



(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)



1-0029929

(51)⁷ G10L 19/008

(13) B

(21) 1-2016-02149

(22) 27/11/2014

(86) PCT/EP2014/075787 27/11/2014

(87) WO2015/078956 A1 04/06/2015

(30) 13194664.2 27/11/2013 EP

(45) 25/11/2021 404

(43) 26/09/2016 342A

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

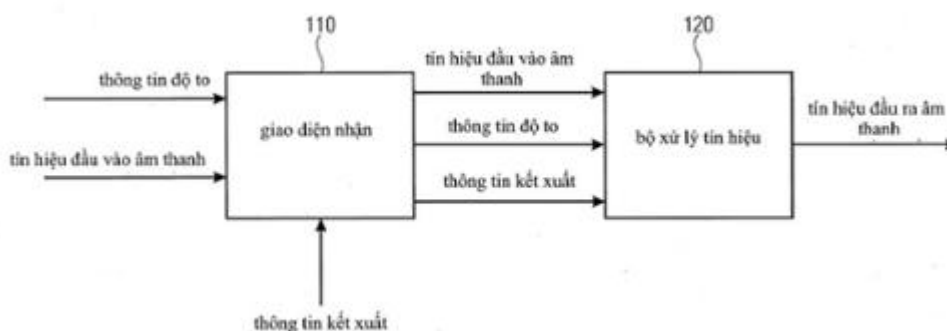
Hansastrasse 27c, 80686 Muenchen, Germany

(72) PAULUS, Jouni (FI); DISCH, Sascha (DE); FUCHS, Harald (DE); GRILL, Bernhard (DE); HELLMUTH, Oliver (DE); MURTAZA, Adrian (RO); RIDDERBUSCH, Falko (DE); TERENTIV, Leon (DE).

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

(54) BỘ GIẢI MÃ TẠO RA TÍN HIỆU ĐẦU RA ÂM THANH, BỘ MÃ HÓA, HỆ THỐNG VÀ PHƯƠNG PHÁP MÃ HÓA VÀ TẠO RA TÍN HIỆU ĐẦU RA ÂM THANH

(57) Sáng chế đề cập đến bộ giải mã tạo ra tín hiệu đầu ra âm thanh, bộ mã hóa, hệ thống và phương pháp mã hóa và tạo ra tín hiệu đầu ra âm thanh. Bộ giải mã tạo ra tín hiệu đầu ra âm thanh bao gồm một hoặc nhiều kênh đầu ra âm thanh được đề xuất. Bộ giải mã bao gồm giao diện nhận (110) để nhận tín hiệu đầu vào âm thanh bao gồm nhiều tín hiệu đối tượng âm thanh, để nhận thông tin độ to trên các tín hiệu đối tượng âm thanh, và để nhận thông tin kết xuất cho biết một hoặc nhiều tín hiệu đối tượng âm thanh sẽ được khuếch đại hoặc được làm suy giảm. Hơn nữa, bộ giải mã bao gồm bộ xử lý tín hiệu (120) để tạo ra một hoặc nhiều kênh đầu ra âm thanh của tín hiệu đầu ra âm thanh. Bộ xử lý tín hiệu (120) được tạo cấu hình để xác định giá trị bù độ to phụ thuộc vào thông tin độ to và phụ thuộc vào thông tin kết xuất. Ngoài ra, bộ xử lý tín hiệu (120) được tạo cấu hình để tạo ra một hoặc nhiều kênh đầu ra âm thanh của tín hiệu đầu ra âm thanh từ tín hiệu đầu vào âm thanh phụ thuộc vào thông tin kết xuất và phụ thuộc vào giá trị bù độ to. Một hoặc nhiều tín hiệu đối tượng âm thanh đường vòng được sử dụng để tạo ra tín hiệu đầu ra âm thanh. Ngoài ra bộ mã hóa được đề xuất.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến việc mã hóa, xử lý và giải mã tín hiệu âm thanh, và cụ thể, sáng chế đề cập đến bộ giải mã, bộ mã hóa và phương pháp ước lượng độ to có thông tin trong hệ thống mã hóa âm thanh dựa trên đối tượng.

Tình trạng kỹ thuật của sáng chế

Gần đây, các kỹ thuật theo tham số để truyền dẫn/lưu trữ tốc độ bit hiệu quả của cảnh âm thanh bao gồm nhiều tín hiệu đối tượng âm thanh đã được đề xuất trong lĩnh vực mã hóa âm thanh [BCC, JSC, SAOC, SAOC1, SAOC2] và tách nguồn có thông tin [ISS1, ISS2, ISS3, ISS4, ISS5, ISS6]. Các kỹ thuật này nhằm khôi phục cảnh âm thanh đầu ra mong muốn hoặc đối tượng nguồn âm thanh mong muốn dựa trên thông tin phụ bổ sung mô tả cảnh âm thanh được lưu trữ/được truyền dẫn và/hoặc các đối tượng nguồn âm thanh trong cảnh âm thanh. Việc khôi phục này diễn ra trong bộ giải mã sử dụng sơ đồ tách nguồn có thông tin. Các đối tượng được khôi phục có thể được kết hợp để tạo ra cảnh âm thanh đầu ra. Phụ thuộc vào cách các đối tượng được kết hợp, độ to cảm nhận được của cảnh đầu ra có thể thay đổi.

Trong truyền thanh TV và vô tuyến, mức âm lượng của các rãnh âm thanh của các chương trình khác nhau có thể được tiêu chuẩn hóa dựa trên các khía cạnh khác nhau, chẳng hạn như mức tín hiệu đỉnh hoặc mức độ to. Phụ thuộc vào các đặc điểm động của các tín hiệu, hai tín hiệu với cùng mức đỉnh có thể có mức chênh lệch rộng về độ to cảm nhận được. Tại thời điểm chuyển đổi giữa các chương trình hoặc kênh, sự chênh lệch trong độ to tín hiệu là rất khó chịu và là nguyên nhân chính gây ra các phản nản của người dùng cuối khi phát sóng.

Trong tình trạng kỹ thuật đã đề xuất chuẩn hóa tất cả các chương trình trên tất cả các kênh giống nhau tới mức tham chiếu chung bằng cách sử dụng phép đo dựa trên

độ to tín hiệu cảm nhận được. Một trong những đề xuất này ở châu Âu là Khuyến nghị EBU R128 [EBU] (sau đây gọi là R128).

Khuyến nghị nêu rằng "độ to chương trình", ví dụ, độ to trung bình trên một chương trình (hoặc một chương trình thương mại, hoặc một số thực thể chương trình có ý nghĩa khác) nên bằng mức xác định (với những độ lệch cho phép nhỏ). Khi ngày càng nhiều đài truyền hình tuân thủ khuyến nghị này và yêu cầu chuẩn hóa, sự chênh lệch về độ to trung bình giữa các chương trình và các kênh sẽ được giảm thiểu.

Việc ước lượng độ to có thể được thực hiện theo các cách khác nhau. Tồn tại nhiều mô hình toán học để ước lượng độ to cảm nhận được của tín hiệu âm thanh. Khuyến nghị EBU R128 tin tưởng vào mô hình được thể hiện trong ITU-R BS.1770 (sau đây được gọi là BS.1770) (xem [ITU]) cho việc ước lượng độ to.

Như đã nêu, ví dụ theo khuyến nghị R128, độ to chương trình, ví dụ độ to trung bình trên một chương trình cần bằng mức xác định với những độ lệch cho phép nhỏ. Tuy nhiên, điều này dẫn đến những vấn đề quan trọng khi việc kết xuất âm thanh được tiến hành, chưa được giải quyết cho đến tình trạng kỹ thuật hiện tại. Việc tiến hành kết xuất âm thanh ở phía bộ giải mã có ảnh hưởng đáng kể đến độ to tổng thể/tổng độ to của tín hiệu đầu vào âm thanh nhận được. Tuy nhiên, mặc dù việc kết xuất cảnh được tiến hành, tổng độ to của tín hiệu âm thanh nhận được sẽ không đổi.

Gần đây, không có giải pháp phía bộ giải mã cụ thể nào cho vấn đề này.

EP 2 146 522 A1 ([EP]) đề cập đến các khái niệm tạo ra các tín hiệu đầu ra âm thanh sử dụng siêu dữ liệu dựa trên đối tượng. Ít nhất một tín hiệu đầu ra âm thanh được tạo ra biểu diễn sự chồng chập của ít nhất hai tín hiệu đầu ra âm thanh khác nhau, nhưng không đề xuất giải pháp cho vấn đề này.

WO 2008/035275 A2 ([BRE]) mô tả hệ thống âm thanh bao gồm bộ mã hóa mà mã hóa các đối tượng âm thanh trong bộ phận mã hóa mà tạo ra tín hiệu âm thanh trộn giảm và dữ liệu tham số biểu diễn nhiều đối tượng âm thanh. Tín hiệu âm thanh trộn giảm và dữ liệu tham số được truyền đến bộ giải mã bao gồm bộ phận giải mã tạo ra các bản sao gần đúng của các đối tượng âm thanh và bộ phận kết xuất tạo ra tín hiệu đầu ra từ các đối tượng âm thanh. Bộ giải mã còn chứa bộ xử lý để tạo ra dữ liệu biến

đổi mã hóa mà được gửi đến bộ mã hóa. Bộ mã hóa sau đó biến đổi mã hóa các đối tượng âm thanh, và cụ thể là biến đổi dữ liệu tham số, để đáp ứng với dữ liệu biến đổi mã hóa. Cách tiếp cận cho phép thao tác các đối tượng âm thanh được kiểm soát bởi bộ giải mã nhưng được thực hiện hoàn toàn hoặc một phần bởi bộ mã hóa. Do đó, thao tác có thể được thực hiện trên đối tượng âm thanh độc lập thực tế hơn là trên các bản sao gần đúng, từ đó mang lại hiệu suất được cải thiện.

EP 2 146 522 A1 ([SCH]) bộc lộ thiết bị tạo ra ít nhất một tín hiệu đầu ra âm thanh biểu diễn sự chồng chập của ít nhất hai đối tượng âm thanh khác nhau bao gồm bộ xử lý để xử lý tín hiệu đầu vào âm thanh để cung cấp phép biểu diễn đối tượng của tín hiệu đầu vào âm thanh, trong đó phép biểu diễn đối tượng này có thể được tạo ra bởi phép tính gần đúng được dẫn hướng theo tham số của các đối tượng ban đầu sử dụng tín hiệu trộn giảm đối tượng. Bộ thao tác đối tượng thao tác một cách riêng rẽ các đối tượng sử dụng đối tượng âm thanh dựa trên siêu dữ liệu đề cập đến các đối tượng âm thanh riêng rẽ để thu được các đối tượng âm thanh được thao tác. Các đối tượng âm thanh được thao tác được trộn sử dụng bộ trộn đối tượng để cuối cùng thu được tín hiệu đầu ra âm thanh có một hoặc nhiều tín hiệu kênh phụ thuộc vào cài đặt kết xuất cụ thể.

