



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

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

(11)



1-0029618

(51)⁷ G10L 19/008; G10L 19/16

(13) B

(21) 1-2017-03274

(22) 01/02/2016

(86) PCT/EP2016/052037 01/02/2016

(87) WO/2016/124524 11/08/2016

(30) 15153486.4 02/02/2015 EP

(45) 25/10/2021 403

(43) 27/11/2017 356A

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

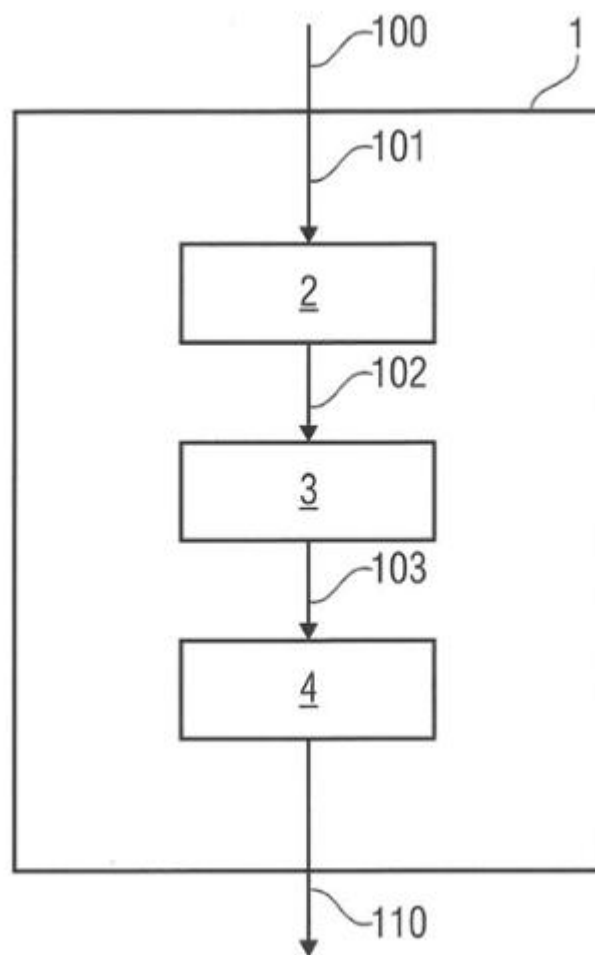
Hansastraße 27c, 80686 Muenchen, Germany

(72) MURTAZA, Adrian (RO); PAULUS, Jouni (FI); FUCHS, Harald (DE);
CAMILLERI, Roberta (MT); TERENTIV, Leon (DE); DISCH, Sascha (DE);
HERRE, Juergen (DE); HELLMUTH, Oliver (DE).

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

(54) THIẾT BỊ VÀ PHƯƠNG PHÁP XỬ LÝ TÍN HIỆU ÂM THANH ĐƯỢC MÃ HÓA

(57) Sáng chế đề cập đến thiết bị và phương pháp xử lý tín hiệu âm thanh được mã hóa. Thiết bị (1) để xử lý tín hiệu âm thanh được mã hóa (100) bao gồm nhiều tín hiệu trộn giảm (101) được kết hợp với nhiều đối tượng âm thanh đầu vào (111) và các tham số đối tượng (E). Thiết bị (1) bao gồm bộ tạo nhóm (2) được tạo cấu hình để tạo nhóm các tín hiệu trộn giảm (101) thành các nhóm các tín hiệu trộn giảm (102) được kết hợp với tập hợp các đối tượng âm thanh đầu vào (111). Thiết bị (1) bao gồm bộ xử lý (3) được tạo cấu hình để thực hiện ít nhất một bước xử lý riêng lẻ trên các tham số đối tượng (E_k) của mỗi tập hợp các đối tượng âm thanh đầu vào (111) để cung cấp các kết quả nhóm (103, 104). Hơn nữa, có bộ tổ hợp (4) được tạo cấu hình để tổ hợp các kết quả nhóm (103, 104) đã nêu hoặc các kết quả nhóm được xử lý để cung cấp tín hiệu âm thanh được giải mã (110). Bộ tạo nhóm (2) được tạo cấu hình để tạo nhóm các tín hiệu trộn giảm (101) để mỗi đối tượng âm thanh đầu vào (111) thuộc về chỉ một tập hợp các đối tượng âm thanh đầu vào (111). Sáng chế cũng đề cập đến phương pháp tương ứng.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến thiết bị và phương pháp xử lý tín hiệu âm thanh được mã hóa.

Tình trạng kỹ thuật của sáng chế

Gần đây, các kỹ thuật để truyền dẫn/lưu trữ hiệu quả tốc độ bit của các ngữ cảnh âm thanh chứa nhiều đối tượng âm thanh đã được đề xuất trong trường mã hóa âm thanh (xem các tài liệu tham khảo sau [BCC, JSC, SAOC, SAOC1, SAOC2]) và sự tách nguồn được thông báo (xem, ví dụ, các tài liệu tham khảo sau [ISS1, ISS2, ISS3, ISS4, ISS5, ISS6]).

Các kỹ thuật nhằm đến việc khôi phục ngữ cảnh âm thanh đầu ra hoặc các đối tượng nguồn âm thanh trên cơ sở thông tin phụ bổ sung mô tả các tín hiệu âm thanh và/hoặc các đối tượng nguồn được truyền dẫn/lưu trữ trong ngữ cảnh âm thanh. Sự khôi phục xảy ra trong bộ giải mã sử dụng sơ đồ tách nguồn được thông báo theo tham số.

Đáng tiếc, các tác giả sáng chế đã phát hiện rằng trong một số trường hợp, sơ đồ tách theo tham số có thể dẫn đến các thành phần lạ có thể nghe rất kém dẫn đến trải nghiệm nghe không hài lòng.

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.

[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, J. Pinel, R. Badeau, L. Girin, 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] S. Zhang and L. Girin: “An Informed Source Separation System for Speech Signals”, INTERSPEECH, 2011.

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

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

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

[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.

[SAOC3D] ISO/IEC, JTC1/SC29/WG11 N14747, Text of ISO/MPEG 23008-3/DIS 3D Audio, Sapporo, July 2014.

[SAOC3D2] J. Herre, J. Hilpert, A. Kuntz, and J. Plogsties, “MPEG-H Audio - The new standard for universal spatial / 3D audio coding,” 137th AES Convention, Los Angeles, 2011.

Bản chất kỹ thuật của sáng chế

Do đó, mục đích của sáng chế là để cải thiện chất lượng âm thanh của các tín hiệu âm thanh được giải mã sử dụng các kỹ thuật mã hóa theo tham số.

Mục đích này đạt được nhờ thiết bị theo điểm 1, và phương pháp tương ứng theo điểm 20.

Mục đích đạt được bằng thiết bị để xử lý tín hiệu âm thanh được mã hóa. Tín hiệu âm thanh được mã hóa bao gồm nhiều tín hiệu trộn giảm được kết hợp với nhiều đối tượng âm thanh đầu vào và các tham số đối tượng (**E**). Thiết bị bao gồm bộ tạo nhóm, bộ xử lý, và bộ tổ hợp.

Bộ tạo nhóm được tạo cấu hình để tạo nhóm nhiều tín hiệu trộn giảm vào nhiều nhóm tín hiệu trộn giảm. Mỗi nhóm gồm các tín hiệu trộn giảm được kết hợp với tập hợp các đối tượng âm thanh đầu vào (hoặc các tín hiệu âm thanh đầu vào) trong số nhiều đối tượng âm thanh đầu vào. Nói cách khác: các nhóm bao phủ các tập hợp con của tập hợp các tín hiệu âm thanh đầu vào được biểu diễn bởi tín hiệu âm thanh được mã hóa. Mỗi nhóm các tín hiệu trộn giảm cũng được kết hợp với một số tham số đối tượng **E** mô tả đối tượng âm thanh đầu vào. Sau đây, các nhóm riêng G_k được nhận biết với chỉ số k với $1 \leq k \leq K$ với K là số các nhóm tín hiệu trộn giảm.

Hơn nữa, bộ xử lý - theo sau nhóm - được tạo cấu hình để thực hiện ít nhất một bước xử lý riêng các tham số đối tượng của mỗi nhóm đối tượng âm thanh. Do đó, ít nhất một bước xử lý được thực hiện không đồng thời trên tất cả các tham số đối tượng nhưng riêng lẻ trên các tham số đối tượng thuộc về nhóm các tín hiệu trộn giảm tương ứng. Trong một phương án, chỉ một bước được thực hiện riêng lẻ. Trong phương án khác nhau, nhiều hơn một bước được thực hiện, trong đó trong phương án thay thế, toàn bộ việc xử lý được thực hiện riêng lẻ trên các nhóm về các tín hiệu trộn giảm. Bộ xử lý cung cấp các kết quả nhóm cho các nhóm riêng.

Trong phương án khác, bộ xử lý - theo sau nhóm - được tạo cấu hình để thực hiện ít nhất một bước xử lý riêng lẻ trên mỗi nhóm trong số nhiều nhóm tín hiệu trộn giảm. Do đó, ít nhất một bước xử lý được thực hiện không đồng thời trên tất cả các tín hiệu trộn giảm nhưng riêng lẻ trên các nhóm tín hiệu trộn giảm tương ứng.

Cuối cùng, bộ tổ hợp được tạo cấu hình để tổ hợp các kết quả nhóm hoặc các kết

quả nhóm được xử lý để cung cấp tín hiệu âm thanh được giải mã. Do đó, các kết quả nhóm hoặc các kết quả của các bước xử lý khác được thực hiện trên các kết quả nhóm được tổ hợp để cung cấp tín hiệu âm thanh được giải mã. Tín hiệu âm thanh được giải mã tương ứng với nhiều đối tượng âm thanh đầu vào mà được mã hóa bởi tín hiệu âm thanh được mã hóa.

Việc tạo nhóm được thực hiện bởi bộ tạo nhóm được thực hiện ít nhất dưới sự co hẹp mà mỗi đối tượng âm thanh đầu vào trong số nhiều đối tượng âm thanh đầu vào thuộc về chỉ hoặc chính xác một tập hợp các đối tượng âm thanh đầu vào. Điều này có nghĩa rằng mỗi đối tượng âm thanh đầu vào thuộc về chỉ một nhóm các tín hiệu trộn giảm. Điều này có nghĩa rằng mỗi tín hiệu trộn giảm thuộc về chỉ một nhóm các tín hiệu trộn giảm.

Theo phương án, bộ tạo nhóm được tạo cấu hình để tạo nhóm nhiều tín hiệu trộn giảm vào nhiều nhóm các tín hiệu trộn giảm để mỗi đối tượng âm thanh đầu vào trong số mỗi tập hợp các đối tượng âm thanh đầu vào hoặc không có mối quan hệ được tạo tín hiệu trong tín hiệu âm thanh được mã hóa với các đối tượng âm thanh đầu vào khác hoặc có mối quan hệ được tạo tín hiệu trong tín hiệu âm thanh được mã hóa chỉ với ít nhất một đối tượng âm thanh đầu vào thuộc về cùng một tập hợp các đối tượng âm thanh đầu vào. Điều này có nghĩa rằng không có đối tượng âm thanh đầu vào nào có mối quan hệ được tạo tín hiệu đến đối tượng âm thanh đầu vào thuộc về nhóm các tín hiệu trộn giảm khác. Mối quan hệ được tạo tín hiệu như vậy là trong một phương án mà hai đối tượng âm thanh đầu vào là các tín hiệu âm lập thể bắt nguồn từ một nguồn đơn.

Thiết bị theo sáng chế xử lý tín hiệu âm thanh được mã hóa bao gồm các tín hiệu trộn giảm. Trộn giảm là phần của quy trình mã hóa số lượng các tín hiệu âm thanh riêng lẻ đã cho và có nghĩa là số lượng nhất định các đối tượng âm thanh đầu vào được tổ hợp thành tín hiệu trộn giảm. Do đó, số lượng các đối tượng âm thanh đầu vào được giảm đến số lượng nhỏ các tín hiệu trộn giảm. Do đây là các tín hiệu trộn giảm được kết hợp với nhiều đối tượng âm thanh đầu vào.

Các tín hiệu trộn giảm được tạo nhóm thành các nhóm các tín hiệu trộn giảm và phải trải qua một cách riêng lẻ - tức là như các nhóm đơn - đến ít nhất một bước xử lý.

Do đó, thiết bị thực hiện ít nhất một bước xử lý được thực hiện không đồng thời trên tất cả các tín hiệu trộn giảm nhưng riêng lẻ trên các nhóm riêng lẻ các tín hiệu trộn giảm. Trong phương án khác, các tham số đối tượng của các nhóm được xử lý riêng biệt để thu được các ma trận sẽ được áp dụng cho tín hiệu âm thanh được mã hóa.

Trong một phương án là thiết bị bộ giải mã của các tín hiệu âm thanh được mã hóa. Thiết bị là phương án thay thế phần của bộ giải mã.

Trong một phương án, mỗi tín hiệu trộn giảm được quy cho một nhóm các tín hiệu trộn giảm và do đó, được xử lý riêng lẻ so với ít nhất một bước xử lý. Trong phương án này, số lượng các nhóm các tín hiệu trộn giảm bằng với số lượng các tín hiệu trộn giảm. Điều này có nghĩa rằng việc tạo nhóm và xử lý riêng lẻ trùng nhau.

Trong một phương án, sự tổ hợp là một trong số các bước cuối cùng của sự xử lý tín hiệu âm thanh được mã hóa. Trong phương án khác, các kết quả nhóm còn phải trải qua các bước xử lý khác nhau mà được thực hiện riêng lẻ hoặc đồng thời trên các kết quả nhóm.

Việc tạo nhóm (hoặc sự phát hiện các nhóm) và việc xử lý riêng lẻ các nhóm đã được thể hiện để dẫn đến sự cải thiện chất lượng âm thanh. Điều này đặc biệt duy trì, ví dụ, các kỹ thuật mã hóa theo tham số.

Theo phương án, bộ tạo nhóm của thiết bị được tạo cấu hình để tạo nhóm nhiều tín hiệu trộn giảm vào nhiều nhóm các tín hiệu trộn giảm trong khi tối thiểu hóa số lượng các tín hiệu trộn giảm nằm trong mỗi nhóm các tín hiệu trộn giảm. Trong phương án này, thiết bị cố gắng giảm số lượng các tín hiệu trộn giảm thuộc về mỗi nhóm. Trong một trường hợp, đến ít nhất một nhóm các tín hiệu trộn giảm thuộc về chỉ một tín hiệu trộn giảm.

Theo phương án, bộ tạo nhóm được tạo cấu hình để tạo nhóm nhiều tín hiệu trộn giảm thành nhiều nhóm các tín hiệu trộn giảm để chỉ một tín hiệu trộn giảm đơn thuộc về một nhóm các tín hiệu trộn giảm. Nói cách khác, việc tạo nhóm dẫn đến nhiều nhóm các tín hiệu trộn giảm trong đó ít nhất một nhóm tín hiệu trộn giảm được cho chỉ một tín hiệu trộn giảm thuộc về. Do đó, ít nhất một nhóm các tín hiệu trộn giảm đề cập đến chỉ một tín hiệu trộn giảm đơn. Trong phương án khác, số lượng các nhóm các tín

hiệu trộn giảm mà chỉ một tín hiệu trộn giảm thuộc về được tối đa hóa.

Trong một phương án, bộ tạo nhóm của thiết bị được tạo cấu hình để tạo nhóm nhiều tín hiệu trộn giảm vào nhiều nhóm các tín hiệu trộn giảm dựa trên thông tin nằm trong tín hiệu âm thanh được mã hóa. Trong phương án khác, thiết bị sử dụng chỉ thông tin nằm trong tín hiệu âm thanh được mã hóa để tạo nhóm các tín hiệu trộn giảm. Sử dụng thông tin nằm trong dòng bit của tín hiệu âm thanh được mã hóa bao gồm - trong một phương án – xem xét thông tin tương quan hoặc hiệp phương sai. Đặc biệt, bộ tạo nhóm, trích thông tin về mối quan hệ giữa các đối tượng âm thanh đầu vào khác nhau từ tín hiệu âm thanh được mã hóa.

Trong một phương án, bộ tạo nhóm được tạo cấu hình để tạo nhóm nhiều tín hiệu trộn giảm đã nêu thành nhiều nhóm tín hiệu trộn giảm đã nêu trên cơ sở trị số bsRelatedTo-values nằm trong tín hiệu âm thanh được mã hóa đã nêu. Liên quan đến các trị số này, ví dụ, đề cập đến WO 2011/039195 A1.

Theo phương án, bộ tạo nhóm được tạo cấu hình để tạo nhóm nhiều tín hiệu trộn giảm vào nhiều nhóm các tín hiệu trộn giảm bằng cách áp dụng ít nhất các bước sau (cho mỗi nhóm tín hiệu trộn giảm):

- phát hiện liệu tín hiệu trộn giảm được gán cho nhóm các tín hiệu trộn giảm hiện có hay không;

- phát hiện liệu ít nhất một đối tượng âm thanh đầu vào trong số nhiều đối tượng âm thanh đầu vào được kết hợp với tín hiệu trộn giảm là phần của tập hợp các đối tượng âm thanh đầu vào được kết hợp với nhóm các tín hiệu trộn giảm hiện có;

- gán tín hiệu trộn giảm vào nhóm các tín hiệu trộn giảm mới

trong trường hợp tín hiệu trộn giảm không bị gán vào nhóm các tín hiệu trộn giảm hiện có (do đó, tín hiệu trộn giảm chưa được gán vào nhóm) và

trong trường hợp tất cả các đối tượng âm thanh đầu vào trong số nhiều đối tượng âm thanh đầu vào được kết hợp với tín hiệu trộn giảm không có sự kết hợp với nhóm các tín hiệu trộn giảm hiện có (do đó, các đối tượng âm thanh đầu vào của tín hiệu trộn giảm - thông qua tín hiệu trộn giảm khác - chưa được gán vào nhóm; và

- tổ hợp tín hiệu trộn giảm với nhóm các tín hiệu trộn giảm hiện có

hoặc trong trường hợp tín hiệu trộn giảm được gán cho nhóm các tín hiệu trộn giảm hiện có hoặc trong trường hợp ít nhất một đối tượng âm thanh đầu vào trong số nhiều đối tượng âm thanh đầu vào được kết hợp với tín hiệu trộn giảm được kết hợp với nhóm các tín hiệu trộn giảm hiện có.

Nếu mỗi quan hệ được tạo tín hiệu trong tín hiệu âm thanh được mã hóa cũng được xem xét, sau đó bước phát hiện khác sẽ được thêm vào dẫn đến yêu cầu thêm cho việc gán và tổ hợp các tín hiệu trộn giảm.

Theo phương án, bộ xử lý được tạo cấu hình để thực hiện nhiều bước xử lý riêng lẻ trên các tham số đối tượng (E_k) của mỗi tập hợp các đối tượng âm thanh đầu vào (hoặc của mỗi nhóm các tín hiệu trộn giảm) để cung cấp các ma trận riêng lẻ như các kết quả nhóm. Bộ tổ hợp được tạo cấu hình để tổ hợp các ma trận riêng lẻ để cung cấp tín hiệu âm thanh được giải mã đã nêu. Các tham số đối tượng (E_k) thuộc về các đối tượng âm thanh đầu vào của nhóm các tín hiệu trộn giảm tương ứng với chỉ số k và được xử lý để thu được các ma trận riêng lẻ cho nhóm có chỉ số k này.

Theo phương án khác, bộ xử lý được tạo cấu hình để thực hiện nhiều bước xử lý riêng lẻ trên mỗi nhóm trong số nhiều nhóm các tín hiệu trộn giảm đã nêu để cung cấp các tín hiệu âm thanh đầu ra như các kết quả nhóm. Bộ tổ hợp được tạo cấu hình để tổ hợp các tín hiệu âm thanh đầu ra để cung cấp tín hiệu âm thanh được giải mã đã nêu.

Trong phương án, các nhóm gồm các tín hiệu trộn giảm được xử lý sao cho các tín hiệu âm thanh đầu ra thu được mà tương ứng với các đối tượng âm thanh đầu vào thuộc về nhóm các tín hiệu trộn giảm tương ứng. Do đó, tổ hợp các tín hiệu âm thanh đầu ra đến các tín hiệu âm thanh được giải mã gần với các bước cuối cùng của các quy trình giải mã được thực hiện trên tín hiệu âm thanh được mã hóa. Do đó, trong phương án này, mỗi nhóm các tín hiệu trộn giảm phải trải qua một cách riêng lẻ tất cả các bước xử lý theo sau sự phát hiện các nhóm các tín hiệu trộn giảm.

Trong phương án khác, bộ xử lý được tạo cấu hình để thực hiện ít nhất một bước xử lý riêng lẻ trên mỗi nhóm trong số nhiều nhóm các tín hiệu trộn giảm đã nêu để cung cấp các tín hiệu được xử lý như các kết quả nhóm. Thiết bị còn bao gồm bộ xử lý sau được tạo cấu hình để xử lý đồng thời các tín hiệu được xử lý đã nêu để cung cấp các tín hiệu âm thanh đầu ra. Bộ tổ hợp được tạo cấu hình để tổ hợp các tín hiệu âm

thanh đầu ra như các kết quả nhóm được xử lý để cung cấp tín hiệu âm thanh được giải mã đã nêu.

Trong phương án này, các nhóm tín hiệu trộn giảm phải trải qua ít nhất một bước xử lý riêng lẻ và đến ít nhất một bước xử lý đồng thời với các nhóm khác. Việc xử lý riêng lẻ dẫn đến các tín hiệu được xử lý mà - trong phương án - được xử lý đồng thời.

