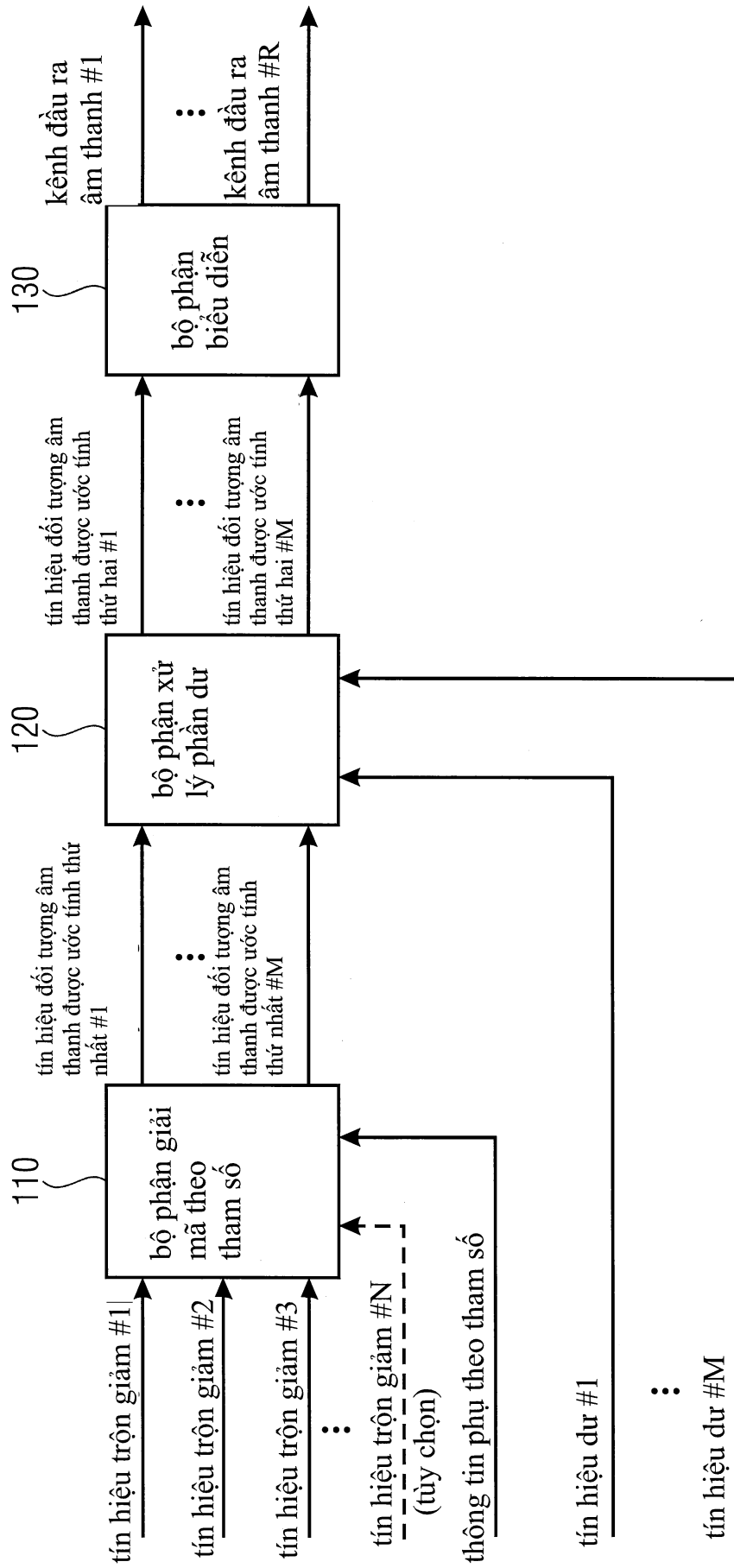


FIG 1B

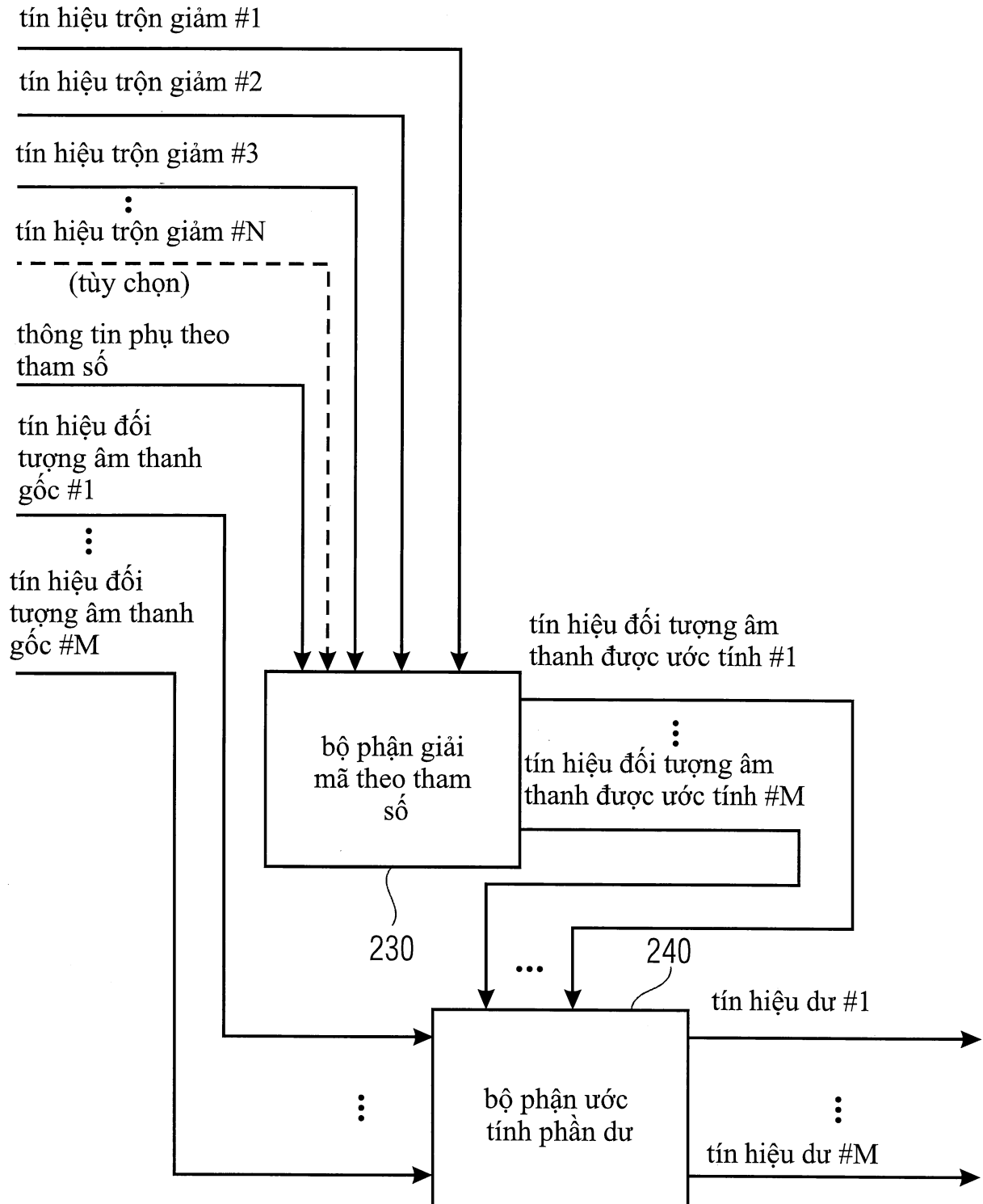


FIG 2A