WO 2008/046531 A1 ([ENG]) bộc lộ bộ mã hóa đối tượng âm thanh để tạo ra tín hiệu đối tượng được mã hóa sử dụng nhiều đối tượng âm thanh bao gồm bộ tạo thông tin trộn giảm để tạo ra thông tin trộn giảm biểu thị sự phân phối của nhiều đối tượng âm thanh thành ít nhất hai kênh trộn giảm, bộ tạo tham số đối tượng âm thanh để tạo ra các tham số đối tượng cho các đối tượng âm thanh, và giao diện đầu ra để tạo ra tín hiệu đầu ra âm thanh được nhập sử dụng thông tin trộn giảm và các tham số đối tượng. Bộ tổng hợp âm thanh sử dụng thông tin trộn giảm để tạo ra dữ liệu đầu ra có thể sử dụng được để tạo nhiều kênh đầu ra của cấu hình đầu ra âm thanh được xác định trước.

Mong muốn ước lượng chính xác độ to trung bình đầu ra hoặc thay đổi độ to trung bình mà không có độ trễ và khi chương trình không thay đổi hoặc cảnh kết xuất không được thay đổi, ước lượng độ to trung bình cũng cần giữ nguyên trạng thái tĩnh.

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.

[EBU] EBU Recommendation R 128 “Loudness normalization and permitted maximum level of audio signals”, Geneva, 2011.

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

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

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

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

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

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

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

[ITU] International Telecommunication Union: “Recommendation ITU-R BS.1770-3 – Algorithms to measure audio programme loudness and true-peak audio level”, Geneva, 2012.

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

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

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

[EP] EP 2146522 A1: S. Schreiner, W. Fiesel, M. Neusinger, O. Hellmuth, R. Sperschneider, ”Apparatus and method for generating audio output signals using object based metadata“, 2010.

[DE] ISO/IEC, “MPEG audio technologies – Part 2: Spatial Audio Object Coding (SAOC) – Amendment 3, Dialogue Enhancement”, ISO/IEC 23003-2:2010/DAM 3, Dialogue Enhancement.

[BRE] WO 2008/035275 A2.

[SCH] EP 2 146 522 A1.

[ENG] WO 2008/046531 A1.

Bản chất kỹ thuật của sáng chế

Mục đích của sáng chế là đề xuất các khái niệm mã hóa, xử lý và giải mã tín hiệu âm thanh. Mục đích của sáng chế được giải quyết bởi bộ giải mã theo điểm 1, bởi bộ mã hóa theo điểm 15, bởi hệ thống theo điểm 18, bởi phương pháp theo điểm 19, bởi phương pháp theo điểm 20 và bởi vật ghi có thể đọc được bằng máy tính bao gồm chương trình máy tính theo điểm 22, 23, 24 và 25.

Sáng chế đề xuất cách có thông tin để ước lượng độ to của đầu ra trong hệ thống mã hóa âm thanh dựa trên đối tượng. Các khái niệm được đề xuất dựa trên thông tin về độ to của các đối tượng trong hỗn hợp âm thanh được cung cấp cho bộ giải mã. Bộ giải

mã sử dụng thông tin này cùng với thông tin kết xuất để ước lượng độ to của tín hiệu đầu ra. Sau đó điều này cho phép, ví dụ, ước lượng sự chênh lệch độ to giữa trộn giảm mặc định và đầu ra được kết xuất. Sau đó có thể bù cho sự chênh lệch để thu được một cách gần đúng độ to không đổi trong đầu ra không chú ý đến thông tin kết xuất. Việc ước lượng độ to trong bộ giải mã diễn ra theo cách tham số đầy đủ, và nó rất nhẹ nhàng và chính xác trong tính toán khi so sánh với các khái niệm ước lượng độ to dựa trên tín hiệu.

Các khái niệm để thu được thông tin về độ to của cảnh đầu ra đặc trưng sử dụng hoàn toàn các khái niệm tham số được đề xuất, mà sau đó cho phép xử lý độ to mà không có ước lượng độ to dựa trên tín hiệu rõ ràng trong bộ giải mã. Hơn nữa, kỹ thuật đặc trưng của Mã hóa đối tượng âm thanh không gian (Spatial Audio Object Coding - SAOC) được tiêu chuẩn hóa bởi MPEG [SAOC] được mô tả, nhưng các khái niệm được đề xuất có thể cũng được sử dụng kết hợp với các kỹ thuật mã hóa đối tượng âm thanh khác.

Bộ giải mã tạo ra tín hiệu đầu ra âm thanh bao gồm một hoặc nhiều kênh đầu ra âm thanh được đề xuất. Bộ giải mã bao gồm giao diện nhận để nhận tín hiệu đầu vào âm thanh bao gồm nhiều tín hiệu đối tượng âm thanh, để nhận thông tin độ to về các tín hiệu đối tượng âm thanh, và để nhận thông tin kết xuất cho biết một hoặc nhiều tín hiệu đối tượng hoặc sẽ được khuếch đại hoặc được làm suy giảm. Hơn nữa, bộ giải mã bao gồm bộ xử lý tín hiệu để tạo ra một hoặc nhiều kênh đầu ra âm thanh của tín hiệu đầu ra âm thanh. Bộ xử lý tín hiệu được tạo cấu hình để xác định giá trị bù độ to phụ thuộc vào thông tin độ to và phụ thuộc vào thông tin kết xuất. Ngoài ra, bộ xử lý tín hiệu được tạo cấu hình để tạo ra một hoặc nhiều kênh đầu ra âm thanh của tín hiệu đầu ra âm thanh từ tín hiệu đầu vào âm thanh phụ thuộc vào thông tin kết xuất và phụ thuộc vào giá trị bù độ to.

Theo một phương án, bộ xử lý tín hiệu có thể được tạo cấu hình để tạo ra một hoặc nhiều kênh đầu ra âm thanh của tín hiệu đầu ra âm thanh từ tín hiệu đầu vào âm thanh phụ thuộc vào thông tin kết xuất và phụ thuộc vào giá trị bù độ to, chẳng hạn như độ to của tín hiệu đầu ra âm thanh bằng với độ to của tín hiệu đầu vào âm thanh, hoặc chẳng hạn như độ to của tín hiệu đầu ra âm thanh gần hơn với độ to của tín hiệu

đầu vào âm thanh so với độ to của tín hiệu âm thanh được biến đổi mà sẽ có được từ việc biến đổi tín hiệu đầu vào âm thanh bằng cách khuếch đại hoặc làm suy giảm các tín hiệu đối tượng âm thanh của tín hiệu đầu vào âm thanh theo thông tin kết xuất.

Theo phương án khác, mỗi tín hiệu đối tượng âm thanh của tín hiệu đầu vào âm thanh có thể được ấn định cho chính xác một nhóm trong số hai hoặc nhiều nhóm, trong đó mỗi nhóm trong số hai hoặc nhiều nhóm có thể bao gồm một hoặc nhiều tín hiệu đối tượng âm thanh của tín hiệu đầu vào âm thanh. Trong phương án này, giao diện nhận có thể được tạo cấu hình để nhận giá trị độ to từ mỗi nhóm của hai hoặc nhiều hơn hai nhóm như thông tin độ to, trong đó giá trị độ to đã nêu cho biết tổng độ to ban đầu của một hoặc nhiều tín hiệu đối tượng âm thanh của nhóm đã nêu. Hơn nữa, giao diện nhận có thể được tạo cấu hình để nhận thông tin kết xuất cho biết ít nhất một nhóm trong số hai hoặc nhiều hơn hai nhóm hoặc một hoặc nhiều tín hiệu đối tượng âm thanh của nhóm đã nêu sẽ được khuếch đại hoặc sẽ được làm suy giảm bằng cách cho biết tổng độ to được biến đổi của một hoặc nhiều tín hiệu đối tượng âm thanh của nhóm đã nêu. Hơn nữa, trong phương án này, bộ xử lý tín hiệu có thể được tạo cấu hình để xác định giá trị bù độ to phụ thuộc vào tổng độ to được biến đổi của mỗi nhóm của ít nhất một nhóm của hai hoặc nhiều hơn hai nhóm và phụ thuộc vào tổng độ to ban đầu của mỗi nhóm trong số hai hoặc nhiều hơn hai nhóm. Ngoài ra, bộ xử lý tín hiệu có thể được tạo cấu hình để tạo ra một hoặc nhiều kênh đầu ra âm thanh của tín hiệu đầu ra âm thanh từ tín hiệu đầu vào âm thanh phụ thuộc vào tổng độ to được biến đổi của mỗi nhóm trong số ít nhất một nhóm trong số hai hoặc nhiều hơn hai nhóm và phụ thuộc vào giá trị bù độ to.

Trong các phương án cụ thể, ít nhất một nhóm trong số hai hoặc nhiều hơn hai nhóm có thể bao gồm hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh.

Ngoài ra bộ mã hóa được đề xuất. Bộ mã hóa bao gồm bộ phận mã hóa dựa trên đối tượng để mã hóa nhiều tín hiệu đối tượng âm thanh để thu được tín hiệu âm thanh được mã hóa bao gồm nhiều tín hiệu đối tượng âm thanh. Hơn nữa, bộ mã hóa bao gồm bộ phận mã hóa độ to đối tượng để mã hóa thông tin độ to trên tín hiệu đối tượng âm thanh. Thông tin độ to bao gồm một hoặc nhiều giá trị độ to, trong đó mỗi giá trị

trong số một hoặc nhiều giá trị độ to phụ thuộc vào một hoặc nhiều tín hiệu đối tượng âm thanh.

Theo phương án, mỗi tín hiệu đối tượng âm thanh của tín hiệu âm thanh được mã hóa có thể được ấn định cho chính xác một nhóm trong số hai hoặc nhiều hơn hai nhóm, trong đó mỗi nhóm trong số hai hoặc nhiều hơn hai nhóm bao gồm một hoặc nhiều tín hiệu đối tượng âm thanh của tín hiệu âm thanh được mã hóa. Bộ phận mã hóa độ to đối tượng có thể được tạo cấu hình để xác định một hoặc nhiều giá trị độ to của thông tin độ to bằng cách xác định giá trị độ to cho mỗi nhóm trong số hai hoặc nhiều hơn hai nhóm, trong đó giá trị độ to đã nêu của nhóm đã nêu cho biết tổng độ to ban đầu của một hoặc nhiều tín hiệu đối tượng âm thanh của nhóm đã nêu.