Đề cập đến các ma trận, trong một phương án, bộ xử lý được tạo cấu hình để thực hiện ít nhất một bước xử lý riêng lẻ trên các tham số đối tượng (E_k) của mỗi tập hợp các đối tượng âm thanh đầu vào để cung cấp các ma trận riêng lẻ. Bộ xử lý sau được bao gồm bởi thiết bị được tạo cấu hình để xử lý đồng thời các tham số đối tượng để cung cấp ít nhất một ma trận toàn thể. Bộ tổ hợp được tạo cấu hình để tổ hợp các ma trận riêng lẻ đã nêu và ít nhất một ma trận toàn thể đã nêu. Trong một phương án, các bộ xử lý sau thực hiện ít nhất một bước xử lý đồng thời trên các ma trận riêng lẻ để thu được ít nhất một ma trận toàn thể.

Các phương án sau đề cập đến các bước xử lý được thực hiện bởi bộ xử lý. Một số bước trong số các bước này cũng phù hợp cho bộ xử lý sau được đề cập trong phương án đã đề cập.

Trong một phương án, bộ xử lý bao gồm bộ không trộn được tạo cấu hình để không trộn các tín hiệu trộn giảm của các nhóm tương ứng của nhiều nhóm các tín hiệu trộn giảm đã nêu. Bằng cách không trộn các tín hiệu trộn giảm, bộ xử lý thu được phép biểu diễn của các đối tượng âm thanh đầu vào ban đầu mà đã được trộn giảm thành tín hiệu trộn giảm

Theo phương án, bộ không trộn được tạo cấu hình để không trộn các tín hiệu trộn giảm của nhóm tương ứng gồm nhiều nhóm các tín hiệu trộn giảm trên cơ sở thuật toán sai số bình phương trung bình tối thiểu (Minimum Mean Squared Error - MMSE). Thuật toán như vậy sẽ được giải thích trong phần mô tả dưới đây.

Trong phương án khác, trong đó bộ xử lý bao gồm bộ không trộn được tạo cấu hình để xử lý các tham số đối tượng của mỗi tập hợp các đối tượng âm thanh đầu vào riêng lẻ để cung cấp các ma trận không trộn riêng lẻ.

Trong một phương án, bộ xử lý bao gồm bộ tính toán được tạo cấu hình để tính

toán riêng lẻ cho mỗi nhóm các tín hiệu trộn giảm các ma trận với các kích thước phụ thuộc vào ít nhất một trong số số lượng các đối tượng âm thanh đầu vào của tập hợp các đối tượng âm thanh đầu vào được kết hợp với nhóm tương ứng gồm các tín hiệu trộn giảm và số lượng các tín hiệu trộn giảm thuộc về nhóm tương ứng gồm các tín hiệu trộn giảm. Vì các nhóm các tín hiệu trộn giảm nhỏ hơn toàn bộ tập hợp các tín hiệu trộn giảm và như các nhóm các tín hiệu trộn giảm đề cập đến số lượng các tín hiệu âm thanh đầu vào ít hơn, các ma trận được sử dụng cho sự xử lý các nhóm các tín hiệu trộn giảm là nhỏ hơn các ma trận được sử dụng trong tình trạng kỹ thuật. Điều này tạo thuận lợi cho sự tính toán.

Theo phương án, bộ tính toán được tạo cấu hình để tính toán cho các ma trận không trộn riêng lẻ như ngưỡng riêng lẻ dựa trên trị số năng lượng tối đa nằm trong nhóm tương ứng gồm các tín hiệu trộn giảm.

Theo phương án, bộ xử lý được tạo cấu hình để tính toán ngưỡng riêng lẻ trên cơ sở trị số năng lượng tối đa nằm trong nhóm tương ứng gồm các tín hiệu trộn giảm đối với mỗi nhóm các tín hiệu trộn giảm riêng lẻ.

Theo phương án, bộ tính toán được tạo cấu hình để tính toán cho bước quy tắc hóa để không trộn các tín hiệu trộn giảm của mỗi nhóm các tín hiệu trộn giảm ngưỡng riêng lẻ trên cơ sở trị số năng lượng tối đa nằm trong nhóm tương ứng các tín hiệu trộn giảm. Các ngưỡng cho các nhóm các tín hiệu trộn giảm được tính toán trong phương án khác bởi bản thân bộ không trộn.

Thảo luận sau đây sẽ cho thấy hiệu quả thú vị của việc tính toán ngưỡng cho các nhóm (một ngưỡng cho mỗi nhóm) và không dành cho tất cả các tín hiệu trộn giảm.

Theo phương án, bộ xử lý bao gồm bộ kết xuất được tạo cấu hình để kết xuất các tín hiệu trộn giảm không được trộn của các nhóm tương ứng cho tình huống đầu ra của tín hiệu âm thanh được giải mã đã nêu để cung cấp các tín hiệu được kết xuất. Việc kết xuất dựa trên đầu vào được cung cấp bởi người nghe hoặc dựa trên dữ liệu về tình huống đầu ra thực tế.

Trong một phương án, bộ xử lý bao gồm bộ kết xuất được tạo cấu hình để xử lý các tham số đối tượng để cung cấp ít nhất một ma trận kết xuất.

Bộ xử lý bao gồm trong phương án bộ trộn sau được tạo cấu hình để xử lý các tham số đối tượng để cung cấp ít nhất một ma trận giải tương quan.

Theo phương án, bộ xử lý bao gồm bộ trộn sau được tạo cấu hình để thực hiện ít nhất một bước giải tương quan trên các tín hiệu được kết xuất đã nêu và được tạo cấu hình để tổ hợp các kết quả (Y_{wet}) của bước giải tương quan được thực hiện với các tín hiệu được kết xuất tương ứng đã nêu (Y_{dry}).

Theo phương án, bộ xử lý được tạo cấu hình để xác định ma trận trộn giảm riêng lẻ (D_k) cho mỗi nhóm các tín hiệu trộn giảm (k là chỉ số của nhóm tương ứng), bộ xử lý được tạo cấu hình để xác định ma trận hiệp phương sai nhóm riêng lẻ (E_k) cho mỗi nhóm các tín hiệu trộn giảm, bộ xử lý được tạo cấu hình để xác định ma trận hiệp phương sai trộn giảm nhóm riêng lẻ (Δ_k) cho mỗi nhóm các tín hiệu trộn giảm trên cơ sở ma trận trộn giảm riêng lẻ (D_k) và ma trận hiệp phương sai nhóm riêng lẻ (E_k), và bộ xử lý được tạo cấu hình để xác định ma trận nhóm nghịch đảo được quy tắc hóa riêng lẻ (J_k) cho mỗi nhóm các tín hiệu trộn giảm.

Theo phương án, bộ tổ hợp được tạo cấu hình để tổ hợp các ma trận nhóm nghịch đảo được quy tắc hóa riêng lẻ (J_k) để thu được ma trận nhóm nghịch đảo được quy tắc hóa toàn thể (J).

Theo một phương án, bộ xử lý được tạo cấu hình để xác định ma trận không trộn tham số nhóm riêng lẻ (U_k) cho mỗi nhóm các tín hiệu trộn giảm trên cơ sở ma trận trộn giảm riêng lẻ (D_k), ma trận hiệp phương sai nhóm riêng lẻ (E_k), và ma trận nhóm nghịch đảo được quy chuẩn hóa riêng lẻ (J_k), và bộ tổ hợp được tạo cấu hình để tổ hợp ma trận không trộn tham số nhóm riêng lẻ (U_k) để thu được ma trận không trộn tham số nhóm toàn thể (U).

Theo một phương án, bộ xử lý được tạo cấu hình để xác định ma trận không trộn tham số nhóm riêng lẻ (U_k) cho mỗi nhóm các tín hiệu trộn giảm trên cơ sở ma trận trộn giảm riêng lẻ (D_k), ma trận hiệp phương sai nhóm riêng lẻ (E_k), và ma trận nhóm nghịch đảo được quy chuẩn hóa riêng lẻ (J_k), và bộ tổ hợp được tạo cấu hình để tổ hợp ma trận không trộn tham số nhóm riêng lẻ (U_k) để thu được ma trận không trộn tham số nhóm toàn thể (U).

Theo một phương án, bộ xử lý được tạo cấu hình để xác định ma trận kết xuất nhóm riêng lẻ ($\mathbf{R}_k$) cho mỗi nhóm gồm các tín hiệu trộn giảm.

Theo một phương án, bộ xử lý được tạo cấu hình để xác định ma trận trộn tăng riêng lẻ ($\mathbf{R}_k\mathbf{U}_k$) cho mỗi nhóm các tín hiệu trộn giảm dựa trên ma trận kết xuất nhóm riêng lẻ ($\mathbf{R}_k$) và ma trận không trộn tham số nhóm riêng lẻ ($\mathbf{U}_k$), và bộ tổ hợp được tạo cấu hình để tổ hợp các ma trận trộn tăng riêng lẻ ($\mathbf{R}_k\mathbf{U}_k$) để thu được ma trận trộn tăng toàn thể ($\mathbf{RU}$).

Theo một phương án, bộ xử lý được tạo cấu hình để xác định ma trận hiệp phương sai nhóm riêng lẻ ($\mathbf{C}_k$) cho mỗi nhóm các tín hiệu trộn giảm dựa trên ma trận kết xuất nhóm riêng lẻ ($\mathbf{R}_k$) và ma trận hiệp phương sai nhóm riêng lẻ ($\mathbf{E}_k$), và bộ tổ hợp để tổ hợp các ma trận hiệp phương sai nhóm riêng lẻ ($\mathbf{C}_k$) để thu được ma trận hiệp phương sai nhóm toàn thể ($\mathbf{C}$).

Theo một phương án, bộ xử lý được tạo cấu hình để xác định ma trận hiệp phương sai nhóm riêng lẻ của tín hiệu được ước lượng theo tham số ($\mathbf{E}_y^{\text{dry}})_k$ dựa trên ma trận kết xuất nhóm riêng lẻ ($\mathbf{R}_k$), ma trận không trộn tham số nhóm riêng lẻ ($\mathbf{U}_k$), ma trận trộn giảm riêng lẻ ($\mathbf{D}_k$), và ma trận hiệp phương sai nhóm riêng lẻ ($\mathbf{E}_k$), và bộ tổ hợp được tạo cấu hình để tổ hợp các ma trận hiệp phương sai nhóm riêng lẻ của tín hiệu được ước lượng theo tham số ($\mathbf{E}_y^{\text{dry}})_k$ để thu được tín hiệu được ước lượng theo tham số toàn thể $\mathbf{E}_y^{\text{dry}}$.

Theo một phương án, bộ xử lý được tạo cấu hình để xác định ma trận nghịch đảo được quy tắc hóa ($\mathbf{J}$) dựa trên sự khai triển trị số suy biến của ma trận hiệp phương sai trộn giảm ($\mathbf{E}_{\text{DMX}}$).

Theo một phương án, bộ xử lý được tạo cấu hình để xác định ma trận con (Δ_k) cho sự xác định ma trận không trộn theo tham số ($\mathbf{U}$), bằng cách lựa chọn các yếu tố ($\Delta(m, n)$) tương ứng với các tín hiệu trộn giảm (m, n) được gán cho nhóm tương ứng (có chỉ số k) gồm các tín hiệu trộn giảm. Mỗi nhóm các tín hiệu trộn giảm bao trùm số lượng cụ thể các tín hiệu trộn giảm và tập hợp được kết hợp gồm các đối tượng âm thanh đầu vào và được biểu thị ở đây bằng chỉ số k .

Theo phương án này, các ma trận con riêng lẻ (Δ_k) thu được nhờ lựa chọn hoặc

chọn phần tử từ ma trận hiệp phương sai trộn giảm Δ mà thuộc về nhóm tương ứng k .

Trong một phương án, các ma trận con riêng lẻ (Δ_k) được nghịch đảo một cách riêng lẻ và các kết quả được tổ hợp trong ma trận nghịch đảo được quy tắc hóa (J).

Trong một phương án khác, ma trận con (Δ_k) thu được sử dụng định nghĩa của chúng như $\Delta_k = D_k E_k D_k^*$ với ma trận trộn giảm riêng lẻ (D_k)

Theo một phương án, bộ tổ hợp được tạo cấu hình để xác định ma trận trộn sau (P) dựa trên các ma trận được xác định một cách riêng lẻ cho mỗi nhóm các tín hiệu trộn giảm và bộ tổ hợp được tạo cấu hình để áp dụng ma trận trộn sau (P) cho nhiều tín hiệu trộn giảm để thu được tín hiệu âm thanh được giải mã. Trong phương án này, từ các tham số đối tượng, ma trận trộn sau được tính toán mà được áp dụng cho tín hiệu âm thanh được mã hóa để thu được tín hiệu âm thanh được giải mã.

Theo một phương án, thiết bị và các thành phần tương ứng của thiết bị được tạo cấu hình để thực hiện đối với mỗi nhóm các tín hiệu trộn giảm riêng lẻ ít nhất một trong số các phép tính sau:

- phép tính ma trận hiệp phương sai nhóm E_k có kích thước N_k lần N_k với các thành phần: $e_{i,j}^k = \sqrt{OLD_i^k OLD_j^k} IOC_{i,j}^k$,

- phép tính ma trận hiệp phương sai trộn giảm nhóm Δ_k có kích thước M_k nhân M_k : $\Delta_k = D_k E_k D_k^*$,

- phép tính sự khai triển trị số suy biến của ma trận hiệp phương sai trộn giảm nhóm $\Delta_k = D_k E_k D_k^*$: $\Delta_k = V_k \Lambda_k V_k^*$,

- phép tính ma trận nhóm nghịch đảo được quy tắc hóa J_k xấp xỉ $J_k \approx \Delta_k^{-1}$: $J_k = V_k \Lambda_k^{inv} V_k^*$, gồm có phép tính ma trận riêng lẻ Λ_k^{inv} (các chi tiết sẽ được nêu ra dưới đây),

- phép tính ma trận không trộn theo tham số nhóm U_k có kích thước N_k nhân M_k : $U_k = E_k D_k^* J_k$,

- phép nhân ma trận kết xuất nhóm R_k có kích thước N_{Upmix} nhân N_k với ma trận không trộn U_k có kích thước N_k nhân M_k : $R_k U_k$,

- phép tính ma trận hiệp phương sai nhóm $\mathbf{C}_k$ có kích thước N_{out} nhân N_{out} : $\mathbf{C}_k = \mathbf{R}_k \mathbf{E}_k \mathbf{R}_k^*$,
- phép tính hiệp phương sai nhóm tín hiệu được ước lượng theo tham số $(\mathbf{E}_y^{\text{dry}})_k$ có kích thước N_{out} nhân N_{out} : $(\mathbf{E}_y^{\text{dry}})_k = \mathbf{R}_k \mathbf{U}_k (\mathbf{D}_k \mathbf{E}_k \mathbf{D}_k^*) \mathbf{U}_k^* \mathbf{R}_k^*$.

Trong phương diện này, k biểu thị chỉ số của nhóm tương ứng các tín hiệu trộn giảm, N_k biểu thị số lượng các đối tượng âm thanh đầu vào của tập hợp các đối tượng âm thanh đầu vào được kết hợp, M_k biểu thị số lượng các tín hiệu trộn giảm thuộc về nhóm các tín hiệu trộn giảm tương ứng, và N_{out} biểu thị số lượng các kênh đầu ra được trộn tăng hoặc được kết xuất.

Các ma trận được tính toán có kích thước nhỏ hơn các ma trận được sử dụng trong tình trạng kỹ thuật. Theo đó, theo một phương án, nhiều bước xử lý nhất có thể được thực hiện một cách riêng lẻ trên các nhóm các tín hiệu trộn giảm.

Mục đích của sáng chế cũng đạt được bằng phương pháp tương ứng để xử lý tín hiệu âm thanh được mã hóa. Tín hiệu âm thanh được mã hóa bao gồm nhiều tín hiệu trộn giảm được kết hợp với nhiều đối tượng âm thanh đầu vào và các tham số đối tượng. Phương pháp bao gồm các bước sau:

- tạo nhóm các tín hiệu trộn giảm thành nhiều nhóm các tín hiệu trộn giảm được kết hợp với tập hợp các đối tượng âm thanh đầu vào trong số nhiều đối tượng âm thanh đầu vào.
- thực hiện ít nhất một bước xử lý riêng lẻ trên các tham số đối tượng của mỗi tập hợp các đối tượng âm thanh đầu vào để cung cấp các kết quả nhóm, và
- tổ hợp các kết quả nhóm đã nêu để cung cấp tín hiệu âm thanh được giải mã.

Việc tạo nhóm được thực hiện với ít nhất sự co hẹp mà mỗi đối tượng âm thanh trong số nhiều đối tượng âm thanh thuộc về chỉ một tập hợp các đối tượng âm thanh đầu vào.

Các phương án đã nêu của thiết bị có thể cũng được thực hiện bằng các bước phương pháp và các phương án tương ứng của phương pháp. Do đó, sự giải thích được đưa ra cho các phương án của thiết bị cũng được dùng cho phương pháp.

Mô tả vắn tắt các hình vẽ

Sáng chế sẽ được giải thích dưới đây với sự viện dẫn đến các hình vẽ kèm theo và các phương án được biểu thị trong các hình vẽ kèm theo, trong đó:

Fig.1 thể hiện hình tổng quát của khái niệm trộn giảm/trộn tăng tham số trên cơ sở MMSE,

Fig.2 thể hiện hệ thống khôi phục tham số với sự giải tương quan được áp dụng trên đầu ra được kết xuất,

Fig.3 thể hiện cấu trúc bộ xử lý trộn giảm,

Fig.4 thể hiện các ảnh phổ của năm đối tượng âm thanh đầu vào (cột bên trái) và các ảnh phổ của các kênh trộn giảm tương ứng (cột bên phải),

Fig.5 thể hiện các ảnh phổ của tín hiệu đầu ra tham chiếu (cột bên trái) và các ảnh phổ của các tín hiệu đầu ra được giải mã và kết xuất 3D SAOC tương ứng,

Fig.6 thể hiện các ảnh phổ của các tín hiệu đầu ra 3D SAOC sử dụng sáng chế,

Fig.7 thể hiện sự xử lý tham số khung theo tình trạng kỹ thuật,

Fig.8 thể hiện sự xử lý tham số khung theo sáng chế,

Fig.9 thể hiện ví dụ của sự thực hiện hàm phát hiện nhóm,

Fig.10 thể hiện theo sơ đồ thiết bị để mã hóa các đối tượng âm thanh đầu vào,

Fig.11 thể hiện theo sơ đồ ví dụ về thiết bị theo sáng chế để xử lý tín hiệu âm thanh được mã hóa,

Fig.12 thể hiện theo sơ đồ ví dụ khác về thiết bị theo sáng chế để xử lý tín hiệu âm thanh được mã hóa,

Fig.13 thể hiện chuỗi các bước phương án theo phương án của sáng chế,

Fig.14 thể hiện theo sơ đồ ví dụ về thiết bị theo sáng chế,

Fig.15 thể hiện theo sơ đồ ví dụ khác về thiết bị,

Fig.16 thể hiện theo sơ đồ bộ xử lý thiết bị theo sáng chế, và

Fig.17 thể hiện theo sơ đồ ứng dụng của thiết bị theo sáng chế,

Mô tả chi tiết sáng chế

Sau đây, sự tổng quát về các sơ đồ tách sẽ được đưa ra, sử dụng ví dụ về công nghệ mã hóa đối tượng âm thanh không gian (Spatial Audio Object Coding - SAOC) MPEG ([SAOC]) và phân xử lý 3D SAOC của âm thanh 3D MPEG-H ([SAOC3D, SAOC3D2]). Các đặc tính toán học của các phương pháp này được xem xét.

Ký hiệu toán học sau đây được sử dụng:

N số lượng các đối tượng âm thanh đầu vào (lựa chọn khác: các đối tượng đầu vào)

N_{dmx} số lượng các kênh (vận chuyển) trộn giảm

N_{out} số lượng các kênh (được kết xuất) trộn tăng

N_{samples} số lượng các mẫu mỗi tín hiệu âm thanh

D ma trận trộn giảm, kích thước **N_{dmx}** nhân **N**

S tín hiệu đối tượng âm thanh đầu vào, kích thước **N** nhân **N_{samples}**

E ma trận hiệp phương sai đối tượng, kích thước **N** nhân **N**, xấp xỉ **E** $\approx$ **SS***

X các tín hiệu âm thanh trộn giảm, kích thước **N_{dmx}** nhân **N_{samples}**, được định nghĩa như **X** = **DS**

E_{DMX} ma trận hiệp phương sai các tín hiệu trộn giảm, kích thước **N_{dmx}** nhân **N_{dmx}**,

được định nghĩa như **E_{DMX}** = **DED***

U ma trận ước lượng nguồn tham số, kích thước **N** nhân **N_{dmx}**,

mà xấp xỉ **U** $\approx$ **ED*** (**DED***)⁻¹

R ma trận kết xuất (được cụ thể hóa tại phía bộ giải mã), kích thước **N_{out}** nhân **N**

Ŝ các tín hiệu đối tượng được khôi phục theo tham số, kích thước **N** nhân **N_{samples}**,

	mà xấp xỉ $\mathbf{S}$ và được xác định như $\hat{\mathbf{S}} = \mathbf{U}\mathbf{X}$,
$\mathbf{Y}_{\text{dry}}$	các tín hiệu đối tượng được khôi phục theo tham số và được kết xuất, kích thước N_{out} nhân N_{samples} , được xác định như $\mathbf{Y}_{\text{dry}} = \mathbf{R}\mathbf{U}\mathbf{X}$
$\mathbf{Y}_{\text{wet}}$	các đầu ra giải tương quan, kích thước N_{out} nhân N_{samples}
$\mathbf{Y}$	đầu ra cuối cùng, kích thước N_{out} nhân N_{samples}
$(\cdot)^*$	toán tử tự phụ hợp (Hermitian), mà biểu diễn sự chuyển vị liên hợp của $(\cdot)$
$\mathbf{F}_{\text{decorr}}(\cdot)$	hàm bộ giải tương quan

Khi không có sự tổn hao chung, để cải thiện khả năng đọc của phương trình, đối với tất cả các biến số được giới thiệu, các chỉ số biểu thị sự phụ thuộc thời gian và tần số được loại bỏ.