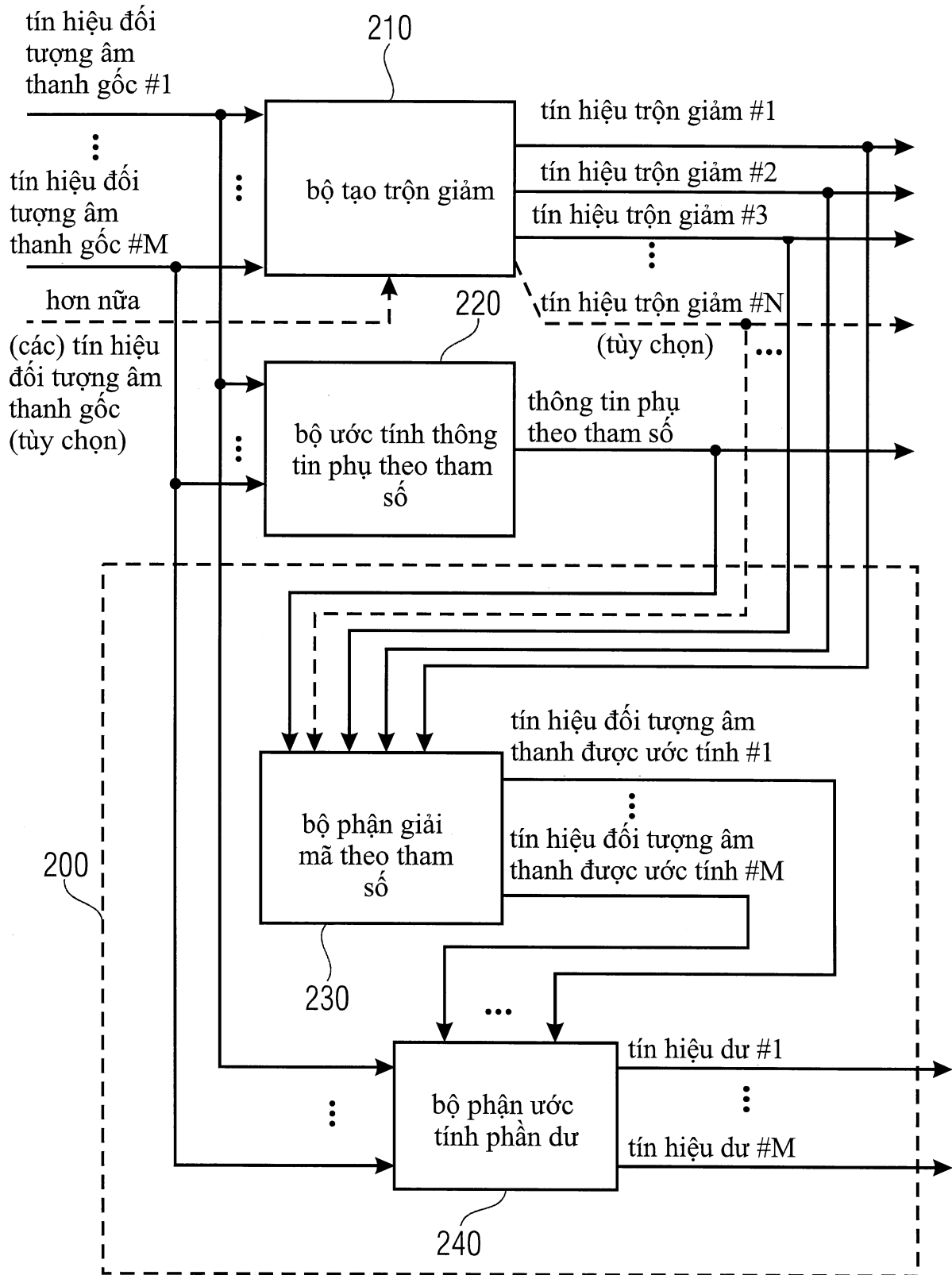


FIG 2B

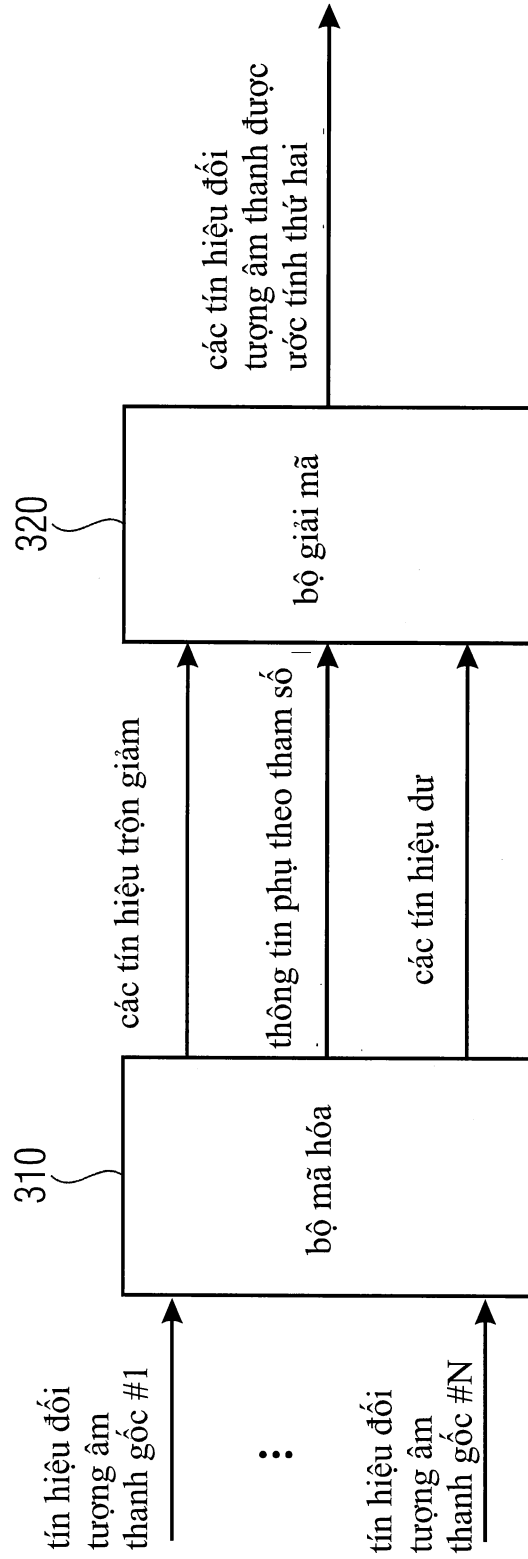


FIG 3