Ngoài ra, hệ thống được đề xuất. Hệ thống bao gồm bộ mã hóa theo một trong số các phương án được mô tả ở trên để mã hóa nhiều tín hiệu đối tượng âm thanh để thu được tín hiệu âm thanh được mã hóa bao gồm nhiều tín hiệu đối tượng âm thanh, và để mã hóa thông tin độ to trên các tín hiệu đối tượng âm thanh. Ngoài ra, hệ thống bao gồm bộ giải mã theo một trong số các phương án được mô tả ở trên để tạo ra tín hiệu đầu ra âm thanh bao gồm một hoặc nhiều kênh đầu ra âm thanh. Bộ giải mã được tạo cấu hình để nhận tín hiệu âm thanh được mã hóa như tín hiệu đầu vào âm thanh và thông tin độ to. Ngoài ra, bộ giải mã còn được tạo cấu hình để nhận thông tin kết xuất. Ngoài ra, bộ giải mã được tạo cấu hình để xác định giá trị bù độ to phụ thuộc vào thông tin độ to và phụ thuộc vào thông tin kết xuất. Hơn nữa, bộ giải mã được tạo cấu hình để tạo ra một hoặc nhiều kênh đầu ra âm thanh của tín hiệu đầu ra âm thanh từ tín hiệu đầu vào âm thanh phụ thuộc vào thông tin kết xuất và phụ thuộc vào giá trị bù độ to.

Ngoài ra, phương pháp tạo ra tín hiệu đầu ra âm thanh bao gồm một hoặc nhiều kênh đầu ra âm thanh được đề xuất. Phương pháp bao gồm các bước:

- Nhận tín hiệu đầu vào âm thanh bao gồm nhiều tín hiệu đối tượng âm thanh.
- Nhận thông tin độ to về tín hiệu đối tượng âm thanh.
- Nhận thông tin kết xuất cho biết một hoặc nhiều tín hiệu đối tượng âm thanh sẽ được khuếch đại hoặc sẽ được làm suy giảm.

- Xác định giá trị bù độ to phụ thuộc vào thông tin độ to và phụ thuộc vào thông tin kết xuất. Và:

- Tạo ra một hoặc nhiều kênh đầu ra âm thanh của tín hiệu đầu ra âm thanh từ tín hiệu đầu vào âm thanh phụ thuộc vào thông tin kết xuất và phụ thuộc vào giá trị bù độ to.

Hơn nữa, phương pháp mã hóa được đề xuất. Phương pháp bao gồm các bước:

- Mã hóa tín hiệu đầu vào âm thanh bao gồm nhiều tín hiệu đối tượng âm thanh.
Và:

- Mã hóa thông tin độ to về tín hiệu đối tượng âm thanh, trong đó thông tin độ to bao gồm một hoặc nhiều giá trị độ to, trong đó mỗi giá trị độ to trong số một hoặc nhiều giá trị độ to phụ thuộc vào một hoặc nhiều tín hiệu đối tượng âm thanh.

Hơn nữa, chương trình máy tính để thực hiện phương pháp được mô tả ở trên khi được chạy trên máy tính hoặc bộ xử lý tín hiệu được đề xuất.

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

Mô tả vắn tắt các hình vẽ

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

Fig.1 minh họa bộ giải mã tạo ra tín hiệu đầu ra âm thanh bao gồm một hoặc nhiều kênh đầu ra âm thanh theo một phương án,

Fig.2 minh họa bộ mã hóa theo một phương án,

Fig.3 minh họa hệ thống mã hóa tín hiệu đối tượng âm thanh và tạo ra tín hiệu đầu ra âm thanh theo một phương án,

Fig.4 minh họa hệ thống mã hóa đối tượng âm thanh không gian bao gồm bộ mã hóa SAOC và bộ giải mã SAOC,

Fig.5 minh họa bộ giải mã SAOC bao gồm bộ giải mã thông tin phụ, bộ tách đối tượng và bộ kết xuất,

Fig.6 minh họa trạng thái của độ to tín hiệu đầu ra ước lượng trên sự thay đổi độ to,

Fig.7 mô tả sự ước lượng độ to có thông tin theo một phương án, minh họa các thành phần của bộ mã hóa và bộ giải mã theo một phương án,

Fig.8 minh họa bộ mã hóa theo phương án khác,

Fig.9 minh họa bộ mã hóa và bộ giải mã theo phương án liên quan đến việc cải thiện âm thanh hội thoại-SAOC, mà bao gồm các kênh đường vòng,

Fig.10 mô tả minh họa thứ nhất của sự thay đổi độ to được đo và kết quả của việc sử dụng các khái niệm được cung cấp để ước lượng sự thay đổi theo độ to và theo cách tham số,

Fig.11 mô tả minh họa thứ hai của sự thay đổi độ to được đo và kết quả của việc sử dụng các khái niệm được cung cấp để ước lượng sự thay đổi theo độ to và theo cách tham số, và

Fig.12 minh họa phương án khác để thực hiện bù độ to theo sáng chế.

Mô tả chi tiết sáng chế

Trước khi các phương án ưu tiên được mô tả chi tiết, ước lượng độ to, mã hóa đối tượng âm thanh không gian (SAOC) và cải thiện âm thanh hội thoại (Dialogue Enhancement - DE) được mô tả.

Trước tiên, việc ước lượng độ to được mô tả.

Như đã nêu ở trên, khuyến nghị EBU R1228 dựa trên mô hình được đưa ra trong ITU-R BS.1770 về ước lượng độ to. Phép đo này sẽ được sử dụng như một ví dụ, nhưng các khái niệm được mô tả dưới đây có thể cũng được áp dụng cho các phép đo độ to khác.

Thao tác ước lượng độ to theo BS.1770 là tương đối đơn giản và nó dựa trên các bước chính sau [ITU]:

- Tín hiệu đầu vào x_i (hoặc các tín hiệu trong trường hợp tín hiệu đa kênh) được lọc với bộ lọc K (tổ hợp của bộ lọc nghiêng và bộ lọc thông cao) để thu được (các) tín hiệu y_i .

- Năng lượng bình phương trung bình z_i của tín hiệu y_i được tính toán.

- Trong trường hợp tín hiệu đa kênh, việc gán trọng số kênh G_i được áp dụng, và các tín hiệu được gán trọng số được tính tổng. Độ to của tín hiệu sau đó được xác định là:

$$L = c + 10 \log_{10} \sum_i G_i z_i$$

với giá trị hằng số $c = -0.691$. Đầu ra sau đó được biểu thị trong các bộ phận của “LKFS” (độ to, được gán trọng số K, liên quan đến tỷ lệ đầy đủ) mà các tỷ lệ tương tự như tỷ lệ dèxiben.

Trong công thức ở trên, G_i có thể, ví dụ, bằng 1 đối với một số kênh, trong khi G_i có thể, ví dụ, bằng 1,41 đối với một số kênh khác. Ví dụ, nếu kênh bên trái, kênh bên phải, kênh trung tâm, kênh xung quanh bên trái và kênh xung quanh bên phải được xem xét, các trọng số tương ứng G_i có thể, ví dụ, bằng 1 đối với kênh bên trái, bên phải và trung tâm, và ví dụ, có thể bằng 1,41 đối với kênh xung quanh bên trái và kênh xung quanh bên phải, xem [ITU].

Có thể thấy rằng, giá trị độ to L liên quan gần với lôgarit của năng lượng tín hiệu.

Sau đây, mã hóa đối tượng âm thanh không gian được mô tả.

Các khái niệm mã hóa âm thanh dựa trên đối tượng cho phép linh hoạt hơn ở phía bộ giải mã của chuỗi. Ví dụ về khái niệm mã hóa âm thanh dựa trên đối tượng là mã hóa đối tượng âm thanh không gian (SAOC).

Fig.4 minh họa hệ thống mã hóa đối tượng âm thanh không gian (SAOC) bao gồm bộ mã hóa SAOC 410 và bộ giải mã SAOC 420.

Bộ mã hóa SAOC 410 nhận N tín hiệu đối tượng âm thanh $S_1, \dots, S_N$ làm đầu vào. Ngoài ra, bộ mã hóa SAOC 410 còn nhận các lệnh "thông tin trộn D " cách các đối tượng này cần được kết hợp để thu được tín hiệu trộn giảm bao gồm M kênh trộn giảm $X_1, \dots, X_M$. Bộ mã hóa SAOC 410 trích xuất một số thông tin phụ từ các đối tượng và từ quy trình trộn giảm, và thông tin phụ này được truyền dẫn và/hoặc được lưu trữ cùng với tín hiệu trộn giảm.

Đặc điểm chính của hệ thống SAOC là tín hiệu trộn giảm X bao gồm các kênh trộn giảm $X_1, \dots, X_M$ tạo thành tín hiệu có nghĩa về mặt ngữ nghĩa. Nói cách khác, có thể nghe được tín hiệu trộn giảm. Ví dụ, nếu bộ nhận không có chức năng bộ giải mã SAOC, bộ nhận có thể vẫn luôn cung cấp tín hiệu trộn giảm làm đầu ra.

Fig.5 minh họa bộ giải mã SAOC bao gồm bộ giải mã thông tin phụ 510, bộ tách đối tượng 520 và bộ kết xuất 530. Bộ giải mã SAOC minh họa trên Fig.5 nhận, ví dụ, tín hiệu trộn giảm và thông tin phụ, từ bộ mã hóa SAOC. Tín hiệu trộn giảm có thể được xem xét như tín hiệu đầu vào âm thanh bao gồm các tín hiệu đối tượng âm thanh, như các tín hiệu đối tượng âm thanh được trộn giảm trong tín hiệu trộn giảm (các tín hiệu đối tượng âm thanh được trộn trong một hoặc nhiều kênh trộn giảm của tín hiệu trộn giảm).