Các hệ thống tách đối tượng tham số:

Các sơ đồ tách tham số chung nhằm ước lượng số lượng các nguồn âm thanh từ việc trộn tín hiệu (trộn giảm) sử dụng thông tin theo tham số hỗ trợ. Giải pháp điển hình cho nhiệm vụ này là dựa trên ứng dụng các thuật toán ước lượng sai số bình phương trung bình tối thiểu (Minimum Mean Squared Error - MMSE). Công nghệ SAOC là một ví dụ về các hệ thống mã hóa âm thanh theo tham số như vậy.

Fig.1 mô tả quy tắc chung của cấu trúc bộ mã hóa/bộ giải mã SAOC.

Việc xử lý trộn giảm/trộn tăng theo tham số chung được thực hiện theo cách lựa chọn thời gian/tần số và có thể được mô tả như chuỗi gồm các bước sau:

- “Bộ mã hóa” được cung cấp với các “đối tượng âm thanh” đầu vào $\mathbf{S}$ và “các tham số trộn” $\mathbf{D}$. “Bộ trộn” trộn giảm “các đối tượng âm thanh” $\mathbf{S}$ thành số lượng các “tín hiệu trộn giảm” $\mathbf{X}$ sử dụng các “tham số trộn” $\mathbf{D}$ (ví dụ, độ khuếch đại trộn giảm).
- “Bộ ước lượng thông tin phụ” trích thông tin phụ mô tả các đặc điểm của “đối tượng âm thanh” đầu vào $\mathbf{S}$ (ví dụ, các đặc tính hiệp phương sai).
- “Các tín hiệu trộn giảm” $\mathbf{X}$ và thông tin phụ được truyền hoặc được lưu trữ. Các tín hiệu âm thanh trộn giảm này có thể còn được nén sử dụng các bộ mã hóa âm

thanh (như Lớp II hoặc III MPEG-1/2 , mã hóa âm thanh cải tiến (Advanced Audio Coding - AAC) MPEG-2/4, mã hóa tiếng nói và âm thanh hợp nhất (Unified Speech and Audio Coding - USAC) MPEG, v.v.). Thông tin phụ có thể cũng được biểu diễn và được mã hóa một cách hiệu quả (ví dụ, như các mối tương quan được mã hóa của các lũy thừa đối tượng và các hệ số tương quan đối tượng).

"Bộ giải mã" chứa "các đối tượng âm thanh" ban đầu từ "các tín hiệu trộn giảm" được giải mã sử dụng thông tin phụ được truyền dẫn (thông tin này cung cấp các tham số đối tượng). "Các bộ xử lý thông tin phụ" ước lượng các hệ số không trộn sẽ được áp dụng trên "các tín hiệu trộn giảm" trong "bộ tách tham số đối tượng" để thu được sự khôi phục đối tượng tham số của S. "Các đối tượng âm thanh" được khôi phục được kết xuất thành cảnh đích (đa kênh), được biểu diễn bởi các kênh đầu ra Y, bằng cách áp dụng "các tham số kết xuất" R.

Quy tắc chung tương tự và các bước liên tiếp được áp dụng trong việc xử lý 3D SAOC, mà kết hợp được giải tương quan bổ sung.

Fig.2 cung cấp hình tổng quát của khái niệm trộn giảm/trộn tăng theo tham số với được giải tương quan được tích hợp.

Sử dụng ví dụ của công nghệ 3D SAOC, phần của âm thanh 3D MPEG-H, các bước xử lý chính của hệ thống tách tham số như vậy có thể được tóm tắt sau đây:

Bộ giải mã 3D SAOC tạo ra đầu ra được kết xuất được cải biên Y như hỗn hợp của tín hiệu được khôi phục và được kết xuất theo tham số (tín hiệu khô) Y_{dry} và phiên bản được giải tương quan của tín hiệu đó (tín hiệu ướt) Y_{wet} .

Các bước xử lý - cho việc thảo luận mà sáng chế liên quan - có thể được phân biệt như được minh họa trên Fig.3:

- Không trộn, mà khôi phục theo tham số các đối tượng âm thanh đầu vào sử dụng ma trận U,
- Kết xuất sử dụng thông tin kết xuất (ma trận R),
- Giải tương quan,
- Trộn sau sử dụng ma trận P, được tính toán dựa trên thông tin được chứa trong

dòng bit.

Sự tách đối tượng tham số thu được từ tín hiệu trộn giảm $\mathbf{X}$ sử dụng ma trận không trộn $\mathbf{U}$ trên cơ sở thông tin phụ bổ sung: $\hat{\mathbf{S}} = \mathbf{UX}$.

Thông tin kết xuất $\mathbf{R}$ được sử dụng để thu được tín hiệu khô như: $\mathbf{Y}_{\text{dry}} = \mathbf{R} \hat{\mathbf{S}} = \mathbf{RUX}$.

Tín hiệu đầu ra cuối cùng $\mathbf{Y}$ được tính toán từ các tín hiệu $\mathbf{Y}_{\text{dry}}$ và $\mathbf{Y}_{\text{wet}}$ như

$$\mathbf{Y} = \mathbf{P} \begin{bmatrix} \mathbf{Y}_{\text{dry}} \\ \mathbf{Y}_{\text{wet}} \end{bmatrix}.$$

Ma trận trộn $\mathbf{P}$ được tính toán, ví dụ, dựa trên thông tin kết xuất, thông tin tương quan, thông tin năng lượng, thông tin hiệp phương sai, v.v.

Trong sáng chế, đây sẽ là ma trận trộn sau được áp dụng cho tín hiệu âm thanh được mã hóa để thu được tín hiệu âm thanh được giải mã.

Dưới đây, thuật toán tách đối tượng tham số chung sử dụng MMSE sẽ được giải thích.

Ma trận không trộn $\mathbf{U}$ thu được dựa trên thông tin được kết xuất từ các biến số nằm trong dòng bit (ví dụ ma trận trộn giảm $\mathbf{D}$ và thông tin hiệp phương sai $\mathbf{E}$), sử dụng thuật toán ước lượng sai số bình phương trung bình tối thiểu (Minimum Mean Squared Error - MMSE): $\mathbf{U} = \mathbf{ED}^*\mathbf{J}$.

Ma trận $\mathbf{J}$ có kích thước N_{dmx} nhân N_{dmx} biểu diễn sự ước lượng của sự nghịch đảo giả của ma trận hiệp phương sai trộn giảm $\mathbf{E}_{\text{DMX}} = \mathbf{DED}^*$ như: $\mathbf{J} \approx \mathbf{E}_{\text{DMX}}^{-1}$.

Sự tính toán của ma trận $\mathbf{J}$ được kết xuất theo: $\mathbf{J} = \mathbf{V} \mathbf{\Lambda}^{\text{inv}} \mathbf{V}^*$,

trong đó các ma trận $\mathbf{V}$ và $\mathbf{\Lambda}$ được xác định sử dụng sự khai triển trị số suy biến (singular value decomposition - SVD) của ma trận $\mathbf{E}_{\text{DMX}}$ như: $\mathbf{E}_{\text{DMX}} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^*$.

Lưu ý rằng các kết quả giống nhau có thể thu được sử dụng các phương pháp khai triển khác nhau như: khai triển trị số đặc trưng, khai triển Schur, v.v..

Thuật toán nghịch đảo được quy tắc hóa $(\cdot)^{\text{inv}}$, được sử dụng cho ma trận trị số suy biến đường chéo $\mathbf{\Lambda}$, có thể được xác định, ví dụ, như được thực hiện trong 3D SAOC, sử dụng việc cắt các trị số suy biến tương ứng với trị số suy biến cao nhất:

$$\Lambda^{inv} = \lambda_{i,j}^{-1} = \begin{cases} \frac{1}{\lambda_{i,i}} & i = j \text{ and } \lambda_{i,i} \geq T_{reg}^{\Lambda}, \\ 0 & \text{otherwise.} \end{cases}$$

Trong phương án khác, các công thức sau được sử dụng:

$$\Lambda^{inv} = \lambda_{i,j}^{-1} = \begin{cases} \frac{1}{\lambda_{i,i}} & i = j \text{ and } abs(\lambda_{i,i}) \geq T_{reg}^{\Lambda}, \\ 0 & \text{otherwise.} \end{cases}$$

Đại lượng vô hướng quy tắc hóa tương đối T_{reg}^{Λ} được xác định sử dụng ngưỡng tuyệt đối T_{reg} và trị số cực đại là Λ như: $T_{reg}^{\Lambda} = \max_i (\lambda_{i,i}) T_{reg}$, ví dụ, với $T_{reg} = 10^{-2}$.

Phụ thuộc vào định nghĩa của các trị số suy biến, $\lambda_{i,i}$ có thể được giới hạn chỉ ở các trị số dương (nếu $\lambda_{i,i} < 0$ thì $\lambda_{i,i} = abs(\lambda_{i,i})$ và $sign(\lambda_{i,i})$ được nhân với véc tơ suy biến bên trái hoặc bên phải tương ứng) hoặc các trị số âm có thể được cho phép.

Trong trường hợp thứ hai với các trị số âm là $\lambda_{i,i}$ đại lượng vô hướng quy tắc hóa tương đối T_{reg}^{Λ} được tính như: $T_{reg}^{\Lambda} = \max_i (abs(\lambda_{i,i})) T_{reg}$.

Để đơn giản hóa, sau đây định nghĩa thứ hai về T_{reg}^{Λ} sẽ được sử dụng.

Các kết quả tương tự có thể thu được sử dụng việc cắt các trị số suy biến tương ứng với trị số tuyệt đối hoặc các phương pháp quy tắc hóa khác được sử dụng cho sự nghịch đảo ma trận.

Sự nghịch đảo của các trị số suy biến rất nhỏ có thể dẫn đến các hệ số không tròn cao và sau đó dẫn đến sự khuếch đại cao của các kênh trộn giảm tương ứng. Trong trường hợp như vậy, các kênh với các mức năng lượng rất nhỏ có thể được khuếch đại sử dụng các độ khuếch đại cao và điều này có thể dẫn đến các thành phần lạ có thể nghe được. Để giảm hiệu quả không mong muốn này, các trị số suy biến nhỏ hơn ngưỡng tương ứng T_{reg}^{Λ} được cắt về 0.

Sau đây, các hạn chế được tìm ra trong công nghệ tách đối tượng tham số của tình trạng kỹ thuật sẽ được giải thích.

Các phương pháp tách đối tượng theo tham số của tình trạng kỹ thuật được mô tả chỉ rõ sử dụng sự nghịch đảo được quy tắc hóa của ma trận hiệp phương sai trộn giảm

để tránh các thành phần lạ tách. Tuy nhiên, đối với một số ngữ cảnh trộn trường hợp sử dụng thực tế, các thành phần lạ có hại gây ra bởi sự quy tắc hóa quá công kích được xác định trong đầu ra của hệ thống.

Sau đây, ví dụ về ngữ cảnh như vậy được xây dựng và được phân tích.

Số lượng $N = 5$ các đối tượng âm thanh đầu vào (S) được mã hóa sử dụng các công nghệ được mô tả (chính xác hơn, phương pháp xử lý 3D SAOC phần của âm thanh 3D MPEG-H) thành số $N_{dmx} = 3$ của các kênh trộn giảm (X).

Các đối tượng âm thanh đầu vào của ví dụ có thể bao gồm:

- một nhóm gồm hai đối tượng âm thanh được tương quan chứa các tín hiệu từ phần đệm ca kịch (Trái và Phải của cặp âm lập thể),
- một nhóm gồm một đối tượng âm thanh độc lập chứa tín hiệu tiếng nói, và
- một nhóm gồm hai đối tượng âm thanh được tương quan chứa bản ghi tiếng piano (Trái và Phải của cặp âm lập thể),

Các tín hiệu đầu vào được trộn giảm thành ba nhóm kênh vận chuyển:

- nhóm G_1 với $M_1 = 1$ kênh trộn giảm, chứa nhóm thứ nhất gồm các đối tượng,
 - nhóm G_2 với $M_2 = 1$ kênh trộn giảm, chứa nhóm thứ hai gồm các đối tượng,
- và
- nhóm G_3 với $M_3 = 1$ kênh trộn giảm, chứa nhóm thứ ba gồm các đối tượng,
- sao cho $N_{dmx} = M_1 + M_2 + M_3$.

Các ma trận trộn giảm D_k tương ứng với mỗi nhóm G_k , đối với $k = 1, 2, 3$, được xây dựng sử dụng các độ khuếch đại trộn đơn nhất, và ma trận trộn giảm hoàn toàn D được cho bởi:

$$D = \begin{bmatrix} D_1 & 0 & 0 \\ 0 & D_2 & 0 \\ 0 & 0 & D_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}, \text{ với } \begin{cases} D_1 = [1 & 1] \\ D_2 = [1] \\ D_3 = [1 & 1] \end{cases}$$

Người ta có thể lưu ý sự vắng mặt của việc trộn chéo giữa nhóm gồm hai tín hiệu đối tượng đầu tiên, tín hiệu đối tượng thứ ba, và nhóm gồm hai tín hiệu đối tượng cuối

cùng. Cũng lưu ý rằng tín hiệu đối tượng thứ ba chứa tiếng nói được trộn một mình thành một kênh trộn giảm. Do đó, sự khôi phục tốt của đối tượng này được trông đợi và theo đó cũng là sự kết xuất tốt. Các ảnh phổ của các tín hiệu đầu vào và tín hiệu trộn giảm thu được được minh họa trên Fig.4.

Sự mã hóa lỗi tín hiệu trộn giảm khả thi được sử dụng trong hệ thống thực tế được loại bỏ ở đây để dễ nêu các đặc điểm về tác dụng không mong muốn. Tại phía bộ giải mã, sự giải mã tham số 3D SAOC được sử dụng để khôi phục và kết xuất các tín hiệu đối tượng âm thanh về thiết lập 3 kênh ($N_{out} = 3$): Các kênh Trái (L), Chính giữa (C), và Phải (R).

Sự trộn lại đơn giản các đối tượng âm thanh đầu vào của ví dụ được sử dụng như sau:

- hai đối tượng âm thanh đầu tiên (phần đệm nhạc) được tắt tiếng (tức là, được kết xuất với độ khuếch đại 0),
- đối tượng đầu vào thứ ba (tiếng nói) được kết xuất đến kênh chính giữa, và
- đối tượng 4 được kết xuất đến kênh bên trái và đối tượng 5 đến kênh bên phải.

Theo đó, ma trận kết xuất được sử dụng được cho bởi:

$$\mathbf{R} = [\mathbf{R}_1 \quad \mathbf{R}_2 \quad \mathbf{R}_3] = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\text{với: } \mathbf{R}_1 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \mathbf{R}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \text{ và } \mathbf{R}_3 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Đầu ra tham chiếu có thể được tính toán bằng cách áp dụng ma trận kết xuất được cụ thể hóa một cách trực tiếp đến các tín hiệu đầu vào: $\mathbf{Y}_{ref} = \mathbf{R}\mathbf{S}$.

Các ảnh phổ của đầu ra tham chiếu và các tín hiệu đầu ra từ sự giải mã 3D SAOC và sự kết xuất được minh họa bằng hai cột của Fig.5.

Từ các ảnh phổ được thể hiện của đầu ra bộ giải mã 3D SAOC, các sự quan sát sau có thể được lưu ý:

- Kênh chính giữa chỉ chứa tín hiệu tiếng nói bị hư hại nghiêm trọng so với tín hiệu tham chiếu. Các lỗ phổ lớn có thể được nhận ra. Các lỗ phổ này (là các vùng thời gian - tần số với năng lượng thiếu) dẫn đến các thành phần lạ có thể nghe được gắt.

- Các khoảng cách phổ nhỏ được thể hiện cũng ở các kênh bên trái và bên phải, đặc biệt ở các vùng tần số thấp, trong đó hầu hết các năng lượng tín hiệu được tập trung. Cũng là các khoảng cách phổ này dẫn đến các thành phần lạ có thể nghe được.

- Không có sự trộn chéo của các nhóm đối tượng trong các kênh trộn giảm, tức là, các đối tượng được trộn trong một kênh trộn giảm không được thể hiện ở bất kỳ kênh trộn giảm nào khác. Kênh trộn giảm thứ hai chỉ chứa một đối tượng (tiếng nói); do đó, các khoảng cách phổ trong đầu ra hệ thống có thể được tạo ra chỉ vì nó được xử lý cùng với các kênh trộn giảm khác.

Dựa trên các quan sát nêu trên, có thể kết luận rằng:

- Hệ thống 3D SAOC không phải là hệ thống "thông qua", tức là, nếu một tín hiệu đầu vào được trộn một mình vào một kênh trộn giảm, chất lượng âm thanh của tín hiệu đầu vào nên được bảo toàn khi giải mã và kết xuất.

- Hệ thống 3D SAOC có thể dẫn các thành phần lạ có thể nghe được do việc xử lý các tín hiệu trộn giảm đa kênh. Chất lượng đầu ra của các đối tượng được chứa trong một nhóm các kênh trộn giảm phụ thuộc vào việc xử lý phần còn lại của các kênh trộn giảm.

Các khoảng cách phổ, cụ thể các khoảng cách ở kênh chính giữa, biểu thị rằng một số thông tin hữu ích được chứa trong các kênh trộn giảm được loại bỏ bằng cách xử lý. Sự tổn hao thông tin có thể được theo vết ngược về bước tách đối tượng tham số, chính xác hơn đến bước quy tắc hóa nghịch đảo ma trận hiệp phương sai trộn giảm.

Bằng sự định nghĩa, ma trận trộn giảm trong ví dụ có cấu trúc đường chéo khối:

$$\mathbf{D} = \begin{bmatrix} \mathbf{D}_1 & 0 & 0 \\ 0 & \mathbf{D}_2 & 0 \\ 0 & 0 & \mathbf{D}_3 \end{bmatrix}$$

Hơn nữa, do mối tương quan cụ thể giữa các đối tượng đầu vào (ví dụ, sự tạo tín hiệu các mối tương quan theo tham số) cũng như ma trận hiệp phương sai tín hiệu đối

tượng đầu vào có sẵn trong bộ giải mã có cấu trúc đường chéo khối.

$$\mathbf{E} = \begin{bmatrix} \mathbf{E}_1 & 0 & 0 \\ 0 & \mathbf{E}_2 & 0 \\ 0 & 0 & \mathbf{E}_3 \end{bmatrix}$$

Kết quả là, ma trận hiệp phương sai trộn giảm có thể biểu diễn ở dạng đường chéo khối:

$$\mathbf{E}_{\text{DMX}} = \begin{bmatrix} \mathbf{E}_1^{\text{DMX}} & 0 & 0 \\ 0 & \mathbf{E}_2^{\text{DMX}} & 0 \\ 0 & 0 & \mathbf{E}_3^{\text{DMX}} \end{bmatrix} = \begin{bmatrix} \mathbf{D}_1 \mathbf{E}_1 \mathbf{D}_1^* & 0 & 0 \\ 0 & \mathbf{D}_2 \mathbf{E}_2 \mathbf{D}_2^* & 0 \\ 0 & 0 & \mathbf{D}_3 \mathbf{E}_3 \mathbf{D}_3^* \end{bmatrix} = \mathbf{D} \mathbf{E} \mathbf{D}^*$$

Trong trường hợp này, ma trận $\mathbf{E}_{\text{DMX}}$ cũng là đường chéo khối, nhưng đối với trường hợp chung, dạng đường chéo khối của nó có thể thu được sau phép hoán vị các hàng/các cột sử dụng toán tử hoán vị Φ : $\bar{\mathbf{E}}_{\text{DMX}} = \Phi \mathbf{E}_{\text{DMX}} \Phi^*$.

Toán tử hoán vị Φ được xác định như ma trận thu được bằng phép hoán vị các hàng của ma trận đồng nhất. Nếu ma trận đối xứng $\mathbf{A}$ có thể được biểu diễn ở dạng đường chéo khối bằng cách hoán vị các hàng và các cột, toán tử hoán vị có thể được sử dụng để diễn tả ma trận thu được $\bar{\mathbf{A}}$ như: $\bar{\mathbf{A}} = \Phi \mathbf{A} \Phi^*$.

Nếu Φ là toán tử hoán vị thì các đặc tính sau sẽ chiếm:

- đầu tiên, nếu $\mathbf{V}$ là ma trận đơn nguyên, khi đó $\mathbf{T} = \Phi \mathbf{V}$ cũng là ma trận đơn nguyên, và
- thứ hai, $\Phi \Phi^* = \Phi^* \Phi = \mathbf{I}$ với ma trận đồng nhất $\mathbf{I}$.

Kết quả là, các toán tử hoán vị là rõ ràng đối với các thuật toán khai triển trị số suy biến. Điều này có nghĩa rằng ma trận ban đầu $\mathbf{A}$ và ma trận được hoán vị $\bar{\mathbf{A}}$ chia sẻ các trị số đơn và các vectơ suy biến được hoán vị giống nhau:

$$\mathbf{V} \Lambda \mathbf{V}^* = \mathbf{A} \Rightarrow \begin{cases} (\Phi \mathbf{V}) \Lambda (\Phi \mathbf{V})^* = \Phi \mathbf{A} \Phi^* \\ (\Phi \mathbf{V}) \Lambda (\Phi \mathbf{V})^* = \bar{\mathbf{A}} \end{cases} \Rightarrow \mathbf{T} \Lambda \mathbf{T}^* = \bar{\mathbf{A}}, \quad \text{with } \mathbf{T} = \Phi \mathbf{V}$$

Do sự biểu diễn đường chéo khối, các trị số đơn của ma trận $\mathbf{E}_{\text{DMX}}$ có thể được tính toán bằng cách áp dụng SVD vào ma trận $\mathbf{E}_{\text{DMX}}$ hoặc bằng cách áp dụng SVD vào các ma trận con đường chéo khối $\mathbf{E}_{\text{DMX}_k}$ và tổ hợp các kết quả:

$$\mathbf{E}_{\text{DMX}} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^* = \begin{bmatrix} \mathbf{V}_1 \mathbf{\Lambda}_1 \mathbf{V}_1^* & 0 & 0 \\ 0 & \mathbf{V}_2 \mathbf{\Lambda}_2 \mathbf{V}_2^* & 0 \\ 0 & 0 & \mathbf{V}_3 \mathbf{\Lambda}_3 \mathbf{V}_3^* \end{bmatrix}$$

với $\mathbf{\Lambda} = \begin{bmatrix} \lambda_{1,1} & 0 & 0 \\ 0 & \lambda_{2,2} & 0 \\ 0 & 0 & \lambda_{3,3} \end{bmatrix}$, $\mathbf{\Lambda}_1 = [\lambda_{1,1}]$, $\mathbf{\Lambda}_2 = [\lambda_{2,2}]$ và $\mathbf{\Lambda}_3 = [\lambda_{3,3}]$.