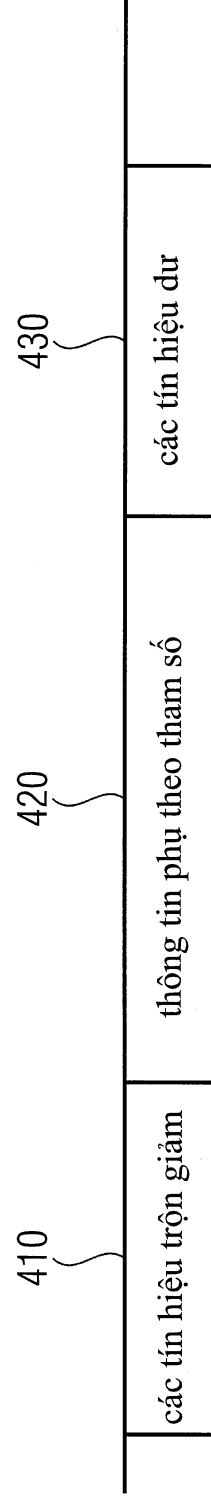


FIG 4

đầu ra đơn, lập thẻ hoặc kênh 5.1

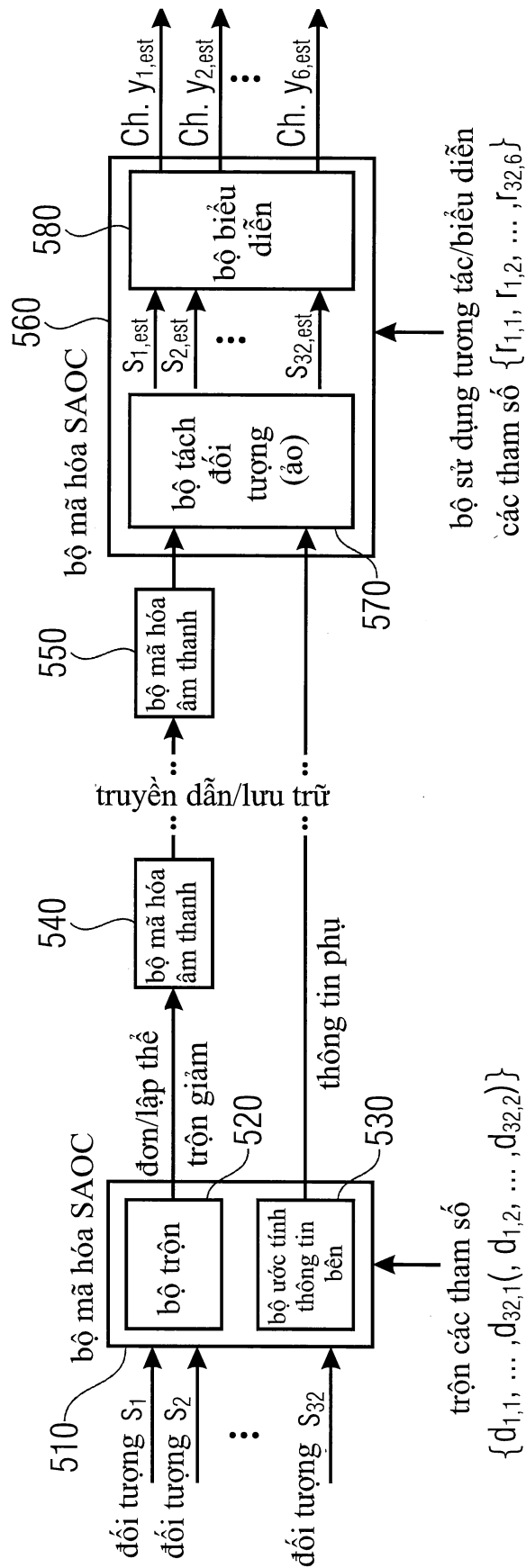


FIG 5