Bộ giải mã SAOC có thể, ví dụ, sau đó cố gắng để khôi phục (ảo) các đối tượng ban đầu, ví dụ, bằng cách sử dụng bộ tách đối tượng 520, ví dụ, sử dụng thông tin phụ được giải mã. Các khôi phục đối tượng (ảo) $\hat{S}_1, \dots, \hat{S}_N$, ví dụ, các tín hiệu đối tượng âm thanh được khôi phục, sau đó được kết hợp dựa trên thông tin kết xuất, ví dụ, ma trận kết xuất $\mathbf{R}$, để tạo ra K kênh đầu ra âm thanh $Y_1, \dots, Y_K$ của tín hiệu đầu ra âm thanh $\mathbf{Y}$.

Trong SAOC, thông thường, các tín hiệu đối tượng âm thanh, ví dụ, được khôi phục, ví dụ, bằng cách sử dụng thông tin hiệp phương sai, ví dụ ma trận hiệp phương sai tín hiệu $\mathbf{E}$, mà được truyền từ bộ mã hóa SAOC đến bộ giải mã SAOC.

Ví dụ, công thức sau có thể được sử dụng để khôi phục các tín hiệu đối tượng âm thanh ở phía bộ giải mã:

$$\hat{\mathbf{S}} = \mathbf{G}\mathbf{X} \text{ với } \mathbf{G} \approx \mathbf{E} \mathbf{D}^H (\mathbf{D} \mathbf{E} \mathbf{D}^H)^{-1}$$

trong đó

N là số tín hiệu đối tượng âm thanh,

$N_{\text{mẫu}}$ là số mẫu được xem xét của tín hiệu đối tượng âm thanh.

M là số kênh trộn giảm,

$\mathbf{X}$ là tín hiệu âm thanh trộn giảm, kích thước $M \times N_{\text{mẫu}}$,

$\mathbf{D}$ là ma trận trộn giảm, kích thước $M \times N$

E là ma trận hiệp phương sai tín hiệu, kích thước $N \times N$ được xác định như $E = X X^H$

$\hat{S}$ là các tín hiệu đối tượng âm thanh N được khôi phục theo tham số, kích thước $N \times N_{\text{mẫu}}$

$(\cdot)^H$ toán tử (Hermit) tự liên hợp mà biểu diễn chuyển vị liên hợp của $(\cdot)$

Sau đó, ma trận kết xuất $\mathbf{R}$ được áp dụng trên các tín hiệu đối tượng âm thanh được khôi phục $\hat{S}$ để thu được các kênh đầu ra âm thanh của tín hiệu đầu ra âm thanh $\mathbf{Y}$, ví dụ, theo công thức:

$$\mathbf{Y} = \mathbf{R} \hat{\mathbf{S}}$$

trong đó

K là số kênh đầu ra âm thanh $Y_1, \dots, Y_K$ của tín hiệu đầu ra âm thanh $\mathbf{Y}$.

$\mathbf{R}$ là ma trận kết xuất có kích thước $K \times N$

$\mathbf{Y}$ là tín hiệu đầu ra âm thanh bao gồm K kênh đầu ra âm thanh,

kích thước $K \times N_{\text{mẫu}}$

Trên Fig.5, quy trình khôi phục đối tượng, ví dụ, được thực hiện bởi bộ tách đối tượng 520, được đề cập đến với khái niệm "ảo" hoặc "tùy chọn", vì có thể không nhất thiết phải diễn ra, nhưng chức năng mong muốn có thể thu được bằng cách tổ hợp sự khôi phục và các bước kết xuất trong miền tham số (tức là, tổ hợp các phương trình).

Nói cách khác, đầu tiên thay vì khôi phục các tín hiệu đối tượng âm thanh sử dụng thông tin trộn $\mathbf{D}$ và thông tin hiệp phương sai $\mathbf{E}$, và sau đó áp dụng thông tin kết xuất $\mathbf{R}$ trên các tín hiệu đối tượng âm thanh được khôi phục để thu được các kênh đầu ra âm thanh $Y_1, \dots, Y_K$, cả hai bước có thể được tiến hành theo bước riêng lẻ, sao cho các kênh đầu ra âm thanh $Y_1, \dots, Y_K$ được tạo ra một cách trực tiếp từ các kênh trộn giảm.

Ví dụ, công thức sau đây có thể được sử dụng:

$$\mathbf{Y} = \mathbf{R} \mathbf{G} \mathbf{X} \quad \text{với} \quad \mathbf{G} \approx \mathbf{E} \mathbf{D}^H (\mathbf{D} \mathbf{E} \mathbf{D}^H)^{-1}$$

Về nguyên tắc, thông tin kết xuất $\mathbf{R}$ có thể yêu cầu bất kỳ tổ hợp nào của các tín hiệu đối tượng âm thanh ban đầu. Tuy nhiên, trong thực tế, việc khôi phục đối tượng

có thể bao gồm các lỗi khôi phục và các cảnh đầu ra được yêu cầu có thể không nhất thiết phải đạt được. Như một nguyên tắc chung cơ bản bao gồm nhiều trường hợp thực tế, càng nhiều cảnh đầu ra được yêu cầu khác với tín hiệu trộn giảm, sẽ có càng nhiều lỗi khôi phục có thể nghe được.

Sau đây, việc cải thiện âm thanh hội thoại (DE) sẽ được mô tả. Kỹ thuật SAOC có thể được sử dụng để hiện thực hóa kịch bản. Lưu ý rằng, mặc dù tên gọi "cải thiện âm thanh hội thoại" gợi ý việc tập trung vào các tín hiệu định hướng hội thoại, nguyên tắc tương tự cũng có thể được sử dụng với các dạng tín hiệu khác.

Trong kịch bản DE, mức độ tự do trong hệ thống được giới hạn từ trường hợp tổng quát.

Ví dụ, các tín hiệu đối tượng âm thanh $S_1, \dots, S_N = \mathbf{S}$ được nhóm (và có thể được trộn) thành hai siêu-đối tượng của đối tượng nổi (foreground object - FGO) $\mathbf{S}_{FGO}$ và đối tượng nền (background object - BGO) $\mathbf{S}_{BGO}$.

Ngoài ra, cảnh đầu ra $Y_1, \dots, Y_K = \mathbf{Y}$ giống với tín hiệu trộn giảm $X_1, \dots, X_M = \mathbf{X}$. Cụ thể hơn, cả hai tín hiệu có cùng hạng bậc, tức là $K = M$, và người dùng cuối chỉ có thể điều khiển mức trộn tương ứng của hai siêu đối tượng FGO và BGO. Nói chính xác hơn, tín hiệu trộn giảm thu được bằng cách trộn FGO và BGO với một số trọng số vô hướng:

$$\mathbf{X} = h_{FGO} \mathbf{S}_{FGO} + h_{BGO} \mathbf{S}_{BGO},$$

và cảnh đầu ra thu được tương tự với một số trọng số vô hướng của FGO và BGO:

$$\mathbf{Y} = g_{FGO} \mathbf{S}_{FGO} + g_{BGO} \mathbf{S}_{BGO}.$$

Phụ thuộc vào các giá trị tương đối của các trọng số trộn, sự cân bằng giữa FGO và BGO có thể thay đổi. Ví dụ, với việc thiết lập

$$\begin{cases} g_{FGO} > h_{FGO} \\ g_{BGO} = h_{BGO} \end{cases}$$

có thể tăng mức tương đối của FGO trong hỗn hợp. Nếu FGO là hội thoại, thiết lập này cung cấp chức năng cải thiện âm thanh hội thoại.

Ví dụ về trường hợp sử dụng, BGO có thể là tạp âm ở sân vận động và âm nền khác trong các sự kiện thể thao và FGO là tiếng nói của người bình luận. Chức năng DE cho phép người dùng cuối khuếch đại hoặc làm suy giảm mức của người bình luận so với nền.

Các phương án dựa trên sự phát hiện ra rằng việc sử dụng kỹ thuật SAOC (hoặc tương tự) trong kịch bản phát thanh cho phép cung cấp chức năng thao tác tín hiệu được mở rộng cho người dùng cuối. Nhiều chức năng hơn so với chỉ thay đổi kênh và điều chỉnh âm lượng phát lại được cung cấp.

Một khả năng để sử dụng kỹ thuật DE được mô tả vắn tắt ở trên. Nếu tín hiệu phát thanh, là tín hiệu trộn giảm cho SAOC, được tiêu chuẩn hóa về mức, ví dụ, theo R128, các chương trình khác nhau có độ to trung bình tương tự khi không có xử lý (SAOC) được áp dụng (hoặc mô tả kết xuất tương tự như mô tả trộn giảm). Tuy nhiên, khi một vài xử lý (SAOC) được áp dụng, tín hiệu đầu ra khác với tín hiệu trộn giảm mặc định và độ to của tín hiệu đầu ra có thể khác với độ to của tín hiệu trộn giảm mặc định. Từ quan điểm của người dùng cuối, điều này có thể dẫn đến tình huống trong đó độ to tín hiệu đầu ra giữa các kênh hoặc các chương trình có thể một lần nữa có bước nhảy không mong muốn hoặc sự chênh lệch. Nói cách khác, lợi ích của việc tiêu chuẩn hóa được áp dụng bởi các đài truyền hình bị mất một phần.

Vấn đề này không chỉ cụ thể cho SAOC hoặc cho kịch bản DE, mà có thể cũng xảy ra với các khái niệm mã hóa âm thanh khác mà cho phép người dùng cuối tương tác với nội dung. Tuy nhiên, trong nhiều trường hợp nó không gây ra bất kỳ tác hại nào nếu tín hiệu đầu ra có độ to khác với trộn giảm mặc định.

Như được nêu ở trên, tổng độ to của chương trình tín hiệu đầu vào âm thanh nên bằng mức đã định với độ lệch được phép nhỏ. Tuy nhiên, như đã nêu, điều này dẫn đến những vấn đề đáng kể khi việc kết xuất âm thanh được thực hiện, như việc kết xuất âm thanh có thể có tác động đáng kể trên độ to tổng thể/tổng độ to của tín hiệu đầu vào âm thanh nhận được. Tuy nhiên, mặc dù việc kết xuất cảnh được thực hiện, tổng độ to của tín hiệu nhận được sẽ vẫn như cũ.

Một cách tiếp cận sẽ là ước lượng độ to của tín hiệu trong khi tín hiệu đang được phát, và với khái niệm tích hợp theo thời gian thích hợp, việc ước lượng có thể hội tụ đến độ to trung bình đúng sau một thời gian. Tuy nhiên thời gian cần thiết cho việc hội tụ vẫn là vấn đề từ quan điểm của người dùng cuối. Khi việc ước lượng độ lớn thay đổi ngay cả khi không có thay đổi nào được áp dụng trên các tín hiệu, phân bù thay đổi độ lớn cũng phải phản ứng và thay đổi trạng thái của nó. Điều này sẽ dẫn đến tín hiệu đầu ra cùng với sự thay đổi theo thời gian độ to trung bình, mà có thể bị coi là khá khó chịu.