Vì các trị số suy biến của ma trận hiệp phương sai trộn giảm liên quan trực tiếp đến các mức năng lượng của các kênh trộn giảm (mà được mô tả bởi đường chéo chính của ma trận $\mathbf{E}_{\text{DMX}}$):

$$\sum_{k=1}^{N_{\text{dmx}}} \lambda_{k,k} = \sum_{k=1}^{N_{\text{dmx}}} \mathbf{E}_{\text{DMX}}(k, k)$$

và các đối tượng nằm trong một kênh không nằm trong bất kỳ kênh trộn giảm khác, người ta có thể kết luận rằng mỗi trị số suy biến tương ứng với một kênh trộn giảm.

Do đó, nếu một kênh trong số các kênh trộn giảm có mức năng lượng nhỏ hơn các kênh trộn giảm còn lại, trị số suy biến tương ứng với kênh này sẽ nhỏ hơn nhiều so với các trị số suy biến còn lại.

Bước cắt được sử dụng trong sự nghịch đảo của ma trận chứa các trị số suy biến của ma trận $\mathbf{E}_{\text{DMX}}$:

$$\mathbf{\Lambda}^{\text{inv}} = \lambda_{i,j}^{-1} = \begin{cases} \frac{1}{\lambda_{i,i}} & i = j \text{ and } \lambda_{i,i} \geq T_{\text{reg}}^{\Lambda}, \\ 0 & \text{otherwise,} \end{cases}$$

hoặc

$$\mathbf{\Lambda}^{\text{inv}} = \lambda_{i,j}^{-1} = \begin{cases} \frac{1}{\lambda_{i,i}} & i = j \text{ and } \text{abs}(\lambda_{i,i}) \geq T_{\text{reg}}^{\Lambda}, \\ 0 & \text{otherwise,} \end{cases}$$

có thể dẫn đến sự cắt các trị số suy biến tương ứng với kênh trộn giảm với mức năng lượng nhỏ (so với kênh trộn giảm với năng lượng lớn nhất). Vì điều này, thông tin có trong kênh trộn giảm này với năng lượng tương đối nhỏ được loại bỏ và các khoảng cách phổ được quan sát trong các hình vẽ ảnh phổ và đầu ra âm thanh được tạo ra.

Để dễ hiểu hơn, phải xem xét việc trộn giảm các đối tượng âm thanh đầu vào xảy ra cho mỗi mẫu và mỗi băng tần một cách riêng biệt. Đặc biệt việc tách thành các băng khác nhau giúp hiểu vì sao các khoảng cách có thể được tìm thấy trong các ảnh phổ của các tín hiệu đầu ra tại các tần số khác nhau.

Vấn đề được nhận biết có thể được tách biệt xuống thực tế rằng ngưỡng quy tắc hóa tương ứng được tính toán cho các trị số suy biến mà không xem xét rằng ma trận sẽ được đảo ngược là đường chéo khối: $T_{reg}^{\Lambda} = \max_i (abs(\lambda_{i,i})) T_{reg}$.

Mỗi ma trận đường chéo khối tương ứng với một nhóm độc lập các kênh trộn giảm. Việc cắt được nhận biết tương ứng với trị số suy biến lớn nhất, nhưng trị số này mô tả chỉ một nhóm các kênh. Do đó, sự khôi phục các đối tượng chứa trong tất cả các nhóm độc lập gồm các kênh trộn giảm trở nên phụ thuộc vào nhóm chứa trị số suy biến lớn nhất này.

Sau đây, sáng chế sẽ được giải thích dựa trên phương án được thảo luận ở trên liên quan đến tình trạng kỹ thuật.

Xem xét ví dụ được mô tả trên đây, ba ma trận hiệp phương sai có thể được kết hợp với ba nhóm khác nhau gồm các kênh trộn giảm G_k với $1 \leq k \leq 3$. Các đối tượng âm thanh hoặc các đối tượng âm thanh đầu vào nằm trong các kênh trộn giảm của mỗi nhóm không nằm trong bất kỳ nhóm nào khác. Ngoài ra, không có mối liên quan nào (ví dụ, sự tương quan) được tạo tín hiệu giữa các đối tượng nằm trong các kênh trộn giảm từ các nhóm khác.

Để giải quyết vấn đề đã được chỉ ra của hệ thống khôi phục tham số, phương pháp theo sáng chế đề xuất áp dụng bước quy tắc hóa một cách độc lập cho mỗi nhóm. Điều này ám chỉ rằng ba ngưỡng khác nhau được tính toán cho sự nghịch đảo của ba ma trận hiệp phương sai trộn giảm độc lập. $T_{reg}^{\Lambda_k} = \max_{i \in G_k} (abs(\lambda_{i,i})) T_{reg}$, trong đó $1 \leq k \leq 3$. Do đó, trong sáng chế, trong một phương án, ngưỡng như vậy được tính toán cho mỗi nhóm một cách riêng biệt và không như trong tình trạng kỹ thuật một ngưỡng tổng cộng cho các băng tần và các mẫu tương ứng.

Sự nghịch đảo cho các trị số đơn thu được theo đó bằng cách áp dụng sự quy tắc hóa một cách độc lập cho các ma trận con E^{DMX_k} , với $1 \leq k \leq 3$:

$$\Lambda_k^{inv} = (\lambda_{i,j}^{-1})_{i,j \in G_k} = \begin{cases} \frac{1}{\lambda_{i,i}} & i = j \text{ and } \lambda_{i,i} \geq T_{reg}^{\Lambda, G_k}, \\ 0 & \text{otherwise.} \end{cases}$$

Trong phương án khác, các công thức sau được sử dụng:

$$\Lambda_k^{inv} = (\lambda_{i,j}^{-1})_{i,j \in G_k} = \begin{cases} \frac{1}{\lambda_{i,i}} & i = j \text{ and } abs(\lambda_{i,i}) \geq T_{reg}^{\Lambda, G_k}, \\ 0 & \text{otherwise.} \end{cases}$$

Sử dụng phương pháp theo sáng chế được đề xuất trong hệ thống 3D SAOC đồng nhất khác cho ví dụ được thảo luận trong phần trước, chất lượng đầu ra âm thanh của đầu ra được giải mã và kết xuất cải thiện. Các tín hiệu thu được được minh họa trên Fig.6.

So sánh các ảnh phổ ở cột bên phải của Fig.5 và Fig.6, có thể quan sát rằng phương pháp theo sáng chế giải quyết các vấn đề được chỉ ra trong hệ thống tách theo tham số của tình trạng kỹ thuật hiện có. Phương pháp theo sáng chế đảm bảo đặc điểm "đi qua" của hệ thống, và quan trọng nhất, các khoảng cách phổ được loại bỏ.

Giải pháp được mô tả để xử lý ba nhóm độc lập gồm các kênh trộn giảm có thể dễ dàng được tạo ra cho bất kỳ số lượng nhóm nào.

Phương pháp theo sáng chế được đề xuất để cải biên công nghệ tách đối tượng theo tham số bằng cách tận dụng thông tin tạo nhóm trong phép nghịch đảo của ma trận hiệp phương sai tín hiệu trộn giảm. Điều này dẫn đến sự cải thiện đáng kể của chất lượng đầu ra âm thanh.

Việc tạo nhóm có thể thu được, ví dụ, từ thông tin trộn và/hoặc tương quan đã có sẵn trong bộ giải mã mà không cần tạo tín hiệu thêm.

Chính xác hơn, một nhóm được xác định trong một phương án bằng tập hợp nhỏ nhất gồm các tín hiệu trộn giảm với hai đặc tính sau trong ví dụ:

- Thứ nhất, các đối tượng âm thanh đầu vào nằm trong các kênh trộn giảm này không nằm trong bất kỳ kênh trộn giảm nào khác.
- Thứ hai, tất cả các tín hiệu đầu vào nằm trong các kênh trộn giảm của một nhóm không liên quan (ví dụ, không có sự liên tương quan nào được tạo tín hiệu trong

tín hiệu âm thanh được mã hóa) đến bất kỳ tín hiệu đầu vào khác nằm trong các kênh trộn giảm của bất kỳ nhóm khác nào. Sự liên tương quan như vậy có nghĩa là sự xử lý được tổ hợp các đối tượng âm thanh tương ứng trong suốt sự giải mã.

Dựa trên định nghĩa nhóm được giới thiệu, số lượng K ($1 \leq K \leq N_{dmx}$) nhóm có thể được xác định: G_k ($1 \leq k \leq K$) và ma trận hiệp phương sai trộn giảm E_{DMX} có thể được biểu thị sử dụng dạng đường chéo khối bằng cách áp dụng toán tử hoán vị Φ :

$$\bar{E}_{DMX} = \Phi E_{DMX} \Phi^* = \begin{bmatrix} E_1^{DMX} & 0 & \dots & 0 \\ 0 & E_2^{DMX} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & E_K^{DMX} \end{bmatrix}$$

Các ma trận phụ E_k^{DMX} được dựng bằng cách lựa chọn các yếu tố của ma trận hiệp phương sai trộn giảm tương ứng với các nhóm độc lập G_k . Đối với mỗi nhóm G_k , ma trận E_k^{DMX} có kích thước M_k nhân M_k được thể hiện sử dụng SVD như: $E_k^{DMX} = V_k \Lambda_k V_k^*$

$$\text{với: } \Lambda_k = \begin{bmatrix} \lambda_{1,1}^k & 0 & \dots & 0 \\ 0 & \lambda_{2,2}^k & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_{M_k, M_k}^k \end{bmatrix} \quad \text{và} \quad \sum_{k=1}^K M_k = N_{dmx}.$$

Sự nghịch đảo giả của ma trận E_k^{DMX} được tính toán như $(E_k^{DMX})^{-1} = V_k \Lambda_k^{inv} V_k^*$ trong đó ma trận nghịch đảo được quy tắc hóa Λ_k^{inv} được cho trong một phương án bởi:

$$\Lambda_k^{inv} = (\lambda_{i,j}^{-1})_{i,j \in G_k} = \begin{cases} \frac{1}{\lambda_{i,i}} & i = j \text{ và } \lambda_{i,i} \geq T_{reg}^{\Lambda_k}, \\ 0 & \text{neueu khong.} \end{cases}$$

và trong phương án khác bởi:

$$\Lambda_k^{inv} = (\lambda_{i,j}^{-1})_{i,j \in G_k} = \begin{cases} \frac{1}{\lambda_{i,i}} & i = j \text{ và } abs(\lambda_{i,i}) \geq T_{reg}^{\Lambda_k}, \\ 0 & \text{neueu khong.} \end{cases}$$

Đại lượng vô hướng được quy tắc hóa tương ứng $T_{reg}^{\Lambda_k}$ được xác định sử dụng ngưỡng tuyệt đối T_{reg} và trị số tối đa của Λ_k là: $T_{reg}^{\Lambda_k} = \max_{i \in G_k} (\lambda_{i,i}) T_{reg}$ ví dụ, với $T_{reg} = 10^{-2}$.

Sự nghịch đảo của ma trận hiệp phương sai trộn giảm được hoán vị $\bar{\mathbf{E}}_{\text{DMX}}$ thu được là:

$$\bar{\mathbf{E}}_{\text{DMX}}^{-1} = \begin{bmatrix} (\mathbf{E}_1^{\text{DMX}})^{-1} & 0 & \dots & 0 \\ 0 & (\mathbf{E}_2^{\text{DMX}})^{-1} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & (\mathbf{E}_K^{\text{DMX}})^{-1} \end{bmatrix}$$

và sự nghịch đảo của ma trận hiệp phương sai trộn giảm được tính toán bằng cách áp dụng toán tử hóa vị nghịch đảo: $\mathbf{E}_{\text{DMX}}^{-1} = \Phi^* \bar{\mathbf{E}}_{\text{DMX}}^{-1} \Phi$.

Ngoài ra, phương pháp theo sáng chế đề xuất trong một phương án để xác định các nhóm dựa hoàn toàn trên thông tin nằm trong dòng bit. Ví dụ, thông tin này có thể được cho bởi thông tin trộn giảm và thông tin tương quan.

Chính xác hơn, một nhóm $\mathbf{G}_k$ được định nghĩa bởi tập hợp nhỏ nhất gồm các kênh trộn giảm với các đặc tính sau:

- Các đối tượng âm thanh đầu vào nằm trong các kênh trộn giảm của nhóm $\mathbf{G}_k$ không nằm trong bất kỳ kênh trộn giảm nào khác. Đối tượng âm thanh đầu vào không nằm trong kênh trộn giảm, ví dụ, nếu độ khuếch đại trộn giảm được cho bởi chỉ số lượng tử hóa nhỏ nhất, hoặc nếu nó bằng 0.
- Tất cả các tín hiệu đầu vào i nằm trong các kênh trộn giảm của nhóm $\mathbf{G}_k$ không liên quan đến bất kỳ tín hiệu đầu vào nào j nằm trong bất kỳ kênh trộn giảm nào trong số bất kỳ nhóm khác. Ví dụ, (so sánh với WO 2011/039195 A1) biến số dòng bit $\text{bsRelatedTo}[i][j]$ có thể được sử dụng để tạo tín hiệu nếu hai đối tượng có liên quan ($\text{bsRelatedTo}[i][j] == 1$) hoặc nếu chúng không liên quan ($\text{bsRelatedTo}[i][j] == 0$). Cũng là các phương pháp khác về việc tạo tín hiệu hai đối tượng liên quan có thể được sử dụng dựa trên ví dụ, thông tin tương quan hoặc hiệp phương sai.

Các nhóm có thể được xác định một lần mỗi khung hoặc một lần mỗi tham số được thiết lập cho tất cả các băng xử lý, hoặc một lần mỗi khung hoặc một lần mỗi tham số được thiết lập cho mỗi băng xử lý.

Phương pháp theo sáng chế cũng cho phép trong một phương án giảm đáng kể độ

phức tạp tính toán của hệ thống tách tham số (ví dụ, bộ giải mã 3D SAOC) bằng cách tận dụng thông tin nhóm trong các thành phần xử lý theo tham số đất về tính toán nhất.

Do đó, phương pháp theo sáng chế đề xuất loại bỏ các phép tính toán mà không mang lại đóng góp nào cho chất lượng âm thanh đầu ra cuối cùng. Các phép tính này có thể được lựa chọn dựa trên thông tin nhóm.

Chính xác hơn, phương pháp theo sáng chế đề xuất để tính toán tất cả các bước xử lý theo tham số một cách độc lập cho mỗi nhóm được xác định trước và cuối cùng để tổ hợp các kết quả.

Sử dụng ví dụ về phần xử lý 3D SAOC của Âm thanh 3D MPEG-H, các thuật toán phức tạp về tính toán được cho bởi:

- phép tính của ma trận hiệp phương sai $\mathbf{E}$ có kích thước N nhân N với các phần tử: $e_{i,j} = \sqrt{OLD_i OLD_j} IOC_{i,j}$,

- phép tính của ma trận hiệp phương sai $\mathbf{\Delta}$ có kích thước N_{dmx} nhân N_{dmx} : $\mathbf{\Delta} = \mathbf{DED}^*$,

- phép tính của phép khai triển trị số suy biến của ma trận $\mathbf{\Delta} = \mathbf{DED}^*$: $\mathbf{\Delta} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^*$,

- phép tính của ma trận nghịch đảo được quy tắc hóa $\mathbf{J}$ xấp xỉ $\mathbf{J} \approx \mathbf{\Delta}^{-1}$: $\mathbf{J} = \mathbf{V} \mathbf{\Lambda}^{inv} \mathbf{V}^*$,

- phép tính của ma trận không trộn theo tham số $\mathbf{U}$ có kích thước N nhân N_{dmx} : $\mathbf{U} = \mathbf{ED}^* \mathbf{J}$,

- phép nhân ma trận kết xuất $\mathbf{R}$ có kích thước N_{out} nhân N với ma trận không trộn $\mathbf{U}$ có kích thước N nhân N_{dmx} : $\mathbf{RU}$,

- phép tính của ma trận hiệp phương sai $\mathbf{C}$ có kích thước N_{out} nhân N_{out} : $\mathbf{C} = \mathbf{RER}^*$,

- phép tính hiệp phương sai của tín hiệu được ước lượng theo tham số $\mathbf{E}_y^{dry}$ có kích thước N_{out} nhân N_{out} : $\mathbf{E}_y^{dry} = \mathbf{RU}(\mathbf{DED}^*) \mathbf{U}^* \mathbf{R}^*$.

Các chênh lệch mức đối tượng (Object Level Differences - OLD) đề cập đến

năng lượng tương ứng của một đối tượng so với đối tượng có năng lượng nhiều nhất đối với băng thời gian và tần số nhất định và sự nhất quán chéo liên đối tượng (Inter-Object Cross Coherence - IOC) mô tả lượng tương tự, hoặc sự tương quan chéo cho hai đối tượng trong băng thời gian và tần số nhất định.

Phương pháp theo sáng chế đề xuất việc giảm độ phức tạp tính toán bằng cách tính toán tất cả các bước xử lý theo tham số cho tất cả K nhóm được định trước $\mathbf{G}_k$ với $1 \leq k \leq K$ một cách độc lập, và tổ hợp các kết quả ở cuối việc xử lý tham số.

Một nhóm $\mathbf{G}_k$ chứa M_k kênh trộn giảm và N_k đối tượng âm thanh đầu vào sao

$$\text{cho: } \sum_{k=1}^K M_k = N_{dmx} \text{ và } \sum_{k=1}^K N_k = N.$$

Đối với mỗi nhóm $\mathbf{G}_k$, ma trận trộn giảm nhóm được định nghĩa là $\mathbf{D}_k$ bằng cách lựa chọn các phần tử của ma trận trộn giảm $\mathbf{D}$ tương ứng với các kênh trộn giảm và các đối tượng âm thanh đầu vào nằm trong nhóm $\mathbf{G}_k$.

Tương tự với ma trận kết xuất nhóm $\mathbf{R}_k$ thu được khỏi ma trận kết xuất $\mathbf{R}$ bằng cách lựa chọn các hàng tương ứng với các đối tượng âm thanh đầu vào nằm trong nhóm $\mathbf{G}_k$.

Tương tự vectơ nhóm OLD^k và ma trận nhóm IOC^k thu được trong số vectơ OLD và ma trận IOC bằng cách lựa chọn các phần tử tương ứng với các đối tượng âm thanh đầu vào nằm trong nhóm $\mathbf{G}_k$.

Đối với mỗi nhóm $\mathbf{G}_k$, các bước xử lý được mô tả được thay thế bằng các bước xử lý được thay thế với các bước xử lý về tính toán như sau:

- phép tính ma trận hiệp phương sai nhóm $\mathbf{E}_k$ có kích thước N_k nhân N_k với các phần tử: $e_{i,j}^k = \sqrt{\text{OLD}_i^k \text{OLD}_j^k} \text{IOC}_{i,j}^k$,

- phép tính ma trận hiệp phương sai trộn giảm nhóm Δ_k có kích thước M_k nhân M_k : $\Delta_k = \mathbf{D}_k \mathbf{E}_k \mathbf{D}_k^*$,

- phép tính phép khai triển trị số suy biến của ma trận hiệp phương sai trộn giảm nhóm $\Delta_k = \mathbf{D}_k \mathbf{E}_k \mathbf{D}_k^*$: $\Delta_k = \mathbf{V}_k \Lambda_k \mathbf{V}_k^*$,

- phép tính ma trận nhóm nghịch đảo được quy tắc hóa $\mathbf{J}_k$ xấp xỉ $\mathbf{J}_k \approx \Delta_k^{-1}$:

$$\mathbf{J}_k = \mathbf{V}_k \mathbf{\Lambda}_k^{inv} \mathbf{V}_k^*,$$

- phép tính ma trận không trộn theo tham số nhóm $\mathbf{U}_k$ có kích thước N_k nhân M_k :

$$\mathbf{U}_k = \mathbf{E}_k \mathbf{D}_k^* \mathbf{J}_k,$$

- phép nhân ma trận kết xuất nhóm $\mathbf{R}_k$ có kích thước N_{upmix} nhân N_k với ma trận không trộn $\mathbf{U}_k$ có kích thước N_k nhân M_k : $\mathbf{R}_k \mathbf{U}_k$,

- phép tính ma trận hiệp phương sai nhóm $\mathbf{C}_k$ có kích thước N_{out} nhân N_{out} : $\mathbf{C}_k = \mathbf{R}_k \mathbf{E}_k \mathbf{R}_k^*$,

- phép tính hiệp phương sai nhóm tín hiệu được ước lượng theo tham số $(\mathbf{E}_Y^{dry})_k$ có kích thước N_{out} nhân N_{out} : $(\mathbf{E}_Y^{dry})_k = \mathbf{R}_k \mathbf{U}_k (\mathbf{D}_k \mathbf{E}_k \mathbf{D}_k^*) \mathbf{U}_k^* \mathbf{R}_k^*$.