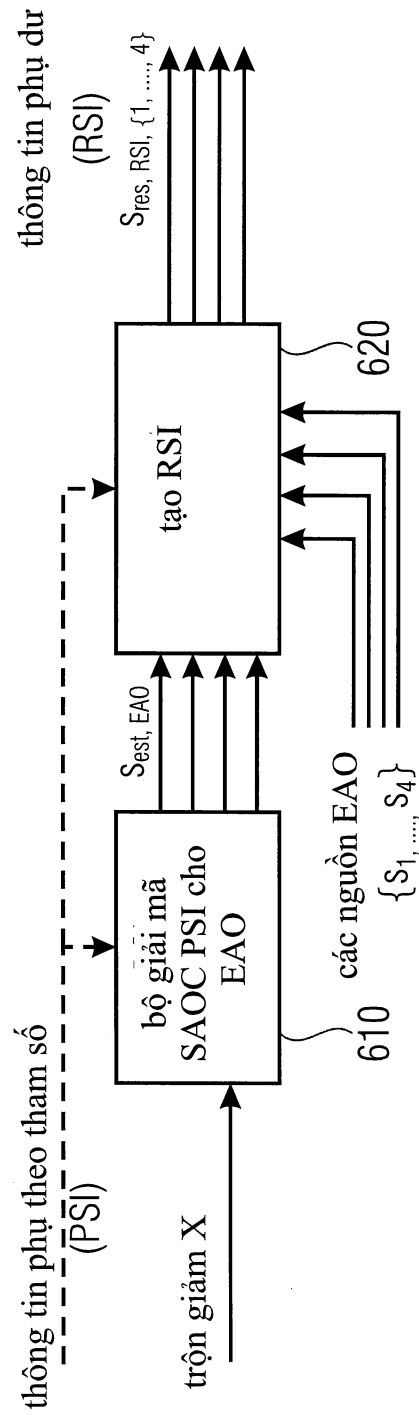


FIG 6

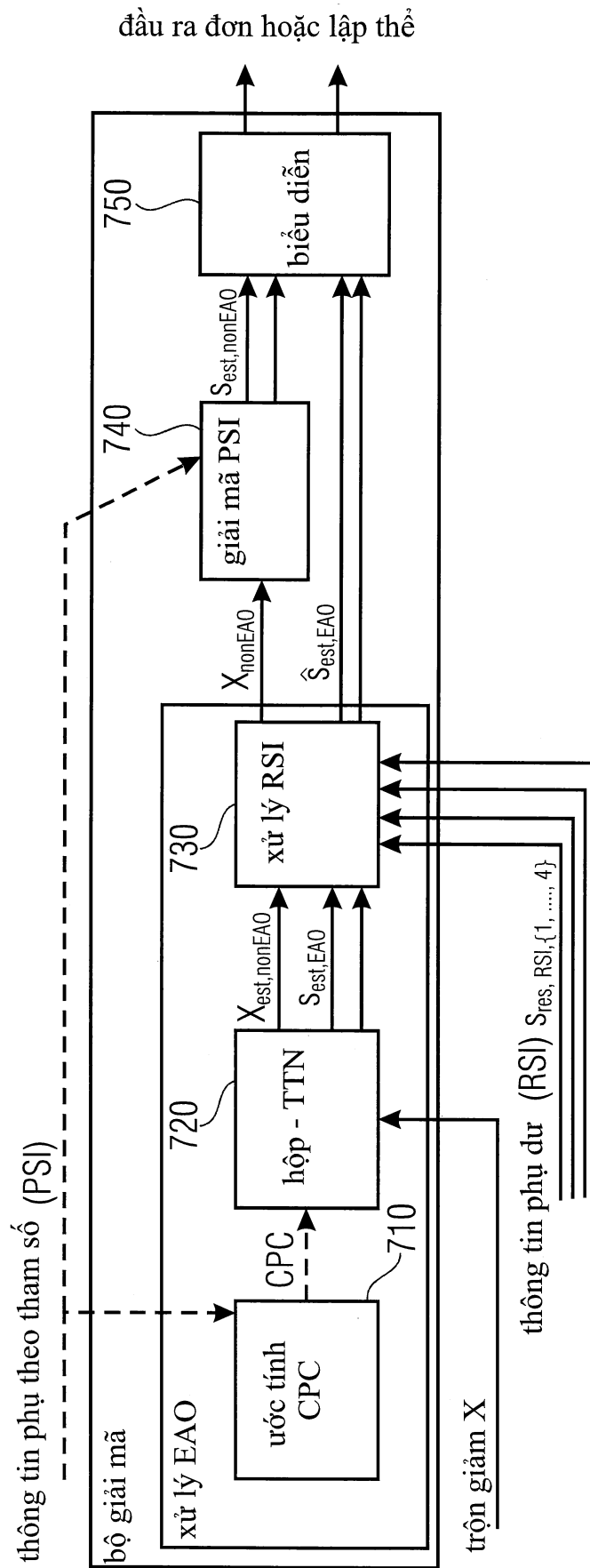


FIG 7