Fig.6 minh họa trạng thái ước lượng độ to tín hiệu đầu ra khi thay đổi độ to. Ngoài ra, ước lượng độ to tín hiệu đầu ra dựa trên tín hiệu được mô tả, mà minh họa hiệu quả của giải pháp như vừa mô tả. Việc ước lượng tiếp cận ước lượng chính xác khá chậm. Thay vì ước lượng độ to tín hiệu đầu ra dựa trên tín hiệu, ước lượng độ to tín hiệu đầu ra có thông tin, mà xác định ngay lập tức độ to tín hiệu đầu ra một cách chính xác sẽ được ưu tiên hơn.

Cụ thể, trên Fig.6, đầu vào người dùng, ví dụ, mức đối tượng hội thoại, thay đổi ngay tại thời điểm T bằng cách tăng về giá trị. Mức tín hiệu đầu ra đúng và độ to tương ứng thay đổi ngay tại cùng thời điểm. Khi việc ước lượng độ to tín hiệu đầu ra được thực hiện từ tín hiệu đầu ra với một số thời gian tích hợp theo thời gian, việc ước lượng này sẽ thay đổi từ từ và đạt đến giá trị chính xác sau độ trễ nhất định. Trong suốt độ trễ này, các giá trị ước lượng đang thay đổi và có thể không được sử dụng để xử lý thêm nữa tín hiệu đầu ra, ví dụ, để hiệu chỉnh mức độ to.

Như đã nêu ở trên, mong muốn có ước lượng chính xác độ to trung bình đầu ra hoặc sự thay đổi về độ to trung bình mà không có độ trễ và khi chương trình không thay đổi hoặc cảnh kết xuất không được thay đổi, ước lượng độ to trung bình cũng cần giữ nguyên trạng thái tĩnh. Nói cách khác, khi một số sự bù thay đổi độ to được áp dụng, tham số bù cần thay đổi chỉ khi chương trình thay đổi hoặc có một số tương tác người dùng.

Trạng thái mong muốn được minh họa trong phần minh họa thấp nhất của Fig.6 (ước lượng độ to tín hiệu đầu ra có thông tin). Ước lượng độ to tín hiệu đầu ra sẽ thay đổi ngay lập tức khi đầu vào người dùng thay đổi.

Fig.2 minh họa bộ mã hóa theo một phương án.

Bộ mã hóa bao gồm bộ phận mã hóa dựa trên đối tượng 210 để mã hóa nhiều tín hiệu đối tượng âm thanh để thu được tín hiệu âm thanh được mã hóa bao gồm nhiều tín hiệu đối tượng âm thanh.

Hơn nữa, bộ mã hóa bao gồm bộ phận mã hóa độ to đối tượng 220 để mã hóa thông tin độ to về tín hiệu đối tượng âm thanh. Thông tin độ to bao gồm một hoặc nhiều giá trị độ to, trong đó mỗi một hoặc nhiều giá trị độ to phụ thuộc vào một hoặc nhiều tín hiệu đối tượng âm thanh.

Theo phương án, mỗi tín hiệu đối tượng âm thanh của tín hiệu âm thanh được mã hóa được ấn định cho chính xác một nhóm trong số hai hoặc nhiều hơn hai nhóm, trong đó mỗi nhóm trong số hai hoặc nhiều hơn hai nhóm bao gồm một hoặc nhiều tín hiệu đối tượng âm thanh của tín hiệu âm thanh được mã hóa. Bộ phận mã hóa độ to đối tượng 220 được tạo cấu hình để xác định một hoặc nhiều giá trị độ to của thông tin độ to bằng cách xác định giá trị độ to của mỗi nhóm trong số hai hoặc nhiều hơn hai nhóm, trong đó giá trị độ to đã nêu của nhóm đã nêu cho biết tổng độ to ban đầu của một hoặc nhiều tín hiệu đối tượng âm thanh của nhóm đã nêu.

Fig.1 minh họa bộ giải mã tạo ra tín hiệu đầu ra âm thanh bao gồm một hoặc nhiều kênh đầu ra âm thanh theo một phương án.

Bộ giải mã bao gồm giao diện nhận 110 để nhận tín hiệu đầu vào âm thanh bao gồm nhiều tín hiệu đối tượng âm thanh, để nhận thông tin độ to trên các tín hiệu đối tượng âm thanh, và để nhận thông tin kết xuất cho biết một hoặc nhiều tín hiệu đối tượng âm thanh hoặc sẽ được khuếch đại hoặc được làm suy giảm.

Hơn nữa, bộ giải mã bao gồm bộ xử lý tín hiệu 120 tạo ra một hoặc nhiều kênh đầu ra âm thanh của tín hiệu đầu ra âm thanh. Bộ xử lý tín hiệu 120 được tạo cấu hình để xác định giá trị bù độ to phụ thuộc vào thông tin độ to và phụ thuộc vào thông tin kết xuất. Ngoài ra, bộ xử lý tín hiệu 120 được tạo cấu hình để tạo ra một hoặc nhiều kênh đầu ra âm thanh của tín hiệu đầu ra âm thanh từ tín hiệu đầu vào âm thanh phụ thuộc vào thông tin kết xuất và phụ thuộc vào giá trị bù độ to.

Theo một phương án, bộ xử lý tín hiệu 110 được tạo cấu hình để tạo ra một hoặc nhiều kênh đầu ra âm thanh của tín hiệu đầu ra âm thanh từ tín hiệu đầu vào âm thanh phụ thuộc vào thông tin kết xuất và phụ thuộc vào giá trị bù độ to, sao cho độ to của tín hiệu đầu ra âm thanh bằng với độ to của tín hiệu đầu vào âm thanh, hoặc sao cho độ to của tín hiệu đầu ra âm thanh gần hơn với độ to của tín hiệu đầu vào âm thanh so với độ to của tín hiệu âm thanh được biến đổi mà sẽ có được từ việc biến đổi tín hiệu đầu vào âm thanh bằng các khuếch đại hoặc làm suy giảm các tín hiệu đối tượng âm thanh của tín hiệu đầu vào âm thanh theo thông tin kết xuất.

Theo phương án khác, mỗi tín hiệu đối tượng âm thanh của tín hiệu đầu vào âm thanh được ấn định cho chính xác một nhóm trong số hai hoặc nhiều hơn hai nhóm, trong đó mỗi nhóm trong số hai hoặc nhiều hơn hai nhóm bao gồm một hoặc nhiều tín hiệu đối tượng âm thanh của tín hiệu đầu vào âm thanh.

Trong phương án này, giao diện nhận 110 được tạo cấu hình để nhận giá trị độ to từ mỗi nhóm trong số hai hoặc nhiều hơn hai nhóm như thông tin độ to, trong đó giá trị độ to đã nêu cho biết tổng độ to ban đầu của một hoặc nhiều tín hiệu đối tượng âm thanh của nhóm đã nêu. Hơn nữa, giao diện nhận 110 được tạo cấu hình để nhận thông tin kết xuất cho biết ít nhất một nhóm trong số hai hoặc nhiều hơn hai nhóm hoặc một hoặc nhiều tín hiệu đối tượng âm thanh của nhóm đã nêu sẽ được khuếch đại hoặc sẽ được làm suy giảm bằng cách cho biết tổng độ to được biến đổi của một hoặc nhiều tín hiệu đối tượng âm thanh của nhóm đã nêu. Hơn nữa, trong phương án này, bộ xử lý tín hiệu 120 được tạo cấu hình để xác định giá trị bù độ to phụ thuộc vào tổng độ to được biến đổi của mỗi nhóm của ít nhất một nhóm đã nêu trong số hai hoặc nhiều hơn hai nhóm và phụ thuộc vào tổng độ to ban đầu của mỗi nhóm trong số hai hoặc nhiều hơn hai nhóm. Ngoài ra, bộ xử lý tín hiệu 120 được tạo cấu hình để tạo ra một hoặc nhiều kênh đầu ra âm thanh của tín hiệu đầu ra âm thanh từ tín hiệu đầu vào âm thanh phụ thuộc vào tổng độ to được biến đổi của mỗi nhóm trong số ít nhất một nhóm đã nêu trong số hai hoặc nhiều hơn hai nhóm và phụ thuộc vào giá trị bù độ to.

Trong các phương án cụ thể, ít nhất một nhóm trong số hai hoặc nhiều hơn hai nhóm bao gồm hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh.

Mối quan hệ trực tiếp tồn tại giữa năng lượng e_i của tín hiệu đối tượng âm thanh i và độ to L_i của tín hiệu đối tượng âm thanh i theo công thức:

$$L_i = c + 10 \log_{10} e_i, \quad e_i = 10^{(L_i - c)/10}$$

trong đó c là giá trị hằng số.

Các phương án được dựa trên các phát hiện sau đây: Các tín hiệu đối tượng âm thanh khác nhau của tín hiệu đầu vào âm thanh có thể có độ to khác nhau và do đó năng lượng khác nhau. Ví dụ, nếu người dùng muốn tăng độ to của một trong số các tín hiệu đối tượng âm thanh, thông tin kết xuất có thể được điều chỉnh tương ứng, và việc tăng độ to của tín hiệu đối tượng âm thanh này tăng năng lượng của đối tượng âm thanh này. Điều này sẽ dẫn đến độ to được tăng của tín hiệu đầu ra âm thanh. Để giữ tổng độ to không đổi, việc bù độ to được tiến hành. Nói cách khác, tín hiệu âm thanh được biến đổi mà sẽ thu được từ việc áp dụng thông tin kết xuất trên tín hiệu đầu vào âm thanh sẽ phải được điều chỉnh. Tuy nhiên, hiệu quả chính xác của việc khuếch đại một trong số các tín hiệu đối tượng âm thanh trên tổng độ to của tín hiệu âm thanh được biến đổi phụ thuộc vào độ to ban đầu của tín hiệu đối tượng âm thanh được khuếch đại, ví dụ, của tín hiệu đối tượng âm thanh mà độ to được tăng. Nếu độ to ban đầu của đối tượng này tương ứng với năng lượng, mà khá thấp, ảnh hưởng trên tổng độ to của tín hiệu đầu vào âm thanh sẽ là không đáng kể. Tuy nhiên, nếu độ to ban đầu của đối tượng này tương ứng với năng lượng, mà là khá cao, ảnh hưởng trên tổng độ to của tín hiệu đầu vào âm thanh sẽ là đáng kể.