Và các kết quả của các bước xử lý nhóm riêng lẻ được tổ hợp ở cuối cùng:

- ma trận trộn tăng $\mathbf{RU}$ có kích thước N_{out} nhân N_{dmx} thu được bằng cách trộn các ma trận nhóm $\mathbf{R}_k \mathbf{U}_k$: $\mathbf{RU} = [\mathbf{R}_1 \mathbf{U}_1 \quad \mathbf{R}_2 \mathbf{U}_2 \quad \dots \quad \mathbf{R}_K \mathbf{U}_K]$,

- ma trận hiệp phương sai $\mathbf{C}$ có kích thước N_{out} nhân N_{out} thu được bằng cách tính tổng các ma trận nhóm $\mathbf{C}_k$: $\mathbf{C} = \sum_{k=1}^K \mathbf{C}_k$,

- hiệp phương sai của tín hiệu được ước lượng theo tham số $\mathbf{E}_Y^{dry}$ có kích thước N_{out} nhân N_{out} thu được bằng cách tính tổng các ma trận nhóm $(\mathbf{E}_Y^{dry})_k$: $\mathbf{E}_Y^{dry} = \sum_{k=1}^K (\mathbf{E}_Y^{dry})_k$

Tính tổng các bước xử lý theo cấu trúc của bộ xử lý trộn giảm được minh họa trên Fig.3, trong khi loại bỏ bước giải tương quan, việc xử lý tham số khung theo tình trạng kỹ thuật có thể được mô tả như trên Fig.7.

Sử dụng phương pháp theo sáng chế được đề xuất, độ phức tạp tính toán được giảm sử dụng sự phát hiện nhóm như được minh họa trên Fig.8.

Ví dụ về việc thực hiện hàm phát hiện nhóm, được gọi là: $[K, G_k] = groupDetect(\mathbf{D}, RelatedTo)$, được cho trên Fig.9 sử dụng mã ANSI C và hàm tính

“getSaocCoreGroups()”.

Phương pháp theo sáng chế được đề xuất chứng tỏ sẽ hiệu quả hơn về tính toán một cách đáng kể so với thực hiện các thuật toán mà không tạo nhóm. Nó cũng cho phép sự cấp phát bộ nhớ tốt hơn và việc sử dụng, hỗ trợ song song hóa phép tính toán, giảm sự tích tụ sai số bằng số, v.v.

Phương pháp theo sáng chế được đề xuất và thiết bị theo sáng chế được đề xuất giải quyết vấn đề hiện có của các hệ thống tách đối tượng theo tham số trong tình trạng kỹ thuật và đưa ra chất lượng âm thanh đầu ra cao hơn đáng kể.

Phương pháp theo sáng chế được đề xuất mô tả phương pháp phát hiện nhóm mà được nhận biết toàn bộ dựa trên thông tin dòng bit hiện có.

Giải pháp tạo nhóm theo sáng chế được đề xuất dẫn đến việc giảm đáng kể độ phức tạp tính toán. Nhìn chung, sự khai triển trị số suy biến là đắt về tính toán và độ phức tạp của nó phát triển theo hàm mũ với kích thước của ma trận sẽ được nghịch đảo: $O(N_{dmx}^3)$.

Đối với số lượng lớn các kênh trộn giảm, tính toán K lần thuật toán SVD đối với ma trận kích thước nhỏ hơn hiệu quả nhiều hơn về tính toán: $\sum_{k=1}^K O(M_k^3)$.

Sử dụng sự cân nhắc tương tự, tất cả các bước xử lý theo tham số trong bộ giải mã có thể được thực hiện một cách hiệu quả bằng cách tính toán tất cả các phép nhân ma trận được mô tả trong hệ thống chỉ cho các nhóm độc lập và tổ hợp các kết quả.

Sự ước lượng việc giảm độ phức tạp cho số lượng khác nhau gồm các đối tượng âm thanh đầu vào, tức là, các đối tượng âm thanh đầu vào, các kênh trộn giảm, và số lượng cố định trong số 24 kênh đầu ra được cho bởi bảng sau:

Số lượng các đối tượng âm thanh đầu vào	8	16	32	60	96	128	256
Số lượng các kênh trộn giảm, N_{dmx}	4	8	16	24	24	32	64

Số lượng các nhóm, K	2	4	4	6	6	8	8
Sự xử lý tham số 3D SAOC [MOPS]	7,5	28	56	464	1000	2022	12000
Sự xử lý tham số phương pháp theo sáng chế [MOPS]	3	3	7,5	10	20	20	81
Giảm độ phức tạp [%]	60,00	89,29	86,61	97,84	98,00	99,01	99,33

Sáng chế thể hiện các lợi ích bổ sung sau:

- Đối với các tình huống khi chỉ một nhóm có thể được tạo ra, đầu ra là giống về bit với trạng thái hiện thời của hệ thống nghệ thuật.
- Việc tạo nhóm duy trì đặc tính "đi qua" của hệ thống. Điều này có nghĩa rằng nếu một đối tượng âm thanh đầu vào được trộn một mình vào một kênh trộn giảm, bộ giải mã có khả năng khôi phục nó một cách hoàn hảo.

Sáng chế dẫn đến các cải biên minh họa được đề xuất sau đối với văn bản tiêu chuẩn.

Thêm vào “9.5.4.2.4 Thuật toán nghịch đảo được quy tắc hóa”:

Ma trận nghịch đảo được quy tắc hóa $\mathbf{J}$ xấp xỉ $\mathbf{J} \approx \Delta^{-1}$ được tính toán như $\mathbf{J} = \mathbf{V} \mathbf{\Lambda}^{inv} \mathbf{V}^*$.

Các ma trận $\mathbf{V}$ và $\mathbf{\Lambda}$ được xác định như sự khai triển trị số suy biến của ma trận Δ như: $\Delta = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^*$.

Sự nghịch đảo được quy tắc hóa $\mathbf{\Lambda}^{inv}$ của ma trận trị số suy biến chéo $\mathbf{\Lambda}$ được tính toán theo 9.5.4.2.5.

Trong trường hợp ma trận Δ được sử dụng trong phép tính của ma trận không trộn theo tham số $\mathbf{U}$, các phép tính được mô tả được áp dụng cho tất cả các ma trận con Δ_k . Ma trận con Δ_k thu được bằng cách lựa chọn các phần tử Δ (m, n) tương ứng với các kênh trộn giảm m và n được gán vào nhóm k.

Nhóm k được xác định bởi tập hợp nhỏ nhất gồm các kênh trộn giảm với các đặc tính sau:

- Các tín hiệu đầu vào nằm trong các kênh trộn giảm của nhóm k không nằm trong bất kỳ kênh trộn giảm nào khác. Tín hiệu đầu vào không nằm trong kênh trộn giảm nếu độ khuếch đại trộn giảm tương ứng được cho bởi chỉ số lượng tử hóa nhỏ nhất (Bảng 49 của ISO/IEC 23003-2:2010).

- Tất cả các tín hiệu đầu vào i nằm trong các kênh trộn giảm của nhóm k không liên quan đến bất kỳ tín hiệu đầu vào nào nằm trong kênh trộn giảm bất kỳ trong số bất kỳ nhóm nào khác (tức là, $bsRelatedTo[i][j] == 0$).

Các kết quả của các thuật toán nghịch đảo được quy tắc hóa độc lập $\mathbf{J}_k \approx \Delta_k^{-1}$ được tổ hợp để thu được ma trận $\mathbf{J}$.

Sáng chế cũng dẫn đến các cải biên minh họa được đề xuất dưới đây cho văn bản tiêu chuẩn.

9.5.4.2.5 Thuật toán nghịch đảo được quy tắc hóa

Ma trận nghịch đảo được quy tắc hóa $\mathbf{J}$ xấp xỉ $\mathbf{J} \approx \Delta^{-1}$ được tính toán như:

$$\mathbf{J} = \mathbf{V} \Delta^{inv} \mathbf{V}^*.$$

Các ma trận $\mathbf{V}$ và Δ được xác định như sự khai triển trị số suy biến của ma trận Δ như:

$$\mathbf{V} \Delta \mathbf{V}^* = \Delta.$$

Sự nghịch đảo được quy tắc hóa Δ^{inv} của ma trận trị số suy biến chéo Δ được tính toán theo 9.5.4.2.6.

Trong trường hợp ma trận Δ được sử dụng trong phép tính ma trận không trộn theo tham số $\mathbf{U}$, các thuật toán được mô tả được áp dụng cho tất cả các ma trận con Δ_q . Ma trận con Δ_q có kích thước $N_g^q \times N_g^q$, với các phần tử $\Delta_q(idx_1, idx_2)$, thu được bằng cách lựa chọn các phần tử $\Delta(ch_1, ch_2)$ tương ứng với các kênh trộn giảm ch_1 và ch_2 được gán cho nhóm $\mathbf{g}_q$ (tức là, $\mathbf{g}_q(idx_1) = ch_1$ và $\mathbf{g}_q(idx_2) = ch_2$).

Nhóm $\mathbf{g}_q$ có kích thước $1 \times N_g^q$ được xác định bởi tập hợp nhỏ nhất các kênh trộn

giảm với các đặc tính sau đây:

- Các tín hiệu đầu vào nằm trong các kênh trộn giảm của nhóm $\mathbf{g}_q$ không nằm trong bất kỳ kênh trộn giảm nào khác. Tín hiệu đầu vào không nằm trong kênh trộn giảm nếu độ khuếch đại trộn giảm tương ứng được cho bởi chỉ số lượng tử hóa nhỏ nhất (Bảng 49 của ISO/IEC 23003-2:2010).

- Tất cả các tín hiệu đầu vào i nằm trong các kênh trộn giảm của nhóm $\mathbf{g}_q$ không liên quan đến bất kỳ tín hiệu đầu vào nào j nằm trong bất kỳ kênh trộn giảm của bất kỳ nhóm nào khác (tức là, $\mathbf{bsRelatedTo}[i][j] == 0$).

Các kết quả của các thuật toán nghịch đảo được quy tắc hóa độc lập $\mathbf{J}_q \approx \Delta_q^{-1}$ được tổ hợp để thu được ma trận $\mathbf{J}$ như:

$$\mathbf{J}(ch_1, ch_2) = \begin{cases} \mathbf{J}_q(id_{x_1}, id_{x_2}), & \text{if } \mathbf{g}_q(id_{x_1}) = ch_1 \text{ and } \mathbf{g}_q(id_{x_2}) = ch_2, \\ 0, & \text{otherwise.} \end{cases}$$

9.5.4.2.6 Phép quy tắc hóa các trị số suy biến

Các thuật toán nghịch đảo được quy tắc hóa $(\cdot)^{inv}$ được sử dụng cho ma trận trị số suy biến chéo Λ được xác định như:

$$\Lambda^{inv} = \lambda_{i,j}^{-1} = \begin{cases} \frac{1}{\lambda_{i,i}} & , \text{neu } i = j \text{ va } abs(\lambda_{i,i}) \geq T_{reg}^{\Lambda}, \\ 0 & , \text{neu khong.} \end{cases}$$

Đại lượng vô hướng quy tắc hóa tương ứng T_{reg}^{Λ} được xác định sử dụng ngưỡng tuyệt đối T_{reg} và trị số tối đa là Λ như sau:

$$T_{reg}^{\Lambda} = \max_i (abs(\lambda_{i,i})) T_{reg}, \text{ với } T_{reg} = 10^{-2}.$$

Trong một số hình vẽ sau đây, tín hiệu riêng lẻ được thể hiện như thu được từ các bước xử lý khác nhau. Điều này được thực hiện cho sự hiểu tốt hơn về sáng chế và là một khả năng để nhận biết sáng chế, tức là, trích các tín hiệu riêng lẻ và thực hiện các bước xử lý trên các tín hiệu này hoặc các tín hiệu được xử lý.

Phương án khác tính toán tất cả các ma trận cần thiết và áp dụng chúng như bước cuối cùng cho tín hiệu âm thanh được mã hóa để thu được tín hiệu âm thanh được giải

mã. Điều này gồm có sự tính toán của các ma trận hác nhau và các tổ hợp tương ứng của chúng.

Phương án tổ hợp cả hai cách.

Fig.10 thể hiện theo sơ đồ thiết bị 10 để xử lý nhiều (ở đây trong ví dụ này là năm) đối tượng âm thanh đầu vào 111 để cung cấp phép biểu diễn của các đối tượng âm thanh đầu vào 111 bằng tín hiệu âm thanh được mã hóa 100.

Các đối tượng âm thanh đầu vào 111 được phân bố hoặc được trộn giảm thành các tín hiệu trộn giảm 101. Trong phương án được thể hiện, bốn trong số năm đối tượng âm thanh đầu vào 111 được gán cho hai tín hiệu trộn giảm 101. Một đối tượng âm thanh đầu vào 111 một mình được gán cho tín hiệu trộn giảm thứ ba 101. Do đó, năm đối tượng âm thanh đầu vào 111 được biểu diễn bởi ba tín hiệu trộn giảm 101.

Các tín hiệu trộn giảm 101 này sau đó - có khả năng theo sau một số bước xử lý không được thể hiện - được tổ hợp thành tín hiệu âm thanh được mã hóa 100.

Tín hiệu âm thanh được mã hóa 100 được nạp vào thiết bị theo sáng chế 1, mà một phương án được thể hiện trên Fig.11.

Từ tín hiệu âm thanh được mã hóa 100, ba tín hiệu trộn giảm 101 (so với Fig.10) được trích.

Các tín hiệu trộn giảm 101 được nhóm lại - trong ví dụ được thể hiện - thành hai nhóm tín hiệu trộn giảm 102.

Vì mỗi tín hiệu trộn giảm 101 được kết hợp với số đã cho gồm các đối tượng âm thanh đầu vào, mỗi nhóm tín hiệu trộn giảm 102 đề cập đến số đã cho gồm các đối tượng âm thanh đầu vào (sự thể hiện tương ứng là đối tượng đầu vào). Do đó, mỗi nhóm gồm các tín hiệu trộn giảm 102 được kết hợp với tập hợp các đối tượng âm thanh đầu vào trong số nhiều đối tượng âm thanh đầu vào mà được mã hóa bởi tín hiệu âm thanh được mã hóa 100 (so với Fig.10).

Việc tạo nhóm xảy ra trong phương án được thể hiện dưới các sự co hẹp:

1. Mỗi đối tượng âm thanh đầu vào 111 thuộc về chỉ một tập hợp các đối tượng âm thanh đầu vào và, do đó, đến một nhóm gồm các tín hiệu trộn giảm 102.

2. Mỗi đối tượng âm thanh đầu vào 111 không có hệ thức nào được tạo tín hiệu trong tín hiệu âm thanh được mã hóa đến đối tượng âm thanh đầu vào 111 thuộc về tập hợp khác được kết hợp với nhóm khác gồm các tín hiệu trộn giảm. Điều này có nghĩa là tín hiệu âm thanh được mã hóa không có thông tin như vậy mà do tiêu chuẩn sẽ dẫn đến phép tính được tổ hợp của các đối tượng âm thanh đầu vào tương ứng.

3. Số lượng các tín hiệu trộn giảm 101 nằm trong các nhóm tương ứng 102 được tối thiểu hóa.

(Ở đây, hai) nhóm các tín hiệu trộn giảm 102 được xử lý một cách riêng lẻ sau đây để thu được năm tín hiệu âm thanh đầu ra 103 tương ứng với năm đối tượng âm thanh đầu vào 111.

Một nhóm các tín hiệu trộn giảm 102 mà được kết hợp với hai tín hiệu trộn giảm 101 bao phủ hai cặp đối tượng âm thanh đầu vào 111 (so với Fig.10) cho phép thu được bốn tín hiệu âm thanh đầu ra 103.

Nhóm khác gồm các tín hiệu trộn giảm 102 dẫn đến một tín hiệu đầu ra 103 như tín hiệu trộn giảm đơn 101 hoặc nhóm gồm các tín hiệu trộn giảm 102 này (hoặc chính xác hơn: nhóm gồm một tín hiệu trộn giảm) đề cập đến một đối tượng âm thanh đầu vào 111 (so với Fig.10).

Năm tín hiệu âm thanh đầu ra 103 được tổ hợp thành một tín hiệu âm thanh được giải mã 110 như đầu ra của thiết bị 1.

Trong một phương án của Fig.11, tất cả các bước được thực hiện một cách riêng lẻ trên các nhóm các tín hiệu trộn giảm 102.

Phương án của thiết bị 1 được thể hiện trên Fig.12 có thể nhận ở đây tín hiệu âm thanh được mã hóa giống nhau 100 như thiết bị 1 được thể hiện trên Fig.11 và thu được bởi thiết bị 10 như được thể hiện trên Fig.10.

Từ tín hiệu âm thanh được mã hóa 100, ba tín hiệu trộn giảm 101 (đối với ba kênh vận chuyển) thu được và được nhóm thành hai nhóm các tín hiệu trộn giảm 102. Các nhóm 102 này được xử lý một cách riêng lẻ để thu được năm tín hiệu được xử lý 104 tương ứng với năm đối tượng âm thanh đầu vào được thể hiện trên Fig.10.

Trong các bước dưới đây, từ năm tín hiệu được xử lý 104 cùng tám tín hiệu âm

thanh đầu ra 103 thu được, ví dụ, được kết xuất để được sử dụng cho tám kênh đầu ra. Các tín hiệu âm thanh đầu ra 103 được tổ hợp thành tín hiệu âm thanh được giải mã 110 mà được phát ra từ thiết bị 1. Trong phương án này, sự xử lý riêng lẻ cũng như cùng nhau được thực hiện trên các nhóm gồm các tín hiệu trộn giảm 102.

Fig.13 thể hiện một số bước phương án của phương pháp theo sáng chế trong đó tín hiệu âm thanh được mã hóa được giải mã.

Trong bước 200, các tín hiệu trộn giảm được trích từ tín hiệu âm thanh được mã hóa. Trong bước 201 sau đây, các tín hiệu trộn giảm được phân bổ đến nhóm các tín hiệu trộn giảm.

Trong bước 202, mỗi nhóm gồm các tín hiệu trộn giảm được xử lý một cách riêng lẻ để cung cấp các kết quả nhóm riêng lẻ. Sự xử lý riêng lẻ các nhóm bao gồm ít nhất sự không trộn để thu được các phép biểu diễn của các tín hiệu âm thanh mà đã được tổ hợp qua sự trộn giảm của các đối tượng âm thanh đầu vào trong quy trình mã hóa. Trong một phương án - không được thể hiện ở đây - việc xử lý riêng lẻ được theo sau bởi việc xử lý cùng nhau.

Trong bước 203, các kết quả nhóm được tổ hợp thành tín hiệu âm thanh được giải mã sẽ được phát ra.

Fig.14 một lần nữa thể hiện phương án của thiết bị 1, trong đó tất cả các bước xử lý theo sau việc tạo nhóm các tín hiệu trộn giảm 101 của tín hiệu âm thanh được mã hóa 100 thành các nhóm gồm các tín hiệu trộn giảm 102 được thực hiện một cách riêng lẻ. Thiết bị 1 mà nhận tín hiệu âm thanh được mã hóa 100 với các tín hiệu trộn giảm 101 bao gồm bộ tạo nhóm 2 mà nhóm các tín hiệu trộn giảm 101 để cung cấp các nhóm các tín hiệu trộn giảm 102. Các nhóm gồm các tín hiệu trộn giảm 102 được xử lý bởi bộ xử lý 3 thực hiện tất cả các bước cần thiết một cách riêng lẻ trên mỗi nhóm gồm các tín hiệu trộn giảm 102. Các kết quả nhóm riêng lẻ của việc xử lý các nhóm gồm các tín hiệu trộn giảm 102 là các tín hiệu âm thanh đầu ra 103 mà được tổ hợp bởi bộ tổ hợp 3 để thu được tín hiệu âm thanh được giải mã 110 sẽ được phát ra bởi thiết bị 1.

Thiết bị 1 được thể hiện trên Fig.15 khác với phương án được thể hiện trên

Fig.14 theo sau việc tạo nhóm các tín hiệu trộn giảm 101. Trong ví dụ, không phải tất cả các bước xử lý được thực hiện một cách riêng lẻ trên các nhóm các tín hiệu trộn giảm 102 nhưng một số bước được thực hiện cùng nhau, do đó xem xét nhiều hơn một nhóm các tín hiệu trộn giảm 102.

Do điều này, bộ xử lý 3 trong phương án này được tạo cấu hình để thực hiện chỉ một số hoặc ít nhất một bước xử lý một cách riêng lẻ. Kết quả của việc xử lý là các tín hiệu được xử lý 104 mà được xử lý cùng nhau bởi bộ xử lý sau 5. Các tín hiệu âm thanh đầu ra thu được 103 cuối cùng được tổ hợp bởi bộ tổ hợp 4 dẫn đến tín hiệu âm thanh được giải mã 110.

Trên Fig.16, bộ xử lý 3 được thể hiện dạng sơ đồ nhận các nhóm gồm các tín hiệu trộn giảm 102 và cung cấp các tín hiệu âm thanh đầu ra 103.

Bộ xử lý 3 bao gồm bộ không trộn 300 được tạo cấu hình để không trộn các tín hiệu trộn giảm 101 của các nhóm tương ứng gồm các tín hiệu trộn giảm 102. Do đó, bộ không trộn 300, khôi phục các đối tượng âm thanh đầu vào riêng lẻ mà đã được tổ hợp bởi bộ mã hóa thành các tín hiệu trộn giảm tương ứng 101.

Các đối tượng âm thanh đầu vào được khôi phục hoặc được tách được chịu sự kiểm tra bộ kết xuất 302. Bộ kết xuất 302 được tạo cấu hình để kết xuất các tín hiệu trộn giảm không được trộn của các nhóm tương ứng cho tình huống đầu ra của tín hiệu âm thanh được giải mã 110 đã nêu để cung cấp các tín hiệu được kết xuất 112. Do đó, các tín hiệu được kết xuất 112, được thích ứng cho loại cảnh phát lại của tín hiệu âm thanh được giải mã. Sự kết xuất phụ thuộc, ví dụ, vào số lượng các loa phát thanh sẽ được sử dụng, đến sự sắp xếp của chúng hoặc loại ảnh hưởng sẽ thu được bằng cách phát tín hiệu âm thanh được giải mã.