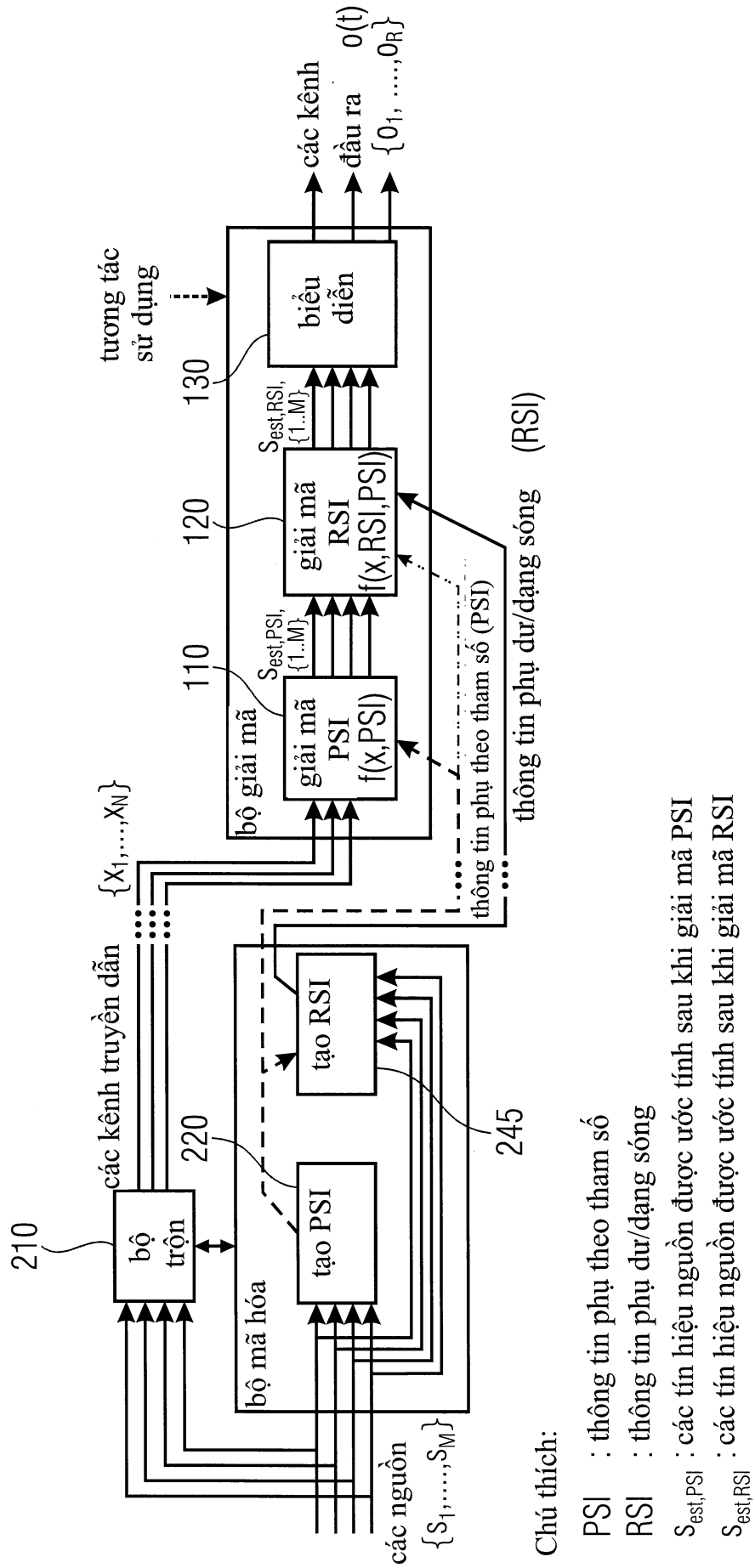


FIG 8

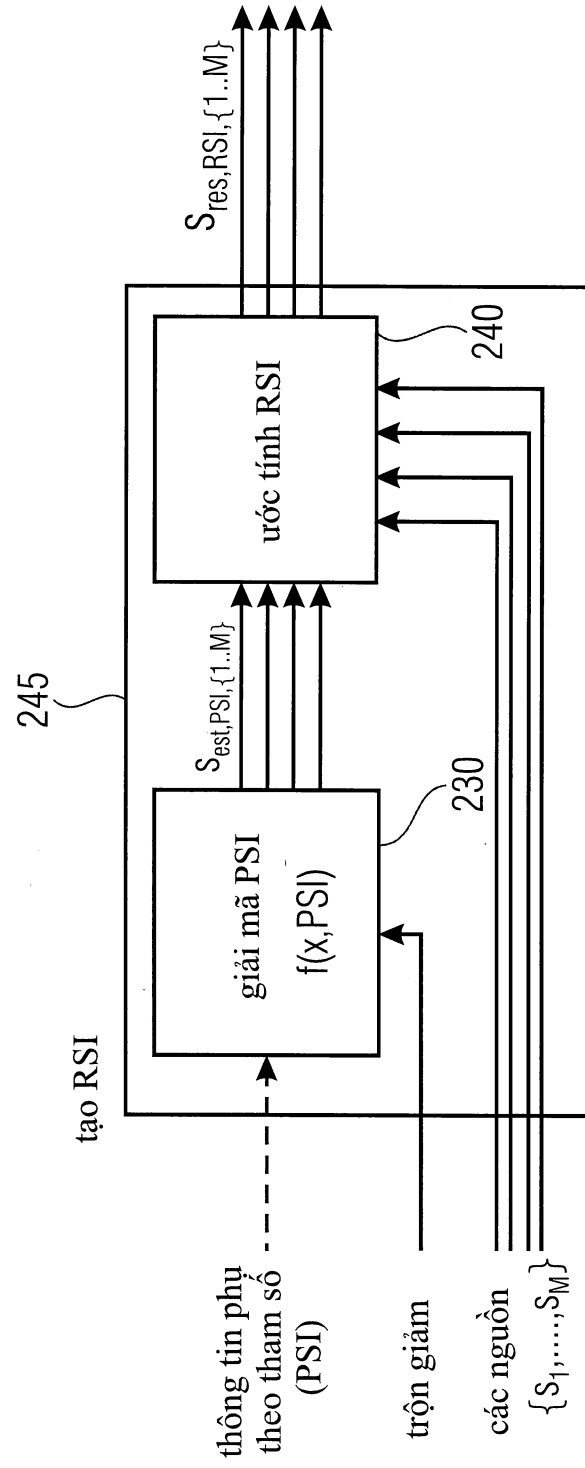


FIG 9