Hai ví dụ có thể được xem xét. Trong cả hai ví dụ, tín hiệu đầu vào âm thanh bao gồm hai tín hiệu đối tượng âm thanh, và trong cả hai ví dụ, bằng cách áp dụng thông tin kết xuất, năng lượng của tín hiệu thứ nhất của đối tượng âm thanh được tăng lên 50%.

Trong ví dụ thứ nhất, tín hiệu đối tượng âm thanh thứ nhất góp 20% và tín hiệu đối tượng âm thanh thứ hai góp 80% vào tổng năng lượng của tín hiệu đầu vào âm thanh. Tuy nhiên, trong ví dụ thứ hai, đối tượng âm thanh thứ nhất, tín hiệu đối tượng âm thanh thứ nhất góp 40% và tín hiệu đối tượng âm thanh thứ hai góp 60% vào tổng năng lượng của tín hiệu đầu vào âm thanh. Trong cả hai ví dụ, những đóng góp này có

được từ thông tin độ to trên các tín hiệu đối tượng âm thanh, vì mối quan hệ trực tiếp tồn tại giữa độ to và năng lượng.

Trong ví dụ thứ nhất, việc tăng 50% năng lượng của đối tượng âm thanh thứ nhất dẫn đến tín hiệu âm thanh được biến đổi mà được tạo ra bằng cách áp dụng thông tin kết xuất trên tín hiệu đầu vào âm thanh có tổng năng lượng $1,5 \times 20\% + 80\% = 110\%$ năng lượng của tín hiệu đầu vào âm thanh.

Trong ví dụ thứ hai, việc tăng 50% năng lượng của đối tượng âm thanh thứ nhất dẫn đến tín hiệu âm thanh được biến đổi mà được tạo ra bằng cách áp dụng thông tin kết xuất trên tín hiệu đầu vào âm thanh có tổng năng lượng $1,5 \times 40\% + 60\% = 120\%$ năng lượng của tín hiệu đầu vào âm thanh.

Do đó, sau khi áp dụng thông tin kết xuất trên tín hiệu đầu vào âm thanh, trong ví dụ thứ nhất, tổng năng lượng của tín hiệu âm thanh được biến đổi phải được giảm chỉ 9% ($10/110$) để thu được năng lượng bằng nhau trong cả hai tín hiệu đầu vào âm thanh và tín hiệu đầu ra âm thanh, trong khi đó trong ví dụ thứ hai, tổng năng lượng của tín hiệu âm thanh được biến đổi phải được giảm 17% ($20/120$). Với mục đích này, giá trị bù độ to có thể được tính toán.

Ví dụ, giá trị bù độ to có thể là vô hướng mà được áp dụng trên tất cả các kênh đầu ra âm thanh của tín hiệu đầu ra âm thanh.

Theo phương án, bộ xử lý tín hiệu được tạo cấu hình để tạo ra tín hiệu âm thanh được biến đổi bằng cách biến đổi tín hiệu đầu vào âm thanh bằng cách khuếch đại hoặc làm suy giảm các tín hiệu đối tượng âm thanh của tín hiệu đầu vào âm thanh theo thông tin kết xuất. Ngoài ra, bộ xử lý tín hiệu được tạo cấu hình để tạo ra tín hiệu đầu ra âm thanh bằng cách áp dụng giá trị bù độ to trên tín hiệu âm thanh được biến đổi, sao cho độ to của tín hiệu đầu ra âm thanh bằng với độ to của tín hiệu đầu vào âm thanh, hoặc sao cho độ to của tín hiệu đầu ra âm thanh gần hơn với độ to của tín hiệu đầu vào âm thanh so với độ to của tín hiệu âm thanh được biến đổi.

Ví dụ, trong ví dụ thứ nhất ở trên, giá trị bù độ to lev ví dụ có thể được đặt thành giá trị $lev = 10/11$, và hệ số nhân $10/11$ có thể được áp dụng trên tất cả các kênh mà thu được từ việc kết xuất các kênh đầu vào âm thanh theo thông tin kết xuất.

Theo đó, ví dụ, trong ví dụ thứ hai ở trên, giá trị bù độ to lev ví dụ có thể được đặt thành giá trị $lev = 10/12 = 5/6$, và hệ số nhân $5/6$ có thể được áp dụng trên tất cả các kênh mà thu được từ việc kết xuất các kênh đầu vào âm thanh theo thông tin kết xuất.

Trong các phương án khác, mỗi tín hiệu đối tượng âm thanh có thể được ấn định cho một trong số nhiều nhóm, và giá trị độ to có thể được truyền cho mỗi nhóm cho biết giá trị tổng độ to của các tín hiệu đối tượng âm thanh của nhóm đã nêu. Nếu thông tin kết xuất chỉ rõ rằng năng lượng của một trong số các nhóm bị suy giảm hoặc khuếch đại, ví dụ, khuếch đại 50% như ở trên, mức tăng tổng năng lượng có thể được tính toán và giá trị bù độ to có thể được xác định như được mô tả ở trên.

Ví dụ, theo phương án, mỗi tín hiệu trong số các tín hiệu đối tượng âm thanh của tín hiệu đầu vào âm thanh được ấn định cho chính xác một nhóm trong chính xác hai nhóm là hai hoặc nhiều hơn hai nhóm. Mỗi tín hiệu trong số các tín hiệu đối tượng âm thanh của tín hiệu đầu vào âm thanh hoặc được ấn định cho nhóm đối tượng nổi của chính xác hai nhóm hoặc được ấn định cho nhóm đối tượng nền của nhóm chính xác. Giao diện nhận 110 được tạo cấu hình để nhận tổng độ to ban đầu của một hoặc nhiều tín hiệu đối tượng âm thanh của nhóm đối tượng nổi. Ngoài ra, giao diện nhận 110 được tạo cấu hình để nhận tổng độ to ban đầu của một hoặc nhiều tín hiệu đối tượng âm thanh của nhóm đối tượng nền. Hơn nữa, giao diện nhận 110 được tạo cấu hình để nhận thông tin kết xuất cho biết ít nhất một nhóm trong số chính xác hai nhóm hoặc một hoặc nhiều tín hiệu đối tượng âm thanh của ít nhất một nhóm đã nêu sẽ được khuếch đại hoặc sẽ được làm suy giảm bằng cách cho biết tổng độ to được biến đổi của một hoặc nhiều tín hiệu đối tượng âm thanh của nhóm đã nêu.

Trong phương án như vậy, bộ xử lý tín hiệu 120 được tạo cấu hình để xác định giá trị bù độ to phụ thuộc vào tổng độ to được biến đổi của mỗi nhóm trong số ít nhất một nhóm đã nêu, phụ thuộc vào tổng độ to ban đầu của một hoặc nhiều tín hiệu đối tượng âm thanh của nhóm đối tượng nổi, và phụ thuộc vào tổng độ to ban đầu của một hoặc nhiều tín hiệu đối tượng âm thanh của nhóm đối tượng nền. Ngoài ra, bộ xử lý tín hiệu 120 được tạo cấu hình để tạo ra một hoặc nhiều kênh đầu ra âm thanh của tín hiệu

đầu ra âm thanh từ tín hiệu đầu vào âm thanh phụ thuộc vào tổng độ to được biến đổi của mỗi nhóm trong số ít nhất một nhóm đã nêu và phụ thuộc vào giá trị bù độ to.

Theo một số phương án, mỗi tín hiệu trong số các tín hiệu đối tượng âm thanh được ấn định cho một trong ba hoặc nhiều hơn ba nhóm, và giao diện nhận có thể được tạo cấu hình để nhận giá trị độ to cho mỗi nhóm trong số ba hoặc nhiều hơn ba nhóm cho biết tổng độ to của tín hiệu đối tượng âm thanh của nhóm đã nêu.

Theo phương án, để xác định giá trị tổng độ to của hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh, ví dụ, giá trị năng lượng tương ứng với giá trị độ to được xác định cho mỗi tín hiệu đối tượng âm thanh, các giá trị năng lượng của tất cả các giá trị độ to được tính tổng vào để thu được tổng năng lượng, và giá trị độ to tương ứng với tổng năng lượng được xác định như giá trị tổng độ to của hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh. Ví dụ, công thức

$$L_i = c + 10 \log_{10} e_i, \quad e_i = 10^{(L_i - c)/10}$$

có thể được sử dụng.

Trong một số phương án, các giá trị độ to được truyền cho mỗi tín hiệu đối tượng âm thanh, hoặc mỗi tín hiệu đối tượng âm thanh được ấn định cho một nhóm hoặc hai hoặc nhiều hơn hai nhóm, trong đó với mỗi nhóm, giá trị độ to được truyền dẫn.

Tuy nhiên, trong một số phương án, với một hoặc nhiều tín hiệu đối tượng âm thanh hoặc với một hoặc nhiều nhóm bao gồm các tín hiệu đối tượng âm thanh, không có giá trị độ to nào được truyền. Thay vào đó, bộ giải mã có thể, ví dụ, giả sử rằng các tín hiệu đối tượng âm thanh hoặc nhóm của các tín hiệu đối tượng âm thanh, mà không có giá trị độ to nào được truyền, có giá trị độ to được xác định trước. Ví dụ, bộ giải mã có thể dựa trên tất cả các xác định khác dựa trên giá trị độ to được xác định trước này.

Theo sáng chế, giao diện nhận 110 được tạo cấu hình để nhận tín hiệu trộn giảm bao gồm một hoặc nhiều kênh trộn giảm làm tín hiệu đầu vào âm thanh, trong đó một hoặc nhiều kênh trộn giảm bao gồm các tín hiệu đối tượng âm thanh, và trong đó số tín hiệu đối tượng âm thanh nhỏ hơn số lượng của một hoặc nhiều kênh trộn giảm. Giao diện nhận 110 được tạo cấu hình để nhận thông tin trộn giảm cho biết tín hiệu đối tượng âm thanh được trộn như thế nào trong một hoặc nhiều kênh trộn giảm. Ngoài ra,

bộ xử lý tín hiệu 120 được tạo cấu hình để tạo ra một hoặc nhiều kênh đầu ra âm thanh của tín hiệu đầu ra âm thanh từ tín hiệu đầu vào âm thanh phụ thuộc vào thông tin trộn giảm, phụ thuộc vào thông tin kết xuất và phụ thuộc vào giá trị bù độ to. Trong phương án cụ thể, bộ xử lý tín hiệu 120 ví dụ có thể được tạo cấu hình để tính toán giá trị bù độ to phụ thuộc vào thông tin trộn giảm.