Các tín hiệu được kết xuất 112, Y_{dry} , còn được đưa đến bộ trộn sau 303 được tạo cấu hình để thực hiện ít nhất một bước giải tương quan trên các tín hiệu được kết xuất 112 đã nêu và được tạo cấu hình để tổ hợp các kết quả Y_{wet} của bước giải tương quan được thực hiện với các tín hiệu được kết xuất 112 tương ứng đã nêu, Y_{dry} . Do đó, bộ trộn sau 303, thực hiện bước giải tương quan các tín hiệu mà được tổ hợp trong một tín hiệu trộn giảm.

Các tín hiệu âm thanh đầu ra thu được 103 cuối cùng được đưa tới bộ tổ hợp như được thể hiện ở trên.

Đối với các bước, bộ xử lý 3 phụ thuộc vào bộ tính toán 301, mà ở đây tách biệt với các bộ phận khác của bộ xử lý 3 nhưng đó là phương án thay thế - không được thể hiện - đặc tính lần lượt của bộ tạo nhóm 300, bộ kết xuất 302, và bộ trộn sau 303.

Liên quan là thực tế, rằng các ma trận cần thiết, các trị số, v.v được tính toán một cách riêng lẻ đối với các nhóm tương ứng của các tín hiệu trộn giảm 102. Điều này ám chỉ rằng, ví dụ, các ma trận sẽ được tính toán là nhỏ hơn các ma trận được sử dụng trong tình trạng kỹ thuật. Các ma trận có các kích thước phụ thuộc vào số lượng các đối tượng âm thanh đầu vào của tập hợp tương ứng gồm các đối tượng âm thanh đầu vào được kết hợp với các nhóm gồm các đối tượng trộn giảm và/hoặc số lượng các tín hiệu trộn giảm thuộc về nhóm tương ứng gồm các tín hiệu trộn giảm.

Trong tình trạng kỹ thuật, ma trận sẽ được sử dụng cho việc không trộn có kích thước là số lượng các đối tượng âm thanh đầu vào hoặc các tín hiệu âm thanh đầu vào nhân với số lượng này. Sáng chế cho phép tính toán ma trận nhỏ hơn với kích thước phụ thuộc vào số lượng các tín hiệu âm thanh đầu vào thuộc về nhóm tương ứng gồm các tín hiệu trộn giảm.

Trên Fig.17, mục đích kết xuất được giải thích.

Thiết bị 1 nhận tín hiệu âm thanh được mã hóa 100 và giải mã nó cung cấp tín hiệu âm thanh được giải mã 110.

Tín hiệu âm thanh được giải mã 110 được phát trong tình huống đầu ra cụ thể hoặc cảnh đầu ra 400. Tín hiệu âm thanh được giải mã 110 là trong ví dụ sẽ được phát ra bởi năm loa phát thanh 401: Trái, Phải, Chính giữa, Bao quanh bên trái, và Bao quanh bên phải. Người nghe 402 ở giữa cảnh 400 đối mặt loa phát thanh chính giữa.

Bộ kết xuất trong thiết bị 1 phân phối các tín hiệu âm thanh được khôi phục sẽ được phát đến các loa phát thanh riêng 401 và, do đó, phân phối phép biểu diễn được khôi phục của các đối tượng âm thanh ban đầu như các nguồn của các tín hiệu âm thanh trong tình huống đầu ra 400 đã cho.

Do đó, sự kết xuất phụ thuộc vào loại tình huống đầu ra 400 và sở thích cá nhân

của người nghe 402.

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

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

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

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

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

Nói cách khác, do đó, phương án của phương pháp theo sáng chế là, chương trình

máy tính có mã chương trình để thực hiện một trong số các phương pháp được mô tả ở đây, khi chương trình máy tính chạy trên máy tính.

Phương án khác của các phương pháp theo sáng chế là, sóng mang dữ liệu (hoặc môi trường lưu trữ số, hoặc vật ghi đọc được bằng máy tính) gồm có, đã được ghi lại trên đó, chương trình máy tính để thực hiện một trong các phương pháp được mô tả ở đây. Vật mang dữ liệu, vật ghi lưu trữ kỹ thuật số hoặc môi trường được ghi lại dạng hữu hình điển hình và/hoặc không chuyển tiếp.

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

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

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

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

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

Thiết bị được mô tả ở đây có thể được thực hiện sử dụng thiết bị phần cứng, hoặc

sử dụng máy tính, hoặc sử dụng tổ hợp của thiết bị phần cứng và máy tính.

Các phương pháp được mô tả ở đây có thể được thực hiện sử dụng thiết bị phần cứng, hoặc sử dụng máy tính, hoặc sử dụng tổ hợp của thiết bị phần cứng và máy tính.

YÊU CẦU BẢO HỘ

1. Thiết bị (1) xử lý tín hiệu âm thanh được mã hóa (100) bao gồm nhiều tín hiệu trộn giảm (101) được kết hợp với nhiều đối tượng âm thanh đầu vào (111) và các tham số đối tượng (E),

thiết bị bao gồm:

bộ tạo nhóm (2) được tạo cấu hình để tạo nhóm nhiều tín hiệu trộn giảm (101) đã nêu thành nhiều nhóm gồm các tín hiệu trộn giảm (102) dựa trên thông tin nằm trong tín hiệu âm thanh được mã hóa (100), trong đó mỗi nhóm gồm các tín hiệu trộn giảm (102) được kết hợp với tập hợp các đối tượng âm thanh đầu vào (111) trong số nhiều đối tượng âm thanh đầu vào (111) đã nêu,

bộ xử lý (3) được tạo cấu hình để thực hiện ít nhất một bước xử lý riêng lẻ trên các tham số đối tượng (E_k) của mỗi tập hợp các đối tượng âm thanh đầu vào (111) để cung cấp các kết quả nhóm (103, 104), và

bộ tổ hợp (4) được tạo cấu hình để tổ hợp các kết quả nhóm (103, 104) đã nêu để cung cấp tín hiệu âm thanh được giải mã (110),

trong đó bộ tạo nhóm (2) đã nêu được tạo cấu hình để tạo nhóm nhiều tín hiệu trộn giảm (101) đã nêu thành nhiều nhóm các tín hiệu trộn giảm (102) để mỗi đối tượng âm thanh đầu vào (111) trong số nhiều đối tượng âm thanh đầu vào (111) đã nêu thuộc về chỉ một tập hợp các đối tượng âm thanh đầu vào (111), và

trong đó bộ tạo nhóm (2) đã nêu được tạo cấu hình để tạo nhóm nhiều tín hiệu trộn giảm (101) đã nêu thành nhiều nhóm các tín hiệu trộn giảm (102) đã nêu để mỗi đối tượng âm thanh đầu vào (111) trong số mỗi tập hợp các đối tượng âm thanh đầu vào hoặc không có mối quan hệ được tạo tín hiệu trong tín hiệu âm thanh được mã hóa (100) với các đối tượng âm thanh đầu vào (111) khác hoặc có mối quan hệ được tạo tín hiệu trong tín hiệu âm thanh được mã hóa (100) chỉ với ít nhất một đối tượng âm thanh đầu vào (111) thuộc về cùng một tập hợp các đối tượng âm thanh đầu vào (111).

2. Thiết bị (1) theo điểm 1, trong đó bộ tạo nhóm (2) đã nêu được tạo cấu hình để tạo nhóm nhiều tín hiệu trộn giảm (101) đã nêu thành nhiều nhóm các tín hiệu trộn giảm (102) đã nêu trong khi giảm thiểu hóa số lượng các tín hiệu trộn giảm (101) nằm

trong mỗi nhóm gồm các tín hiệu trộn giảm (102).

3. Thiết bị (1) theo điểm 1 hoặc 2, trong đó bộ tạo nhóm (2) đã nêu được tạo cấu hình để tạo nhóm nhiều tín hiệu trộn giảm (101) đã nêu thành nhiều nhóm các tín hiệu trộn giảm (102) đã nêu để chỉ một tín hiệu trộn giảm đơn (101) thuộc về một nhóm các tín hiệu trộn giảm (102).

4. Thiết bị (1) theo một điểm trong số các điểm từ 1 đến 3,

trong đó bộ tạo nhóm (2) đã nêu được tạo cấu hình để tạo nhóm nhiều tín hiệu trộn giảm (101) đã nêu thành nhiều nhóm các tín hiệu trộn giảm (102) đã nêu bằng cách áp dụng ít nhất các bước sau:

phát hiện liệu tín hiệu trộn giảm (101) được gán cho nhóm các tín hiệu trộn giảm (102) hiện có hay không;

phát hiện liệu ít nhất một đối tượng âm thanh đầu vào (111) trong số nhiều đối tượng âm thanh đầu vào (111) được kết hợp với tín hiệu trộn giảm (101) là phần của tập hợp các đối tượng âm thanh đầu vào (111) được kết hợp với nhóm các tín hiệu trộn giảm (102) hiện có hay không;

gán tín hiệu trộn giảm (101) vào nhóm mới gồm các tín hiệu trộn giảm (102) trong trường hợp tín hiệu trộn giảm (101) không bị gán vào nhóm các tín hiệu trộn giảm (102) hiện có và trong trường hợp tất cả các đối tượng âm thanh đầu vào (111) trong số nhiều đối tượng âm thanh đầu vào được kết hợp với tín hiệu trộn giảm (101) không có sự kết hợp với nhóm các tín hiệu trộn giảm (102) hiện có; và

tổ hợp tín hiệu trộn giảm (101) với nhóm các tín hiệu trộn giảm (102) hiện có hoặc trong trường hợp tín hiệu trộn giảm (101) được gán cho nhóm các tín hiệu trộn giảm (102) hiện có hoặc trong trường hợp ít nhất một đối tượng âm thanh đầu vào (111) trong số nhiều đối tượng âm thanh đầu vào được kết hợp với tín hiệu trộn giảm (101) được kết hợp với nhóm các tín hiệu trộn giảm (102) hiện có.

5. Thiết bị (1) theo một điểm trong số các điểm từ 1 đến 4,

trong đó bộ xử lý (3) đã nêu được tạo cấu hình để thực hiện nhiều bước xử lý riêng lẻ trên các tham số đối tượng (E_k) của mỗi tập hợp các đối tượng âm thanh đầu vào (111) để cung cấp các ma trận riêng lẻ như các kết quả nhóm, và

trong đó bộ tổ hợp (4) đã nêu được tạo cấu hình để tổ hợp các ma trận riêng lẻ đã nêu.

6. Thiết bị (1) theo một điểm trong số các điểm từ 1 đến 5,

trong đó bộ xử lý (3) đã nêu được tạo cấu hình để thực hiện ít nhất một bước xử lý riêng lẻ trên các tham số đối tượng (E_k) của mỗi tập hợp các đối tượng âm thanh đầu vào (111) để cung cấp các ma trận riêng lẻ,

trong đó thiết bị (1) đã nêu bao gồm bộ xử lý sau (5) được tạo cấu hình để xử lý đồng thời các tham số đối tượng để cung cấp ít nhất một ma trận toàn thể, và

trong đó bộ tổ hợp (4) đã nêu được tạo cấu hình để tổ hợp các ma trận riêng lẻ đã nêu và ít nhất một ma trận toàn thể đã nêu.

7. Thiết bị (1) theo một điểm trong số các điểm từ 1 đến 6,

trong đó bộ xử lý (3) đã nêu bao gồm bộ tính toán (301) được tạo cấu hình để tính toán riêng lẻ cho mỗi nhóm tín hiệu trộn giảm (102) các ma trận với các kích thước phụ thuộc vào ít nhất một số lượng các đối tượng âm thanh đầu vào của tập hợp các đối tượng âm thanh đầu vào được kết hợp với nhóm tương ứng gồm các tín hiệu trộn giảm (102) và số lượng các tín hiệu trộn giảm (101) thuộc về nhóm tương ứng gồm các tín hiệu trộn giảm (102).

8. Thiết bị (1) theo một điểm trong số các điểm từ 1 đến 7,

trong đó bộ xử lý (3) được tạo cấu hình để tính toán cho mỗi nhóm các tín hiệu trộn giảm ngưỡng riêng lẻ dựa trên trị số năng lượng tối đa nằm trong nhóm tương ứng gồm các tín hiệu trộn giảm (102).

9. Thiết bị (1) theo một điểm trong số các điểm từ 1 đến 8,

trong đó bộ xử lý (3) đã nêu được tạo cấu hình để xác định ma trận trộn giảm riêng lẻ (D_k) cho mỗi nhóm các tín hiệu trộn giảm (102),

trong đó bộ xử lý (3) đã nêu được tạo cấu hình để xác định ma trận hiệp phương sai nhóm riêng lẻ (E_k) cho mỗi nhóm các tín hiệu trộn giảm (102),

trong đó bộ xử lý (3) đã nêu được tạo cấu hình để xác định ma trận hiệp phương sai nhóm riêng lẻ (Δ_k) cho mỗi nhóm các tín hiệu trộn giảm (102) dựa trên ma trận trộn

giảm riêng lẻ ($\mathbf{D}_k$) và ma trận hiệp phương sai nhóm riêng lẻ ($\mathbf{E}_k$), và

trong đó bộ xử lý (3) đã nêu được tạo cấu hình để xác định ma trận nhóm nghịch đảo được quy tắc hóa riêng lẻ ($\mathbf{J}_k$) cho mỗi nhóm các tín hiệu trộn giảm (102).

10. Thiết bị (1) theo điểm 9,

trong đó bộ tổ hợp (4) đã nêu được tạo cấu hình để tổ hợp các ma trận nhóm nghịch đảo được quy tắc hóa riêng lẻ ($\mathbf{J}_k$) để thu được ma trận nhóm nghịch đảo được quy tắc hóa toàn thể ($\mathbf{J}$).

11. Thiết bị (1) theo điểm 9 hoặc 10,

trong đó bộ xử lý đã nêu (3) được tạo cấu hình để xác định ma trận không trộn theo tham số nhóm riêng lẻ ($\mathbf{U}_k$) cho mỗi nhóm gồm các tín hiệu trộn giảm (102) dựa trên ma trận trộn giảm riêng lẻ ($\mathbf{D}_k$), ma trận hiệp phương sai nhóm riêng lẻ ($\mathbf{E}_k$), và ma trận nhóm nghịch đảo được quy tắc hóa riêng lẻ ($\mathbf{J}_k$), và

trong đó bộ tổ hợp (4) đã nêu được tạo cấu hình để tổ hợp ma trận không trộn theo tham số nhóm riêng lẻ ($\mathbf{U}_k$) để thu được ma trận không trộn theo tham số nhóm toàn thể ($\mathbf{U}$).

12. Thiết bị (1) theo điểm 11,

trong đó bộ xử lý đã nêu (3) được tạo cấu hình để xác định ma trận không trộn theo tham số nhóm riêng lẻ ($\mathbf{U}_k$) cho mỗi nhóm gồm các tín hiệu trộn giảm (102) dựa trên ma trận trộn giảm riêng lẻ ($\mathbf{D}_k$), ma trận hiệp phương sai nhóm riêng lẻ ($\mathbf{E}_k$), và ma trận nhóm nghịch đảo được quy tắc hóa riêng lẻ ($\mathbf{J}_k$), và

trong đó bộ tổ hợp (4) đã nêu được tạo cấu hình để tổ hợp ma trận không trộn theo tham số nhóm riêng lẻ ($\mathbf{U}_k$) để thu được ma trận không trộn theo tham số nhóm toàn thể ($\mathbf{U}$).

13. Thiết bị (1) theo một điểm trong số các điểm từ 1 đến 12,

trong đó bộ xử lý (3) đã nêu được tạo cấu hình để xác định ma trận kết xuất nhóm riêng lẻ ($\mathbf{R}_k$) cho mỗi nhóm các tín hiệu trộn giảm (102).

14. Thiết bị (1) theo điểm 13,

trong đó bộ xử lý (3) đã nêu được tạo cấu hình để xác định ma trận trộn tăng

riêng lẻ ($\mathbf{R}_k \mathbf{U}_k$) cho mỗi nhóm các tín hiệu trộn giảm (102) dựa trên ma trận kết xuất nhóm riêng lẻ ($\mathbf{R}_k$) và ma trận không trộn theo tham số nhóm riêng lẻ ($\mathbf{U}_k$), và

trong đó bộ tổ hợp (4) đã nêu được tạo cấu hình để tổ hợp các ma trận trộn tăng riêng lẻ ($\mathbf{R}_k \mathbf{U}_k$) để thu được ma trận trộn tăng toàn thể ($\mathbf{RU}$).

15. Thiết bị (1) theo điểm 13 hoặc 14,

trong đó bộ xử lý (3) đã nêu được tạo cấu hình để xác định ma trận hiệp phương sai nhóm riêng lẻ ($\mathbf{C}_k$) cho mỗi nhóm gồm các tín hiệu trộn giảm (102) dựa trên ma trận kết xuất nhóm riêng lẻ ($\mathbf{R}_k$) và ma trận hiệp phương sai nhóm riêng lẻ ($\mathbf{E}_k$), và

trong đó bộ tổ hợp (4) đã nêu được tạo cấu hình để tổ hợp các ma trận hiệp phương sai nhóm riêng lẻ ($\mathbf{C}_k$) để thu được ma trận hiệp phương sai nhóm toàn thể ($\mathbf{C}$).

16. Thiết bị (1) theo một điểm trong số các điểm từ 13 đến 15,

trong đó bộ xử lý (3) đã nêu được tạo cấu hình để xác định ma trận hiệp phương sai nhóm riêng lẻ của tín hiệu được ước lượng theo tham số ($\mathbf{E}_y^{\text{dry}})_k$ dựa trên ma trận kết xuất nhóm riêng lẻ ($\mathbf{R}_k$), ma trận không trộn theo tham số nhóm riêng lẻ ($\mathbf{U}_k$), ma trận trộn giảm riêng lẻ ($\mathbf{D}_k$), và ma trận hiệp phương sai nhóm riêng lẻ ($\mathbf{E}_k$), và

trong đó bộ tổ hợp (4) đã nêu được tạo cấu hình để tổ hợp các ma trận hiệp phương sai nhóm riêng lẻ của tín hiệu được ước lượng theo tham số ($\mathbf{E}_y^{\text{dry}})_k$ để thu được tín hiệu được ước lượng theo tham số toàn thể $\mathbf{E}_y^{\text{dry}}$.

17. Thiết bị (1) theo một điểm trong số các điểm từ 1 đến 16,

trong đó bộ xử lý (3) đã nêu được tạo cấu hình để xác định ma trận nghịch đảo được quy tắc hóa ($\mathbf{J}$) dựa trên sự khai triển trị số suy biến của ma trận hiệp phương sai trộn giảm ($\mathbf{E}_{\text{DMX}}$).

18. Thiết bị (1) theo một điểm trong số các điểm từ 1 đến 17,

trong đó bộ xử lý (3) đã nêu được tạo cấu hình để xác định cho sự xác định ma trận không trộn theo tham số ($\mathbf{U}$) ma trận con (Δ_k) bằng cách lựa chọn các phần tử ($\Delta(m, n)$) tương ứng với các tín hiệu trộn giảm (m, n) được gán cho nhóm (k) tương ứng gồm các tín hiệu trộn giảm (102).

19. Thiết bị (1) theo một điểm bất kỳ trong số các điểm nêu trên,

trong đó bộ tổ hợp (4) đã nêu được tạo cấu hình để xác định ma trận trộn sau (P) dựa trên các ma trận được xác định riêng lẻ cho mỗi nhóm các tín hiệu trộn giảm (102) và

trong đó bộ tổ hợp (4) đã nêu được tạo cấu hình để áp dụng ma trận trộn sau (P) cho nhiều tín hiệu trộn giảm (101) để thu được tín hiệu âm thanh được giải mã (110).

20. Phương pháp xử lý tín hiệu âm thanh được mã hóa (100) bao gồm nhiều tín hiệu trộn giảm (101) được kết hợp với nhiều đối tượng âm thanh đầu vào (111) và các tham số đối tượng (E),

phương pháp đã nêu bao gồm:

tạo nhóm nhiều tín hiệu trộn giảm (101) đã nêu thành nhiều nhóm gồm các tín hiệu trộn giảm (102) dựa trên thông tin nằm trong tín hiệu âm thanh được mã hóa (100) đã nêu, trong đó mỗi nhóm gồm các tín hiệu trộn giảm (102) được kết hợp với tập hợp các đối tượng âm thanh đầu vào (111) trong số nhiều đối tượng âm thanh đầu vào (111) đã nêu,

thực hiện ít nhất một bước xử lý riêng lẻ trên các tham số đối tượng (E_k) của mỗi tập hợp các đối tượng âm thanh đầu vào (111) để cung cấp các kết quả nhóm (103, 104), và

tổ hợp các kết quả nhóm (103, 104) đã nêu để cung cấp tín hiệu âm thanh được giải mã (110),

trong đó tạo nhóm nhiều tín hiệu trộn giảm (101) đã nêu thành nhiều nhóm các tín hiệu trộn giảm (102) để mỗi đối tượng âm thanh đầu vào (111) trong số nhiều đối tượng âm thanh đầu vào (111) đã nêu thuộc về chỉ một tập hợp các đối tượng âm thanh đầu vào (111), và

trong đó tạo nhóm nhiều tín hiệu trộn giảm (101) đã nêu thành nhiều nhóm các tín hiệu trộn giảm (102) đã nêu để mỗi đối tượng âm thanh đầu vào (111) trong số mỗi tập hợp các đối tượng âm thanh đầu vào hoặc không có mối quan hệ được tạo tín hiệu trong tín hiệu âm thanh được mã hóa (100) với các đối tượng âm thanh đầu vào (111) khác hoặc có mối quan hệ được tạo tín hiệu trong tín hiệu âm thanh được mã hóa (100)

chỉ với ít nhất một đối tượng âm thanh đầu vào (111) thuộc về cùng một tập hợp các đối tượng âm thanh đầu vào (111).