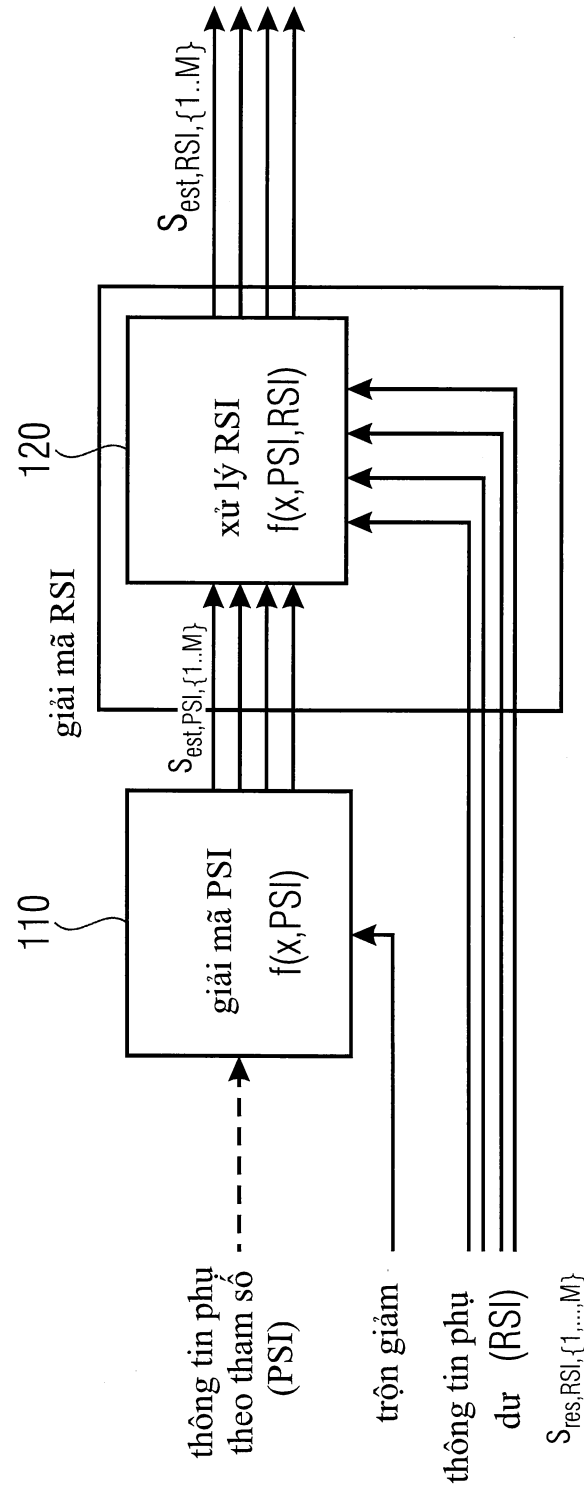


FIG 10

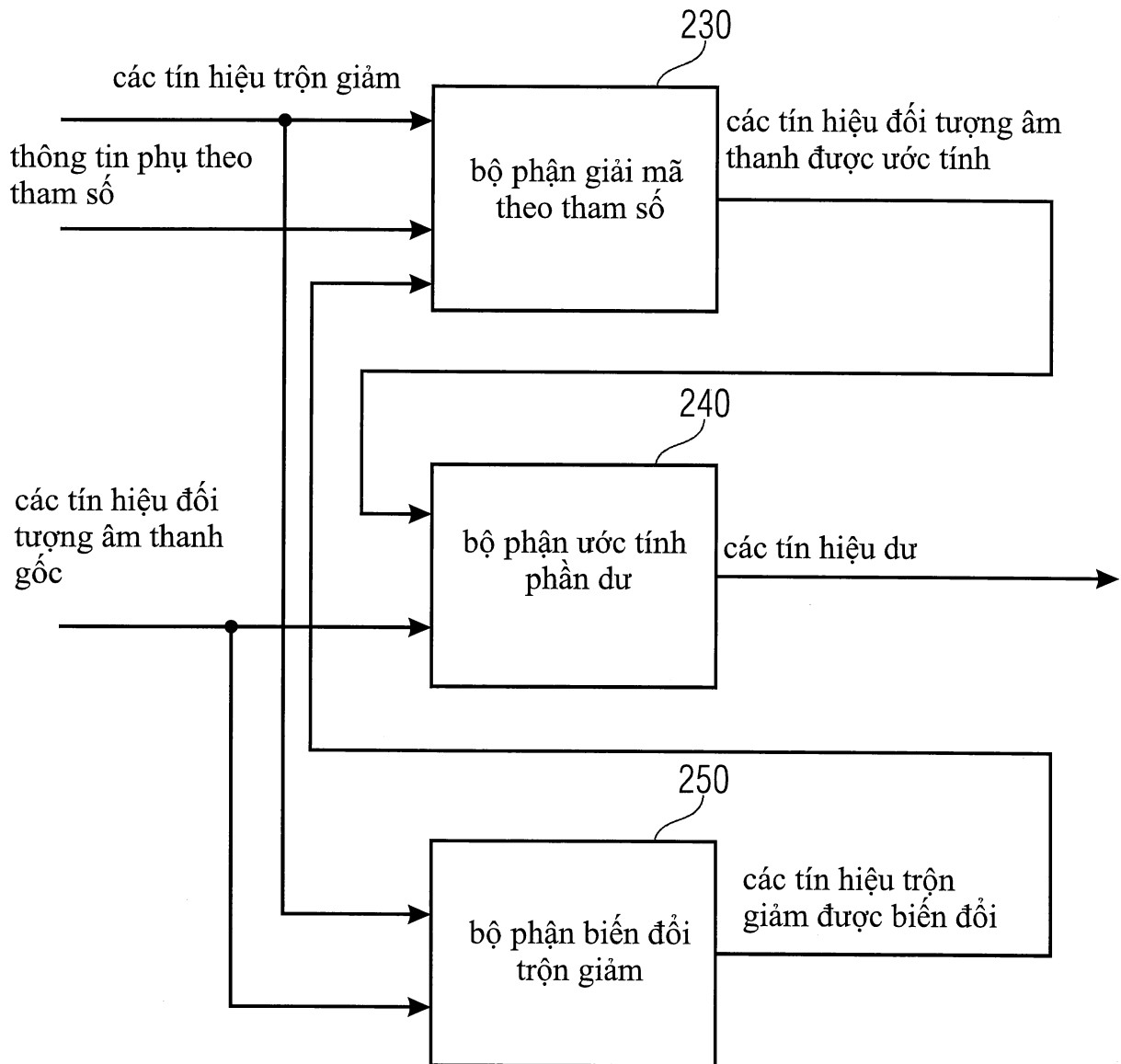


FIG 11

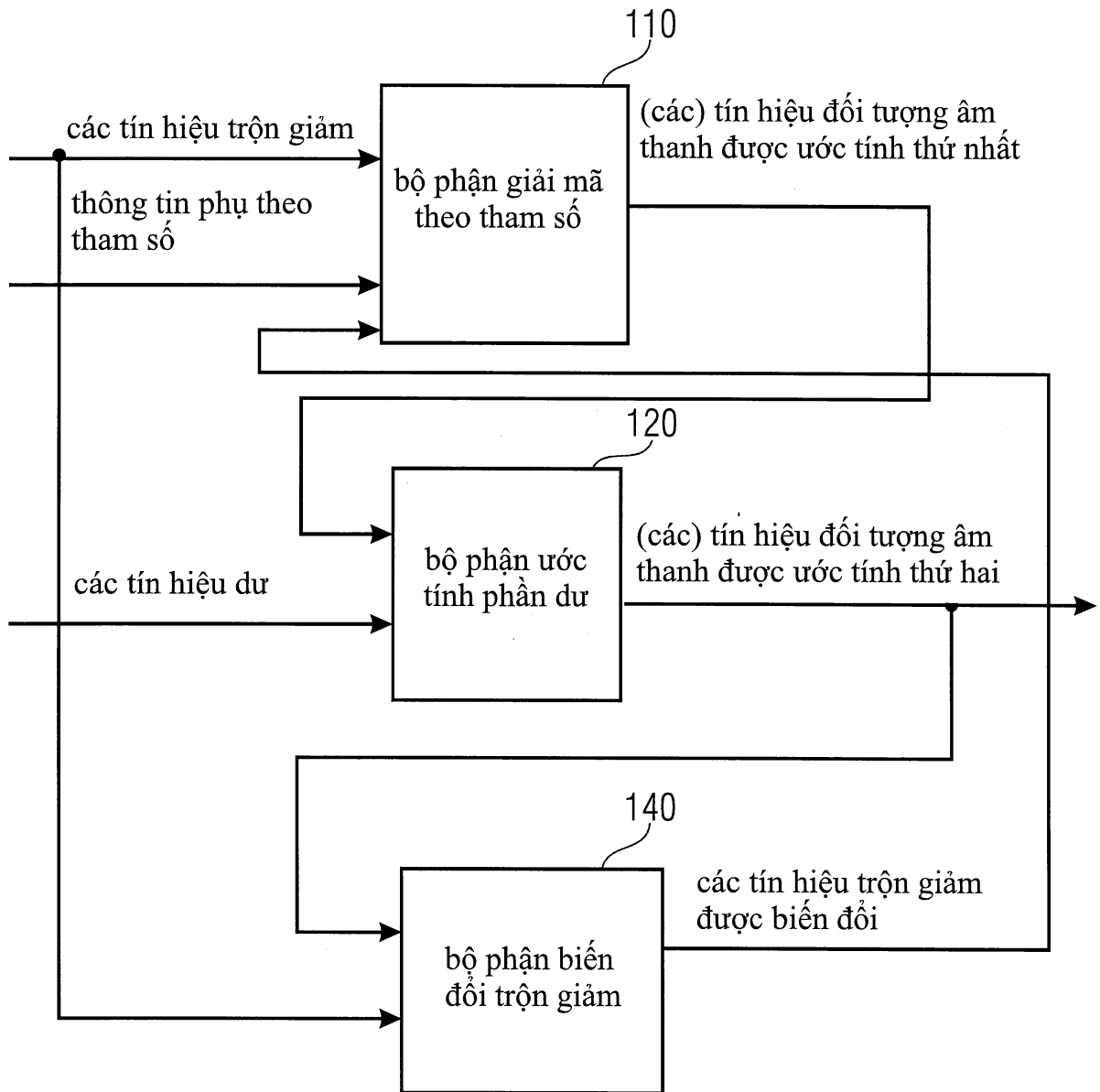