Ví dụ, thông tin trộn giảm có thể là ma trận trộn giảm. Trong các phương án, bộ giải mã có thể là bộ giải mã SAOC. Trong các phương án như vậy, giao diện nhận 110 có thể, ví dụ còn được tạo cấu hình để nhận thông tin hiệp phương sai, ví dụ, ma trận hiệp phương sai như được mô tả ở trên.

Đối với thông tin kết xuất cho biết một hoặc nhiều tín hiệu đối tượng âm thanh hoặc sẽ được khuếch đại hoặc sẽ được làm suy giảm, cần lưu ý rằng, ví dụ, thông tin mà cho biết một hoặc nhiều tín hiệu đối tượng âm thanh sẽ được khuếch đại hoặc làm suy giảm như thế nào, là thông tin kết xuất. Ví dụ, ma trận kết xuất **R**, ví dụ, ma trận kết xuất SAOC, là thông tin kết xuất.

Fig.3 minh họa hệ thống mã hóa tín hiệu đối tượng âm thanh và tạo ra tín hiệu đầu ra âm thanh theo một phương án,

Hệ thống bao gồm bộ mã hóa 310 theo một trong số các phương án được mô tả ở trên để mã hóa nhiều tín hiệu đối tượng âm thanh để thu được tín hiệu âm thanh được mã hóa bao gồm nhiều tín hiệu đối tượng âm thanh.

Ngoài ra, hệ thống bao gồm bộ giải mã 320 theo một trong số các phương án được mô tả ở trên để tạo ra tín hiệu đầu ra âm thanh bao gồm một hoặc nhiều kênh đầu ra âm thanh. Bộ giải mã được tạo cấu hình để nhận tín hiệu âm thanh được mã hóa làm tín hiệu đầu vào âm thanh và thông tin độ to. Ngoài ra, bộ giải mã 320 còn được tạo cấu hình để nhận thông tin kết xuất. Ngoài ra, bộ giải mã 320 được tạo cấu hình để xác định giá trị bù độ to phụ thuộc vào thông tin độ to và phụ thuộc vào thông tin kết xuất. Hơn nữa, bộ giải mã 320 được tạo cấu hình để tạo ra một hoặc nhiều kênh đầu ra âm thanh của tín hiệu đầu ra âm thanh từ tín hiệu đầu vào âm thanh phụ thuộc vào thông tin kết xuất và phụ thuộc vào giá trị bù độ to.

Fig.7 minh họa việc ước lượng độ to có thông tin theo một phương án. Ở phía bên trái của đường truyền 730, các thành phần của bộ mã hóa mã hóa âm thanh dựa trên đối tượng được minh họa. Cụ thể, bộ phận mã hóa dựa trên đối tượng 710 ("bộ mã hóa âm thanh dựa trên đối tượng") và bộ phận mã hóa độ to đối tượng 720 được minh họa ("ước lượng độ to đối tượng").

Chính bản thân đường truyền 730 chứa thông tin độ to **L**, thông tin trộn giảm **D** và đầu ra của bộ mã hóa âm thanh dựa trên đối tượng 710 **B**.

Ở phía bên phải của đường truyền 730, các thành phần của bộ xử lý tín hiệu của bộ giải mã mã hóa âm thanh dựa trên đối tượng được minh họa. Giao diện nhận của bộ giải mã không được minh họa. Bộ ước lượng độ to đầu ra 740 và bộ phận giải mã âm thanh dựa trên đối tượng 750 được mô tả. Bộ ước lượng độ to đầu ra 740 có thể được tạo cấu hình để xác định giá trị bù độ to. Bộ phận giải mã âm thanh dựa trên đối tượng 750 có thể được tạo cấu hình để xác định tín hiệu âm thanh được biến đổi từ tín hiệu âm thanh, là đầu vào vào bộ giải mã, bằng cách áp dụng thông tin kết xuất **R**. Việc áp dụng giá trị bù độ to trên tín hiệu âm thanh được biến đổi để bù sự thay đổi tổng độ to gây ra bởi sự kết xuất không được thể hiện trên Fig.7.

Đầu vào vào bộ mã hóa chứa các đối tượng đầu vào **S** ở mức tối thiểu. Hệ thống ước lượng độ to của mỗi đối tượng (hoặc một số thông tin liên quan đến độ to khác, chẳng hạn như các năng lượng đối tượng), ví dụ, bởi bộ phận mã hóa độ to đối tượng 720, và thông tin **L** này được truyền dẫn và/hoặc được lưu trữ. (Cũng có thể là độ to của các đối tượng được cung cấp làm đầu vào vào hệ thống, và bước ước lượng trong hệ thống có thể được bỏ qua).

Trong phương án của Fig.7, bộ giải mã nhận ít nhất là thông tin độ to đối tượng và, ví dụ, thông tin kết xuất **R** mô tả việc trộn các đối tượng vào thành tín hiệu đầu ra. Dựa trên điều này, ví dụ, bộ ước lượng độ to đầu ra 740, ước lượng độ to của tín hiệu đầu ra và cung cấp thông tin này như đầu ra của nó.

Thông tin trộn giảm **D** có thể được cung cấp làm thông tin kết xuất, trong trường hợp việc ước lượng độ to cung cấp ước lượng của độ to tín hiệu trộn giảm. Cũng có thể cung cấp thông tin trộn giảm làm đầu vào vào ước lượng độ to đối tượng, và truyền dẫn và/hoặc lưu trữ nó cùng với thông tin độ to đối tượng. Việc ước lượng độ to đầu ra

có thể sau đó ước lượng đồng thời độ to của tín hiệu trộn giảm và đầu ra được kết xuất và cung cấp hai giá trị này hoặc sự chênh lệch của chúng như thông tin độ to đầu ra. Giá trị chênh lệch (hoặc nghịch đảo của nó) mô tả giá trị bù cần thiết mà nên được áp dụng trên tín hiệu đầu ra được kết xuất để làm cho độ to của nó tương tự như độ to của tín hiệu trộn giảm. Thông tin độ to đối tượng có thể còn bao gồm thông tin về các hệ số tương quan giữa các đối tượng khác nhau và thông tin tương quan này có thể được sử dụng trong việc ước lượng độ to đầu ra cho việc ước lượng chính xác hơn.

Dưới đây, phương án được ưu tiên cho ứng dụng cải thiện âm thanh hội thoại được mô tả.

Trong ứng dụng cải thiện âm thanh hội thoại, như được mô tả ở trên, các tín hiệu đối tượng âm thanh đầu vào được nhóm lại và được trộn giảm từng phần để tạo ra hai siêu đối tượng, FGO và BGO, mà sau đó có thể được tính tổng không đáng kể để thu được tín hiệu trộn giảm cuối cùng.

Theo mô tả của SAOC [SAOC], N tín hiệu đối tượng đầu vào được biểu diễn dưới dạng ma trận $\mathbf{S}$ có kích thước $N \times N_{mẫu}$, và thông tin trộn giảm dưới dạng ma trận $\mathbf{D}$ có kích thước $M \times N$. Các tín hiệu trộn giảm sau đó có thể thu được là $\mathbf{X} = \mathbf{D}\mathbf{S}$.

Thông tin trộn giảm $\mathbf{D}$ lúc này có thể được chia thành hai phần:

$$\mathbf{D} = \mathbf{D}_{FGO} + \mathbf{D}_{BGO}$$

cho các siêu đối tượng.

Mỗi cột của ma trận $\mathbf{D}$ tương ứng với tín hiệu đối tượng âm thanh ban đầu, hai ma trận trộn giảm thành phần có thể thu được bằng cách thiết lập các cột, mà tương ứng với siêu đối tượng khác thành không (giả sử rằng không có đối tượng ban đầu nào có thể có mặt trong cả hai siêu đối tượng). Nói cách khác, các cột tương ứng với siêu đối tượng BGO được đặt thành không trong $\mathbf{D}_{FGO}$, và ngược lại.

Ma trận trộn giảm mới này mô tả cách hai siêu đối tượng có thể thu được từ các đối tượng đầu vào, cụ thể:

$$\mathbf{S}_{FGO} = \mathbf{D}_{FGO}\mathbf{S} \quad \text{và} \quad \mathbf{S}_{BGO} = \mathbf{D}_{BGO}\mathbf{S},$$

và phép trộn giảm thực tế được đơn giản hóa thành:

$$\mathbf{X} = \mathbf{S}_{FGO} + \mathbf{S}_{BGO}.$$

Cũng có thể được xem xét rằng bộ giải mã đối tượng (ví dụ SAOC) cố gắng khôi phục các siêu đối tượng:

$$\tilde{\mathbf{S}}_{FGO} \approx \mathbf{S}_{FGO} \quad \text{và} \quad \tilde{\mathbf{S}}_{BGO} \approx \mathbf{S}_{BGO},$$

và kết xuất riêng DE có thể được viết dưới dạng tổ hợp của sự khôi phục hai siêu đối tượng này:

$$\mathbf{Y} = g_{FGO} \mathbf{S}_{FGO} + g_{BGO} \mathbf{S}_{BGO} \approx g_{FGO} \tilde{\mathbf{S}}_{FGO} + g_{BGO} \tilde{\mathbf{S}}_{BGO}.$$

Sự ước lượng độ to đối tượng nhận hai siêu đối tượng $\mathbf{S}_{FGO}$ và $\mathbf{S}_{BGO}$ làm đầu vào và ước lượng độ to của mỗi siêu đối tượng trong số chúng: L_{FGO} là độ to (tổng/toàn phần) của $\mathbf{S}_{FGO}$, và L_{BGO} là độ to (tổng số/toàn phần) của $\mathbf{S}_{BGO}$. Các giá trị độ to này được truyền dẫn và/hoặc được lưu trữ.

Thay vào đó, bằng cách sử dụng một trong các siêu đối tượng, ví dụ FGO, làm tham chiếu, có thể tính toán sự khác biệt về độ to của hai đối tượng này, ví dụ:

$$\Delta L_{FGO} = L_{BGO} - L_{FGO}.$$

Các giá trị đơn lẻ này sau đó được truyền dẫn và/hoặc được lưu trữ.

Fig.8 minh họa bộ mã hóa theo phương án khác. Bộ mã hóa của Fig.8 bao gồm bộ trộn giảm đối tượng 811 và bộ ước lượng thông tin phụ đối tượng 812. Ngoài ra, bộ mã hóa của Fig.8 còn bao gồm bộ phận mã hóa độ to đối tượng 820. Hơn nữa, bộ mã hóa của Fig.8 bao gồm bộ trộn siêu đối tượng âm thanh 805.