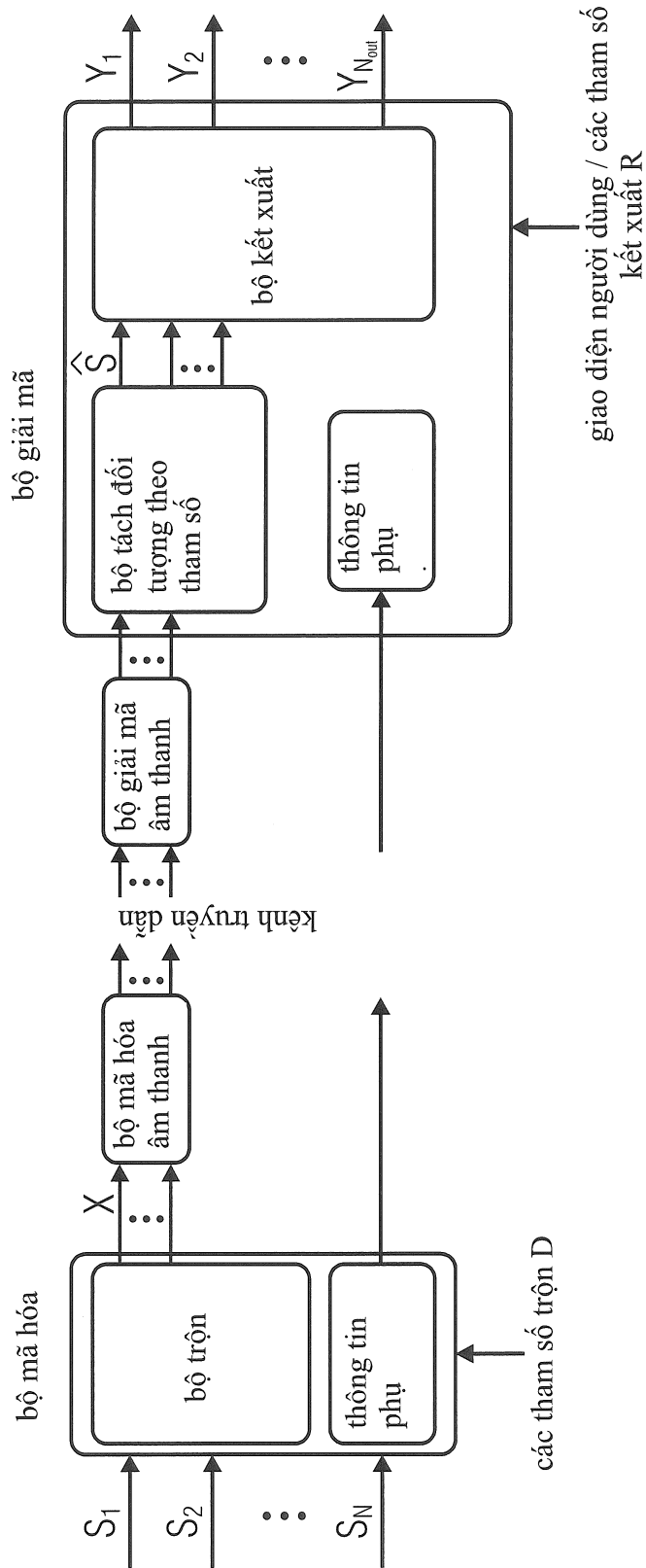


Fig. 1

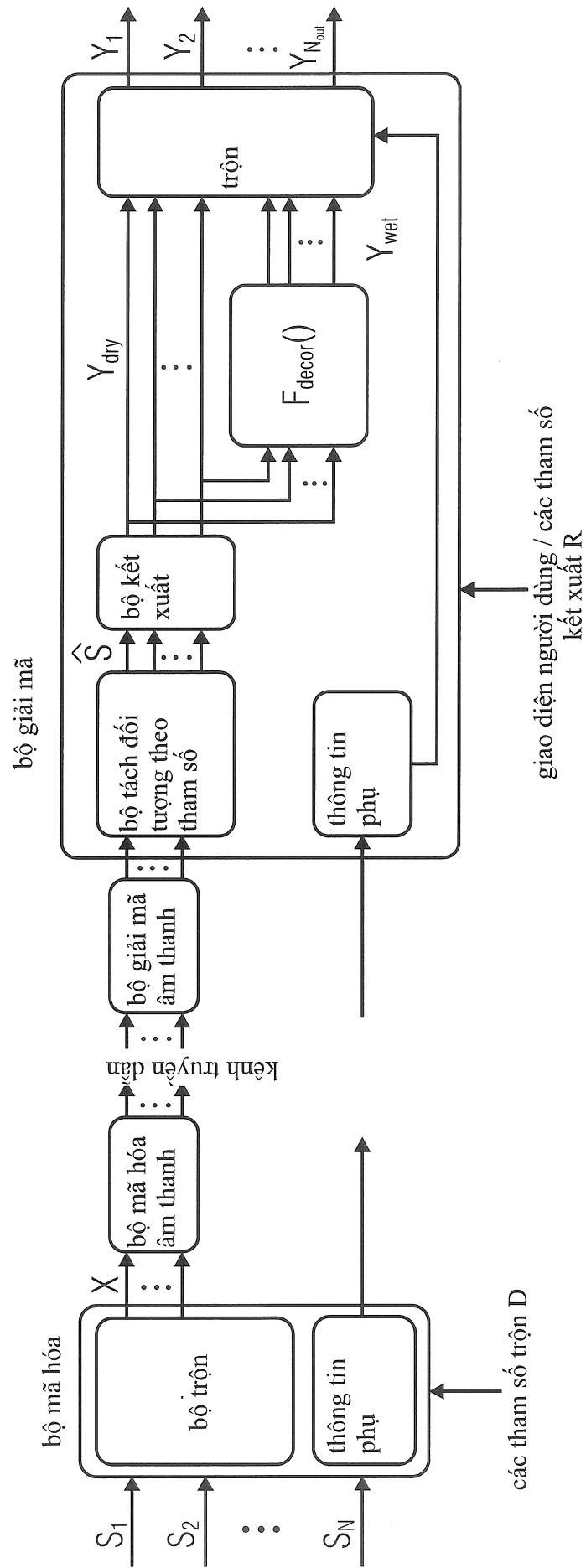


Fig. 2

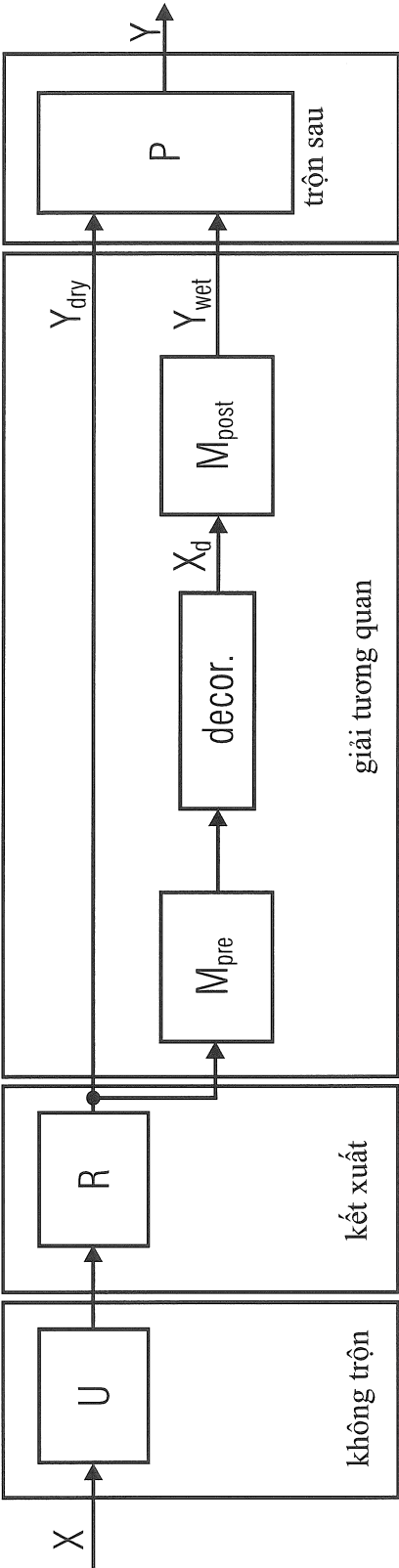


Fig. 3

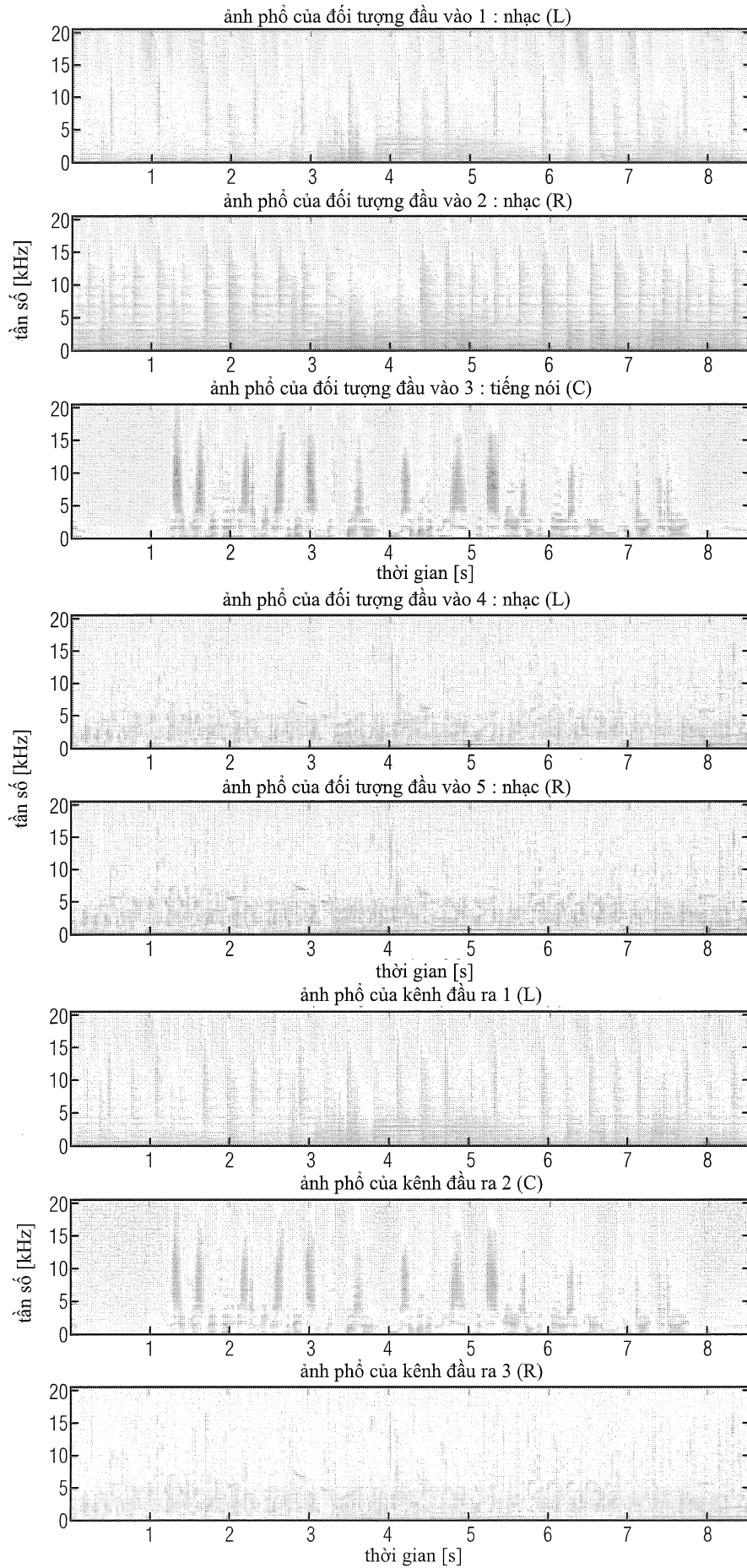


Fig. 4

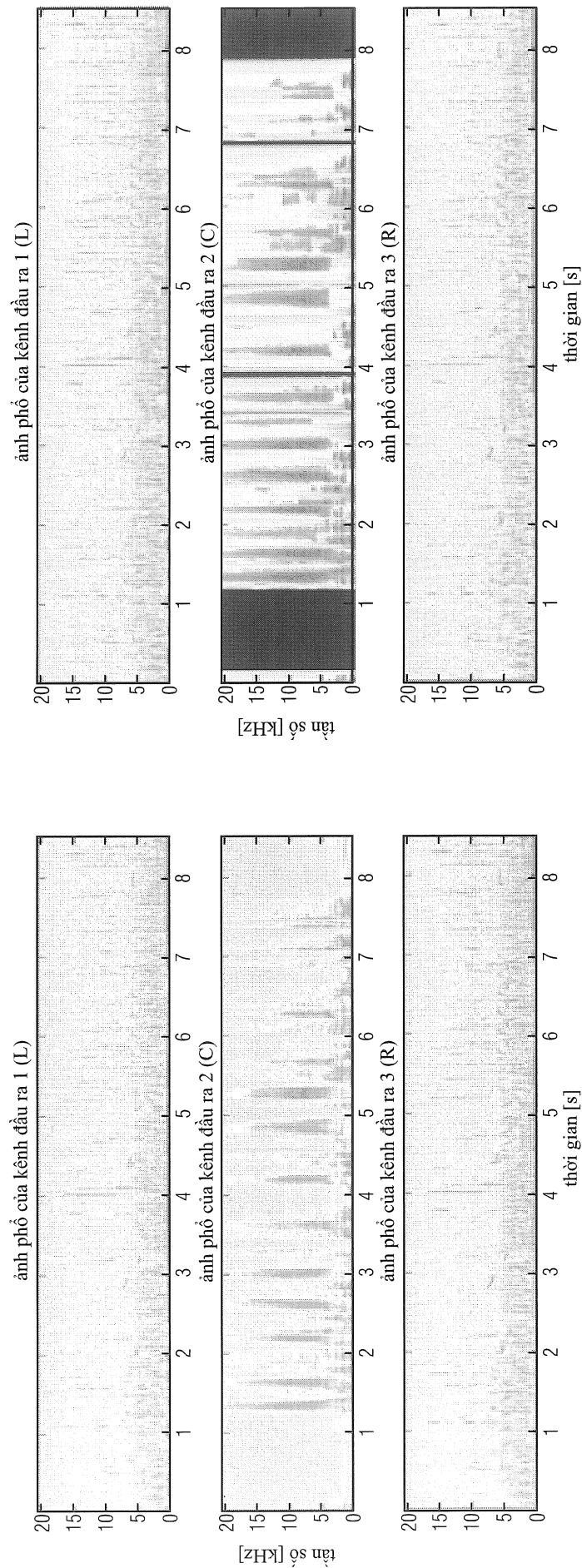


Fig. 5

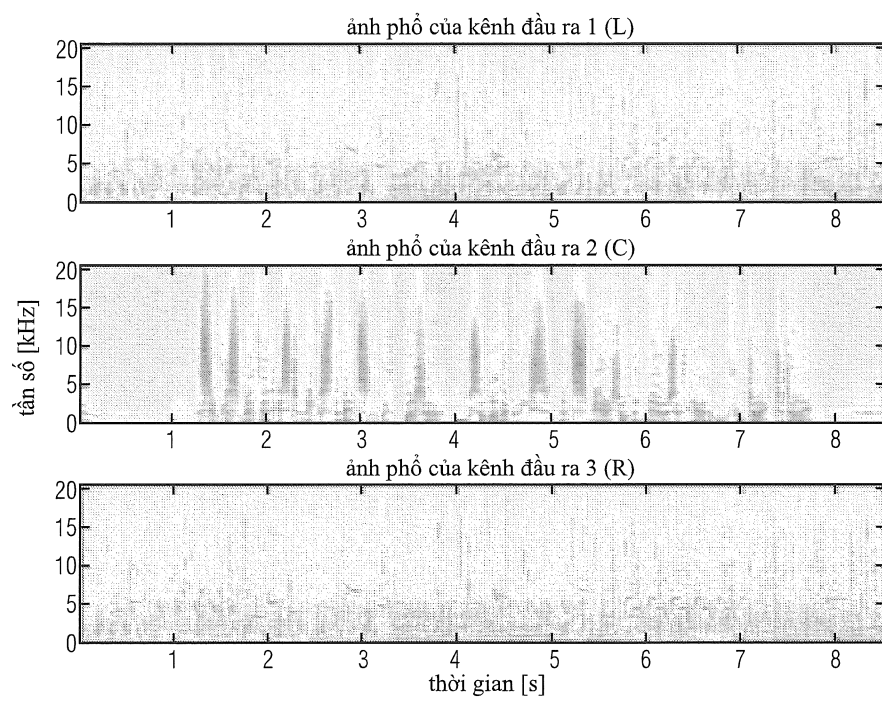


Fig. 6

$e_{i,j} = \sqrt{OLD_i OLD_j} OC_{i,j}$ $\Delta = DED^*$ $[V, \Lambda] = svd(\Delta)$ $J = V \Lambda_{inv}^* V^*$ $U = ED^* J$ không trộn	RU kết xuất	$C = RER^*$ $E_Y^{dry} = RU(DED^*)U^*R^*$ trộn sau $P = \text{function}(E_Y^{dry}, C)$
--	---------------------------	---

Fig. 7

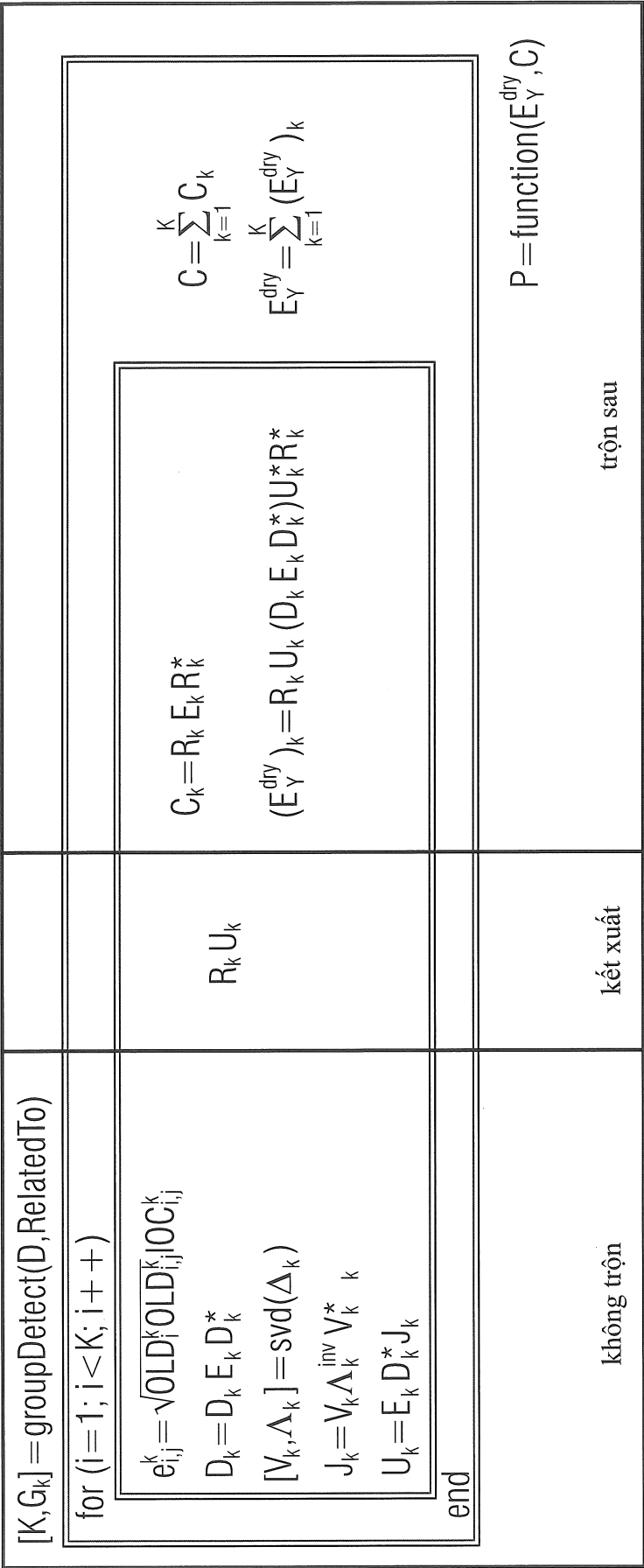


Fig. 8

```

#define SAOC_MAX_OBJECTS          128
#define SAOC_MAX_DMx_CHANNELS    32

#define SAOC_MAX_GROUP_OBJECTS    32 /* Maximum number of objects in one group */

#define SAOC_MAX_GROUP_DMx_CHANNELS 16 /* Maximum number of downmix channels in one group */

typedef struct saocCoreGroupList {
    int numDmx;
    int numObj;
    int dmxList[SAOC_MAX_DMx_CHANNELS];
    int objList[SAOC_MAX_OBJECTS];
} saocCoreGroupList;

typedef struct saocCoreGroup {
    int numGroups;
    saocCoreGroupList saocGroup[SAOC_MAX_DMx_CHANNELS];
} saocCoreGroup;

static int getSaocCoreGroups(int CrossCov[SAOC_MAX_DMx_CHANNELS][SAOC_MAX_OBJECTS],
                             int numObjects,
                             int numDmxChannels,
                             saocCoreGroup *groupInfo)
{
    int iGr, iObj, jObj, iDmx, jDmx, indexObj, indexChan, groupIndex, foundGroup;

    int channelsFlag[SAOC_MAX_DMx_CHANNELS];
    int objectsFlag[SAOC_MAX_OBJECTS];

    /* Init empty grouping info */
    for (iGr = 0; iGr < SAOC_MAX_DMx_CHANNELS; iGr++) {
        for (iObj = 0; iObj < numObjects; iObj++){
            groupInfo->saocGroup[iGr].objList[iObj] = -1;
        }
        for (iDmx = 0; iDmx < SAOC_MAX_DMx_CHANNELS; iDmx++) {
            groupInfo->saocGroup[iGr].dmxList[iDmx] = -1;
        }
        groupInfo->saocGroup[iGr].numDmx = 0;
        groupInfo->saocGroup[iGr].numObj = 0;
    }
    groupInfo->numGroups = 0;

    for (iObj = 0; iObj < numObjects; iObj++){
        objectsFlag[iObj] = -1;
    }
    for (jDmx = 0; jDmx < numDmxChannels; jDmx++) {
        channelsFlag[jDmx] = jDmx;
    }
}