FIG 12

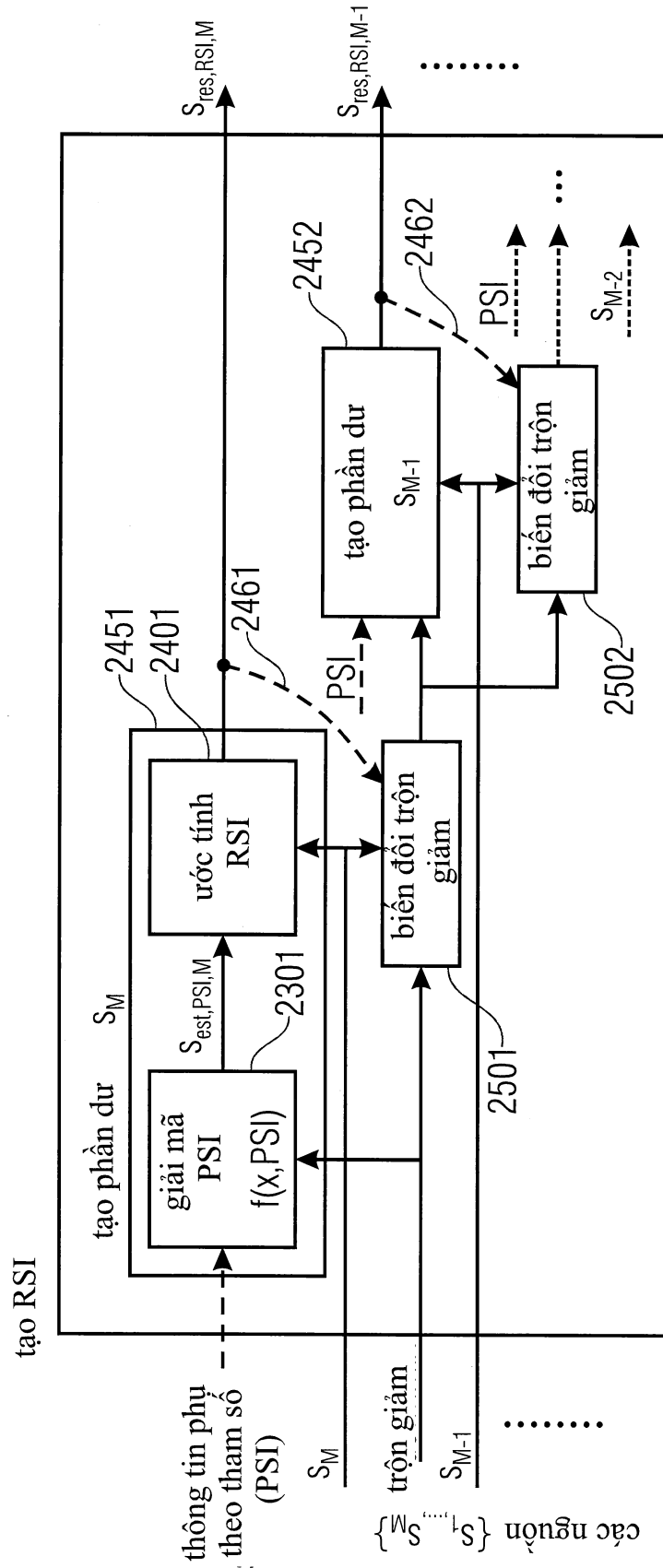


FIG 13

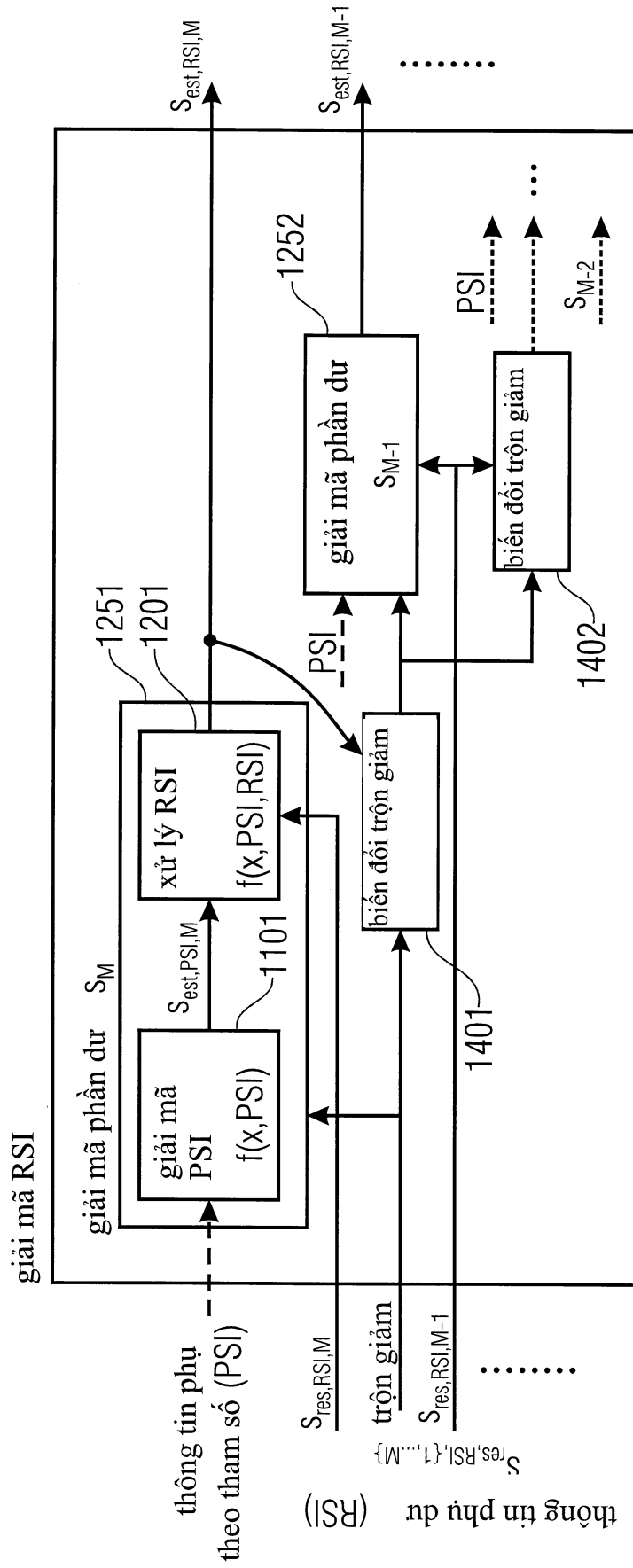


FIG 14

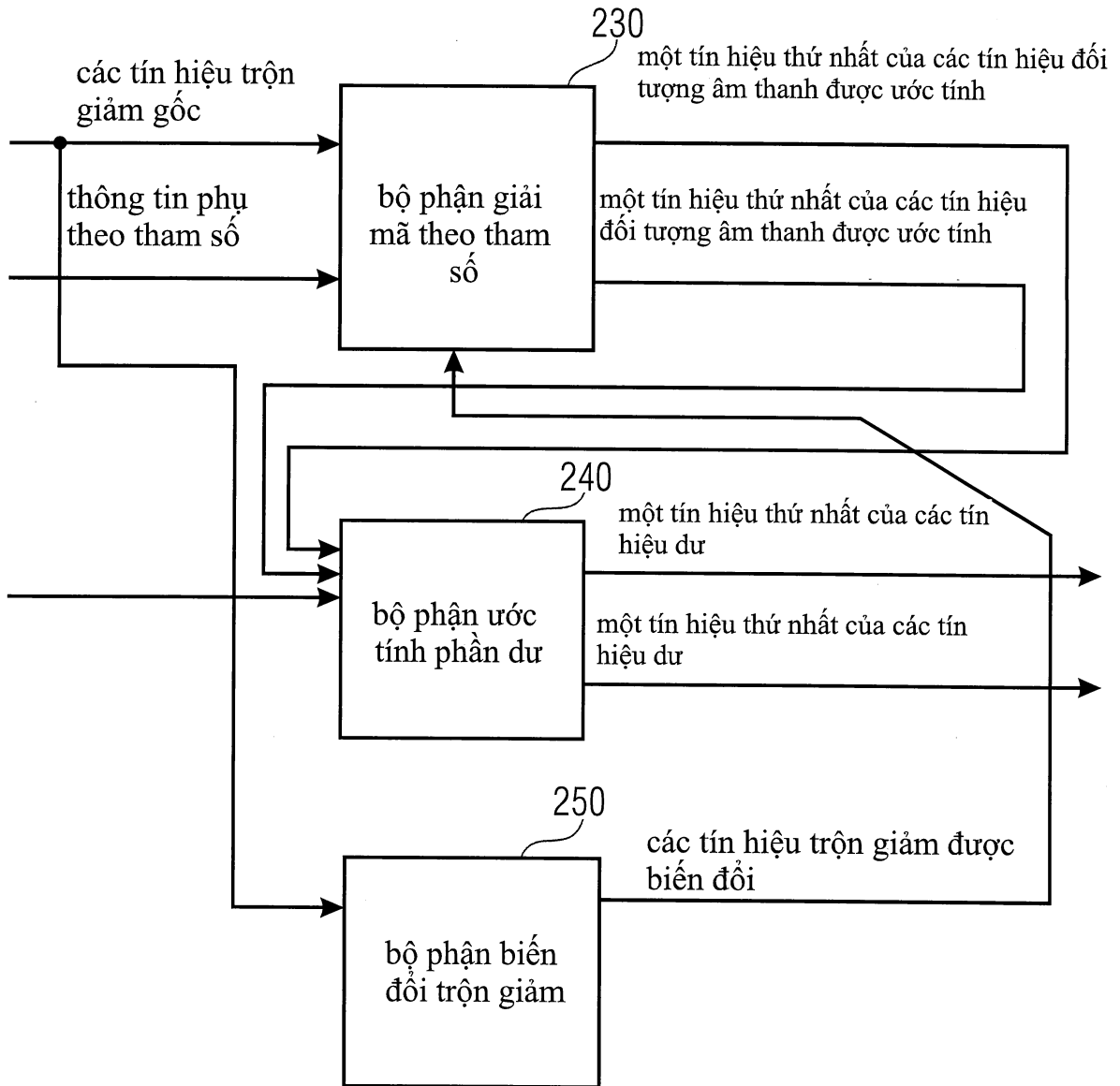


FIG 15

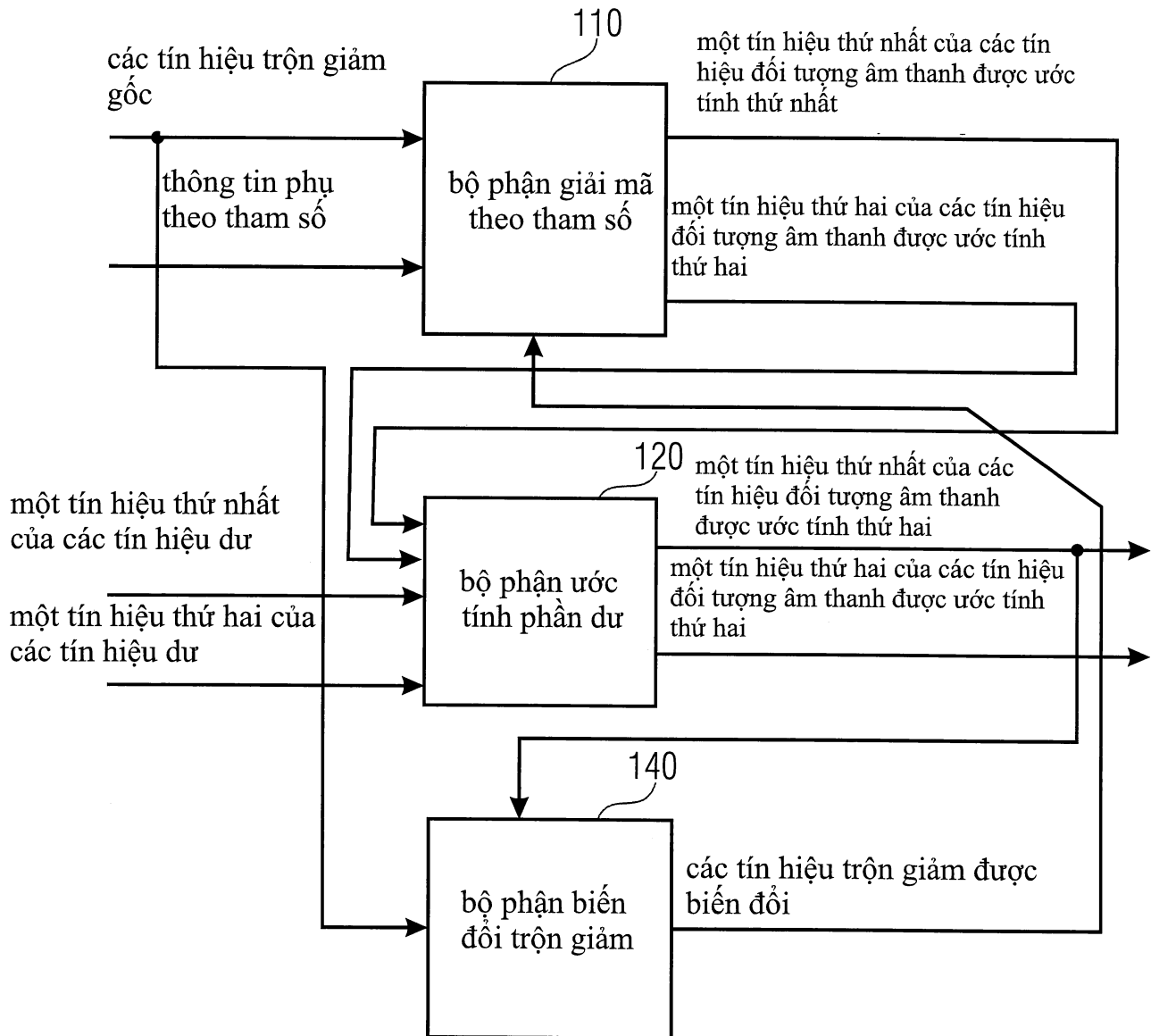


FIG 16