```

Fig. 9 (1)

```

for (iDmx = 0; iDmx < numDmxChannels; iDmx++) {
    for (iObj = 0; iObj < numObjects; iObj++) {
        if (CrossCov[iDmx][iObj] > 0) {
            if (objectsFlag[iObj] == -1) {
                objectsFlag[iObj] = iDmx;
            } else {
                foundGroup = objectsFlag[iObj];
                for (jObj = 0; jObj < numObjects; jObj++) {
                    if (objectsFlag[jObj] == foundGroup) {
                        objectsFlag[jObj] = iDmx;
                    }
                }
            }
        }
        for (jDmx = 0; jDmx < iDmx; jDmx++) {
            if (channelsFlag[jDmx] == foundGroup) {
                channelsFlag[jDmx] = iDmx;
            }
        }
    }
}

for (iObj = 0; iObj < numObjects; iObj++) {
    iGr = groupInfo->numGroups;
    if (objectsFlag[iObj] > -1) {
        groupIndex = objectsFlag[iObj];
        for (jObj = 0; jObj < numObjects; jObj++) {
            if (objectsFlag[jObj] == groupIndex) {
                indexObj = groupInfo->saocGroup[iGr].numObj++;
                groupInfo->saocGroup[iGr].objList[indexObj] = jObj;
                objectsFlag[jObj] = -1;
            }
        }
        for (iDmx = 0; iDmx < numDmxChannels; iDmx++) {
            if (channelsFlag[iDmx] == groupIndex) {
                indexChan = groupInfo->saocGroup[iGr].numDmx++;
                groupInfo->saocGroup[iGr].dmxList[indexChan] = iDmx;
                channelsFlag[iDmx] = -1;
            }
        }
        groupInfo->numGroups++;
    }
}

```

/* search for not assigned objects */

/* iObj not assigned to any group */

/* assign iObj to iDmx group */

/* iObj already assigned to a group */

/* assign all objects from this existing group to the new one */

/* assign all channels from this existing group to the new one */

/* object not assigned to any group */

/* assign all objects from this group into the object list */

/* mark object as assigned to an object list */

/* assign all chnls from this group into the list */

/* mark channel as assigned to a channel list */

Fig. 9 (2)

```

for (iGr = 0; iGr < groupInfo->numGroups; iGr++) {
    assert(groupInfo->saocGroup[iGr].numDmx < SAOC_MAX_GROUP_DMx_CHANNELS);
    assert(groupInfo->saocGroup[iGr].numObj < SAOC_MAX_GROUP_OBJECTS);
}

return 0;
}

static void getSaocDmxPossitions(float Dmx[SAOC_MAX_DMx_CHANNELS][SAOC_MAX_OBJECTS],
                                int RelatedToMatrix[SAOC_MAX_OBJECTS][SAOC_MAX_OBJECTS],
                                int CrossCov[SAOC_MAX_DMx_CHANNELS][SAOC_MAX_OBJECTS],
                                int numObjects,
                                int firstObject,
                                int numDmxChannels,
                                int firstDmxChannel)
{
    int ch, dCh, obj, obj2, objIdx, objIdx2;
    int D[SAOC_MAX_DMx_CHANNELS][SAOC_MAX_OBJECTS] = {{0}};
    float minVal = (float)pow(10.0f, -150.0f/20.0f) + SAOC_EPSILON;

    for (ch=0; ch<numDmxChannels; ch++) {
        dCh = firstDmxChannel + ch;
        for (obj=0; obj<numObjects; obj++) {
            D[ch][obj] = (minVal < Dmx[dCh][firstObject + obj] ? 1 : 0);
        }
    }

    for (ch=0; ch<numDmxChannels; ch++) {
        dCh = firstDmxChannel + ch;
        for (obj=0; obj<numObjects; obj++) {
            objIdx = firstObject + obj;
            CrossCov[ch][obj] = 0.0f;
            for (obj2=0; obj2<numObjects; obj2++) {
                objIdx2 = firstObject + obj2;
                CrossCov[ch][obj] += D[ch][obj] * RelatedToMatrix[objIdx][objIdx2];
            }
        }
    }
}

static int getSaocCoreGrouping(float Dmx[][SAOC_MAX_OBJECTS],
                                int RelatedToMatrix[SAOC_MAX_OBJECTS][SAOC_MAX_OBJECTS],
                                int numObjects,
                                int numDmxChannels,
                                saocCoreGroup *groupInfo)
{
    int CrossCov[SAOC_MAX_DMx_CHANNELS][SAOC_MAX_OBJECTS] = {{0}};

    getSaocDmxPossitions(Dmx, RelatedToMatrix, CrossCov, numObjects, 0, numDmxChannels, 0);
    getSaocCoreGroups(CrossCov, numObjects, numDmxChannels, groupInfo);

    return 0;
}

```

Fig. 9 (3)

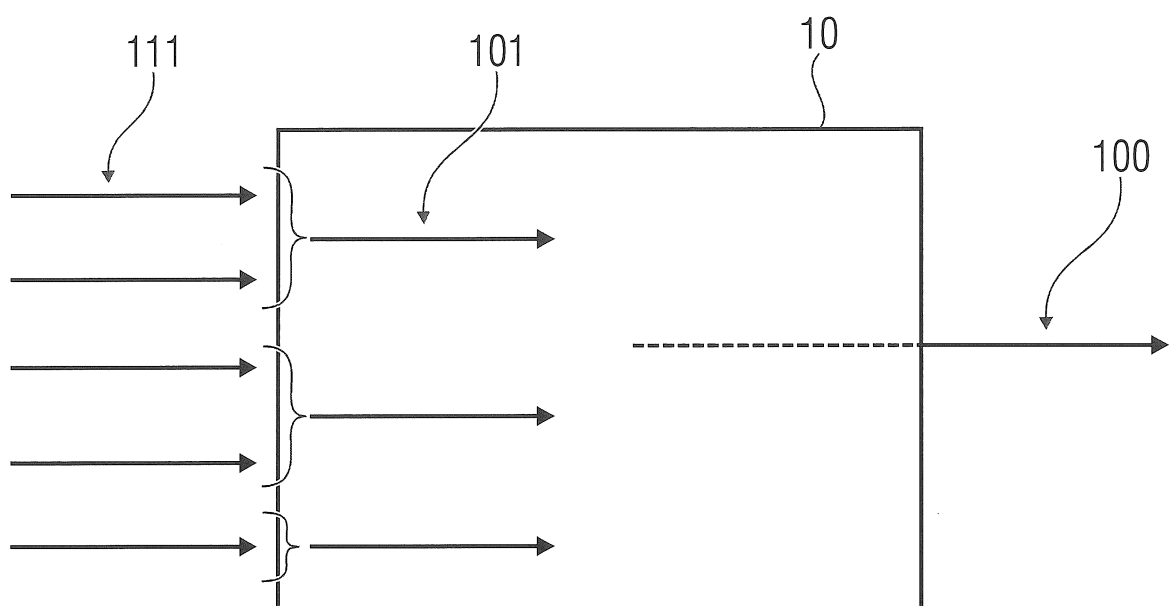


Fig. 10

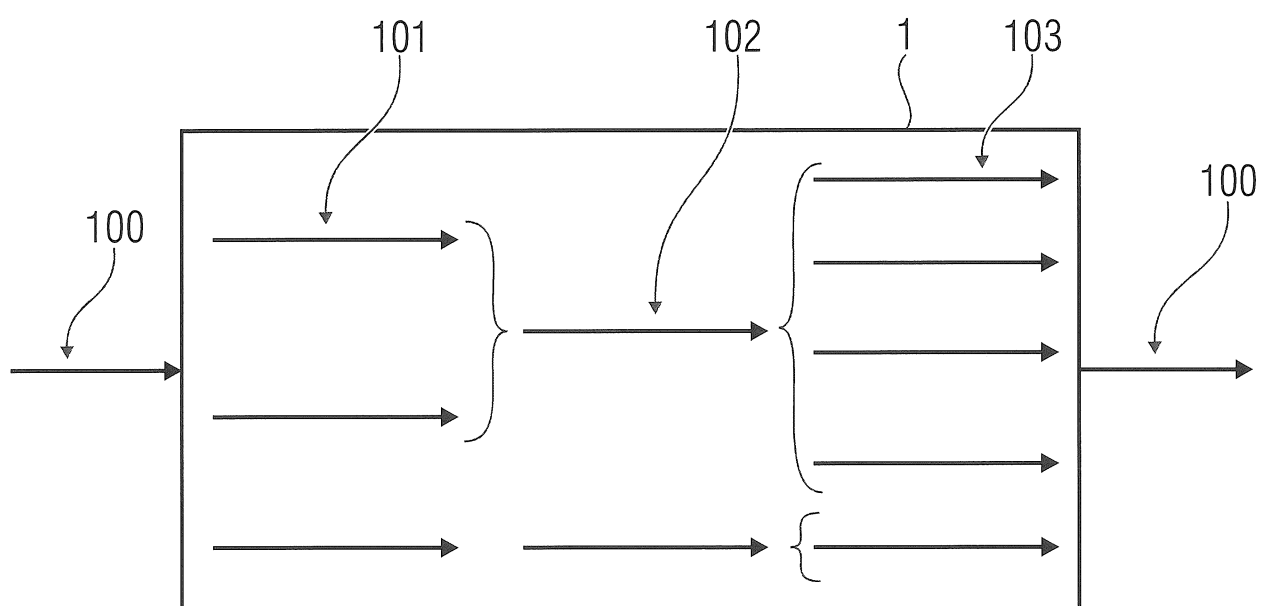


Fig. 11

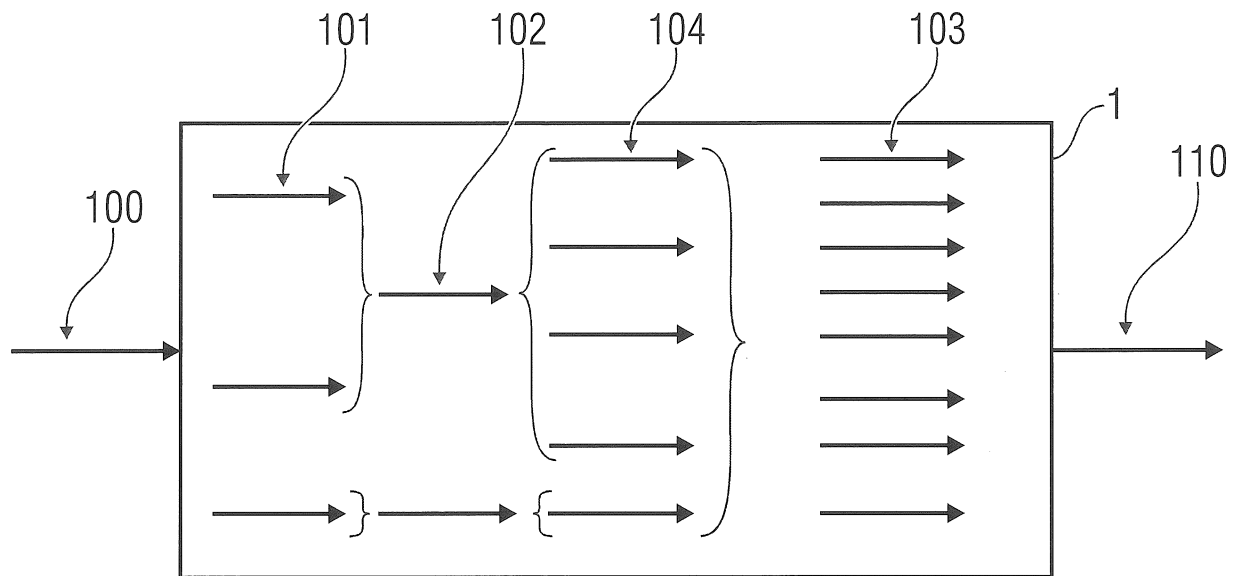


Fig. 12

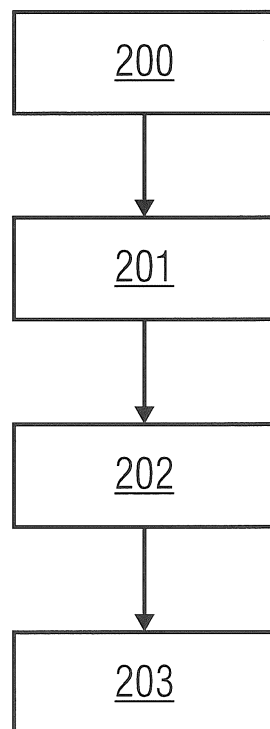


Fig. 13

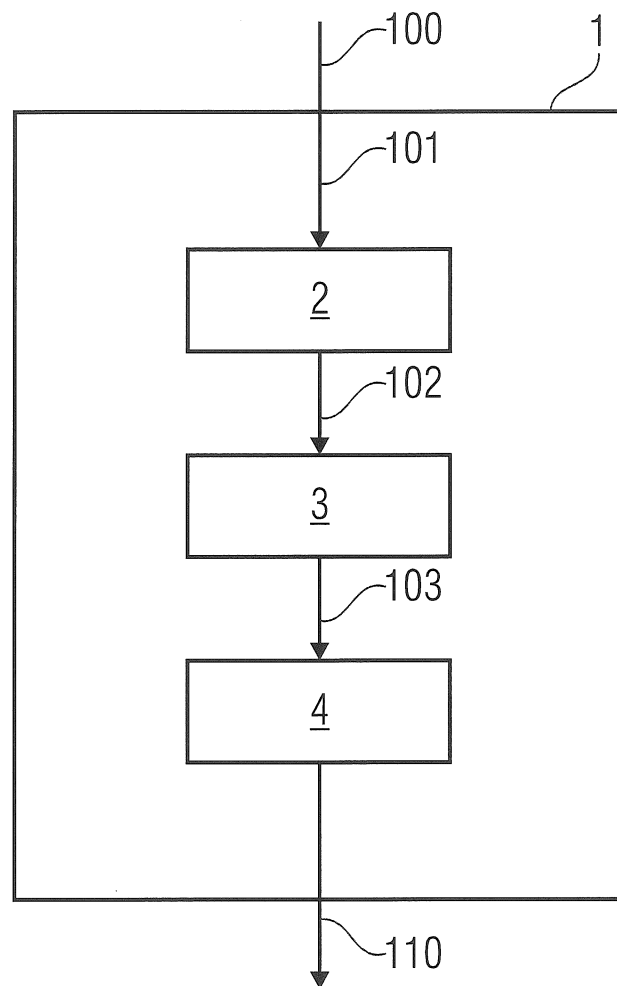


Fig. 14

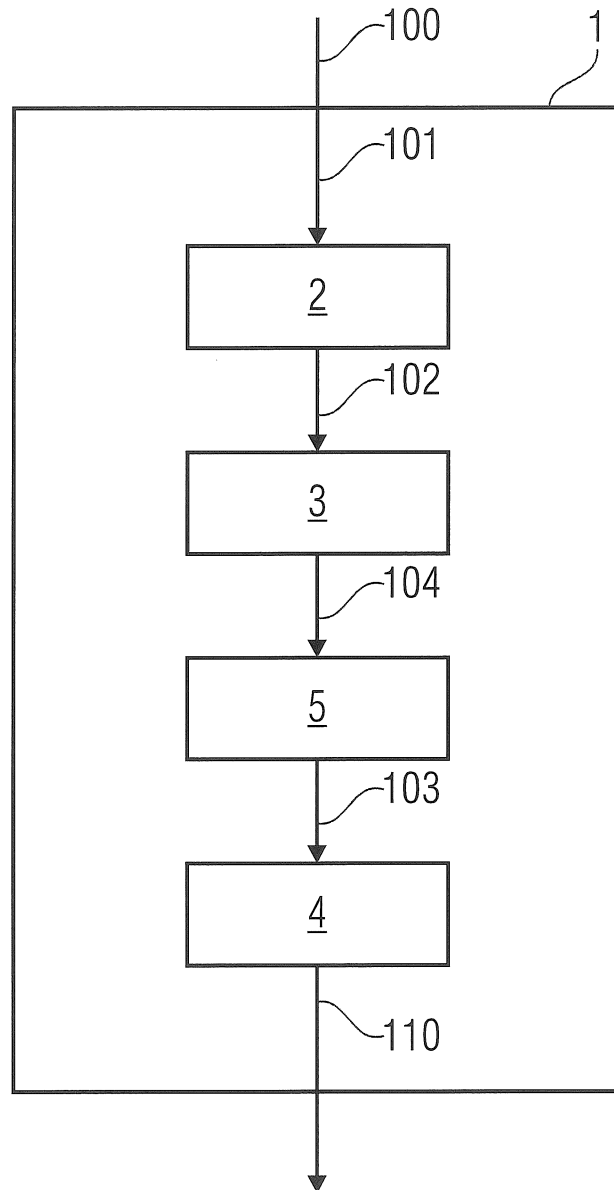


Fig. 15

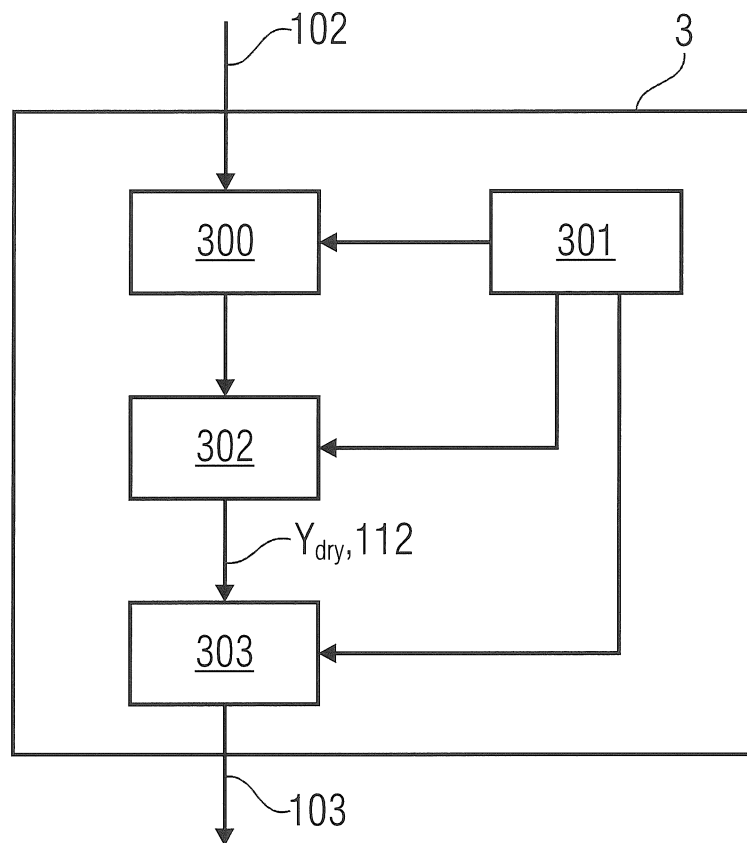


Fig. 16

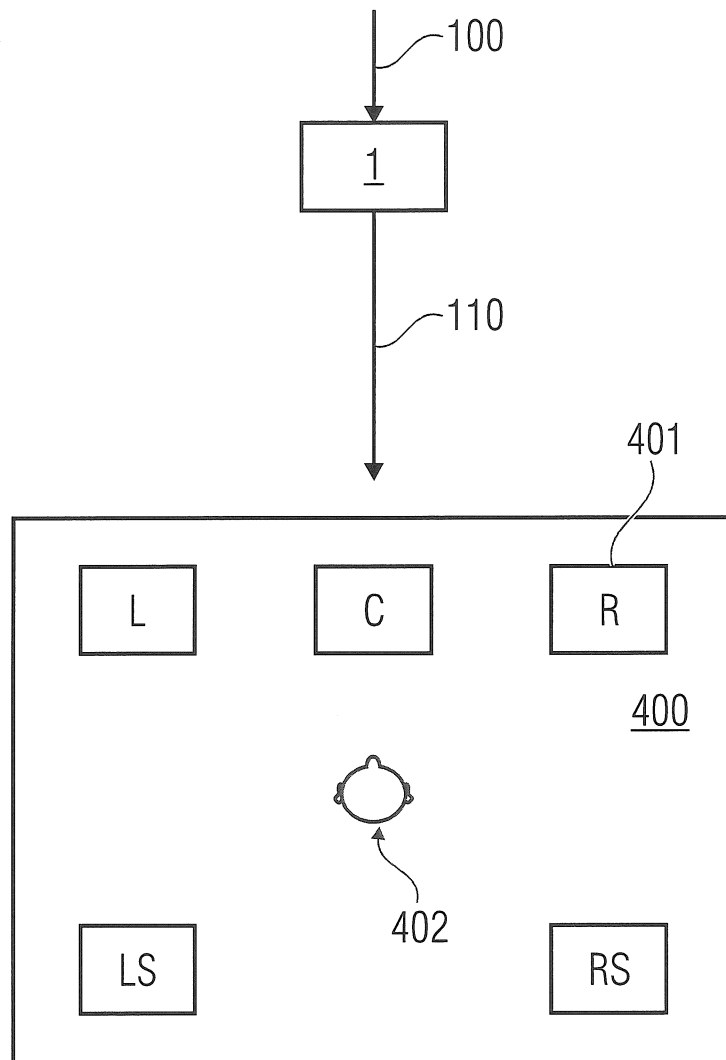


Fig. 17