Bộ mã hóa của Fig.8 sử dụng các siêu đối tượng âm thanh trung gian làm đầu vào cho ước lượng độ to đối tượng. Trong các phương án, bộ mã hóa của Fig.8 có thể được tạo cấu hình để tạo ra hai siêu đối tượng âm thanh. Trong các phương án khác, bộ mã hóa của Fig.8 có thể được tạo cấu hình để tạo ra ba hoặc nhiều hơn ba siêu đối tượng âm thanh.

Ngoài ra, các khái niệm được đề xuất cung cấp đặc tính mới mà bộ mã hóa có thể, ví dụ, ước lượng độ to trung bình của tất cả các đối tượng đầu vào. Các đối tượng

có thể, ví dụ, được trộn thành tín hiệu trộn giảm mà được truyền dẫn. Các khái niệm được đề xuất hơn nữa còn cung cấp đặc tính mới mà độ to đối tượng và thông tin trộn giảm có thể, ví dụ, được chứa trong thông tin phụ mã hóa đối tượng mà được truyền dẫn.

Bộ giải mã có thể, ví dụ, sử dụng thông tin phụ mã hóa đối tượng cho việc tách (ảo) các đối tượng và tổ hợp lại các đối tượng sử dụng thông tin kết xuất.

Hơn nữa, các khái niệm được đề xuất cung cấp đặc tính mới mà thông tin trộn giảm có thể được sử dụng để ước lượng độ to của tín hiệu trộn giảm mặc định, thông tin kết xuất và độ to đối tượng nhận được có thể được sử dụng để ước lượng độ to trung bình của tín hiệu đầu ra, và/hoặc sự thay đổi độ to có thể được ước lượng từ hai giá trị này. Hoặc thông tin kết xuất và thông tin trộn giảm có thể được sử dụng để ước lượng sự thay đổi độ to từ trộn giảm mặc định, đặc tính mới khác nữa của các khái niệm được đề xuất.

Ngoài ra, các khái niệm được đề xuất cung cấp đặc tính mới là đầu ra bộ giải mã có thể được biến đổi để bù cho sự thay đổi về độ to để mà độ to trung bình của tín hiệu được biến đổi khớp với độ to trung bình của trộn giảm mặc định.

Phương án cụ thể liên quan đến SAOC-DE được minh họa trên Fig.9. Hệ thống nhận các tín hiệu đối tượng âm thanh đầu vào, thông tin trộn giảm, và thông tin của việc nhóm các đối tượng thành các siêu đối tượng. Dựa trên điều này, bộ trộn siêu đối tượng âm thanh 905 tạo thành hai siêu đối tượng S_{FGO} và S_{BGO} . Có thể phân tín hiệu mà được xử lý với SAOC không cấu thành tín hiệu trộn vẹn. Ví dụ, trong cấu hình kênh 5.1, SAOC có thể được triển khai trên tập hợp con của các kênh, như trên các kênh phía trước (bên trái, bên phải và trung tâm), trong khi các kênh khác (xung quanh bên trái, xung quanh bên phải, và các hiệu ứng tần số thấp) được định tuyến xung quanh, SAOC (đường vòng) và được phân phối như vậy. Các kênh này không được xử lý bởi SAOC được kí hiệu với X_{BYPASS} . Các kênh đường vòng có thể cần được cung cấp cho bộ mã hóa cho việc ước lượng chính xác hơn thông tin độ to.

Các kênh đường vòng có thể được xử lý theo nhiều cách khác nhau.

Ví dụ, các kênh đường vòng, có thể, ví dụ, tạo ra siêu đối tượng độc lập. Điều này cho phép xác định kết xuất để cả ba siêu đối tượng được định tỷ lệ độc lập.

Hoặc chẳng hạn, các kênh đường vòng ví dụ có thể được kết hợp với một trong hai siêu đối tượng khác. Các thiết lập kết xuất của siêu đối tượng này cũng kiểm soát phần kênh đường vòng. Ví dụ, trong kịch bản cải thiện âm thanh hội thoại, có thể có ý nghĩa để kết hợp các kênh đường vòng với siêu đối tượng nền: $\mathbf{X}_{BGO} = \mathbf{S}_{BGO} + \mathbf{X}_{BYPASS}$.

Hoặc ví dụ, các kênh đường vòng ví dụ có thể được bỏ qua.

Theo các phương án, bộ phận mã hóa dựa trên đối tượng 210 của bộ mã hóa được tạo cấu hình để nhận các tín hiệu đối tượng âm thanh, trong đó mỗi tín hiệu đối tượng âm thanh được ấn định cho chính xác một nhóm trong số chính xác hai nhóm, trong đó mỗi nhóm trong số chính xác hai nhóm bao gồm một hoặc nhiều tín hiệu đối tượng âm thanh. Ngoài ra, bộ phận mã hóa dựa trên đối tượng 210 được tạo cấu hình để trộn giảm các tín hiệu đối tượng âm thanh, có trong chính xác hai nhóm, để thu được tín hiệu trộn giảm bao gồm một hoặc nhiều kênh âm thanh trộn giảm làm tín hiệu âm thanh được mã hóa, trong đó số lượng của một hoặc nhiều kênh trộn giảm nhỏ hơn số lượng của các tín hiệu đối tượng âm thanh có trong chính xác hai nhóm. Bộ phận mã hóa độ to đối tượng 220 được ấn định để nhận một hoặc nhiều tín hiệu đối tượng âm thanh đường vòng thêm, trong đó mỗi tín hiệu trong số một hoặc nhiều tín hiệu đối tượng âm thanh đường vòng thêm được ấn định cho nhóm thứ ba, trong đó mỗi tín hiệu trong số một hoặc nhiều tín hiệu đối tượng âm thanh đường vòng thêm không có trong nhóm thứ nhất và không có trong nhóm thứ hai, trong đó bộ phận mã hóa dựa trên đối tượng 210 được tạo cấu hình để không trộn giảm một hoặc nhiều tín hiệu đối tượng âm thanh đường vòng thêm trong tín hiệu trộn giảm.

Trong phương án, bộ phận mã hóa độ to đối tượng 220 được tạo cấu hình để xác định giá trị độ to thứ nhất, giá trị độ to thứ hai và giá trị độ to thứ ba của thông tin độ to, giá trị độ to thứ nhất cho biết tổng độ to của một hoặc nhiều tín hiệu đối tượng âm thanh của nhóm thứ nhất, và giá trị độ to thứ hai cho biết tổng độ to của một hoặc nhiều tín hiệu đối tượng âm thanh của nhóm thứ hai, và giá trị độ to thứ ba cho biết tổng độ to của một hoặc nhiều tín hiệu đối tượng âm thanh đường vòng thêm của nhóm thứ ba. Trong phương án khác, bộ phận mã hóa độ to đối tượng 220 được tạo

cấu hình để xác định giá trị độ to thứ nhất và giá trị độ to thứ hai của thông tin độ to, giá trị độ to thứ nhất cho biết tổng độ to của một hoặc nhiều tín hiệu đối tượng âm thanh của nhóm thứ nhất, và giá trị độ to thứ hai cho biết tổng độ to của một hoặc nhiều tín hiệu đối tượng âm thanh của nhóm thứ hai và của một hoặc nhiều tín hiệu đối tượng âm thanh đường vòng thêm của nhóm thứ ba.

Theo phương án, giao diện nhận 110 của bộ giải mã được tạo cấu hình để nhận tín hiệu trộn giảm. Ngoài ra, giao diện nhận 110 được tạo cấu hình để nhận một hoặc nhiều tín hiệu đối tượng âm thanh đường vòng thêm, trong đó một hoặc nhiều tín hiệu đối tượng âm thanh đường vòng thêm không được trộn trong tín hiệu trộn giảm. Hơn nữa, giao diện nhận 110 được tạo cấu hình để nhận thông tin độ to cho biết thông tin về độ to của các tín hiệu đối tượng âm thanh mà được trộn trong tín hiệu trộn giảm và cho biết thông tin về độ to của một hoặc nhiều tín hiệu đối tượng âm thanh đường vòng thêm mà được trộn trong tín hiệu trộn giảm. Hơn nữa, bộ xử lý tín hiệu 120 được tạo cấu hình để xác định giá trị bù độ to phụ thuộc vào thông tin về độ to của các tín hiệu đối tượng âm thanh mà được trộn trong tín hiệu trộn giảm, và phụ thuộc vào thông tin về độ to của một hoặc nhiều tín hiệu đối tượng âm thanh đường vòng thêm mà không được trộn trong tín hiệu trộn giảm.

Fig.9 minh họa bộ mã hóa và bộ giải mã theo phương án liên quan SAOC-DE, mà bao gồm các kênh đường vòng. Ngoài ra, bộ mã hóa của Fig.9 bao gồm bộ mã hóa SAOC 902.

Trong phương án của Fig.9, tổ hợp có thể có của các kênh đường vòng với các siêu đối tượng khác diễn ra trong hai khối "chứa đường vòng" 913, 914, tạo ra các siêu đối tượng X_{FGO} và X_{BGO} với những phần được xác định từ các kênh đường vòng được chứa.

Độ to cảm nhận được L_{BYPASS} , L_{FGO} , và L_{BGO} của cả hai siêu đối tượng này được ước lượng trong các bộ phận ước lượng độ to 921, 922, 923. Thông tin độ to này sau đó được biến đổi thành mã hóa tương thích trong bộ ước lượng thông tin độ to siêu đối tượng 925 và sau đó được truyền dẫn và/hoặc được lưu trữ.

Bộ mã hóa và bộ giải mã SAOC thực tế hoạt động như mong đợi để trích xuất thông tin phụ đối tượng từ các đối tượng, tạo ra tín hiệu trộn giảm X , và truyền dẫn và/hoặc lưu trữ thông tin đến bộ giải mã. Các kênh đường vòng có thể có được truyền dẫn và/hoặc lưu trữ cùng với thông tin khác đến bộ giải mã.

Bộ giải mã SAOC-DE 945 nhận giá trị khuếch đại "khuếch đại hội thoại" làm đầu vào người dùng. Dựa trên đầu vào này và thông tin trộn giảm nhận được, bộ giải mã SAOC 945 xác định thông tin kết xuất. Bộ giải mã SAOC 945 sau đó tạo ra cảnh đầu ra được kết xuất làm tín hiệu Y . Bên cạnh đó, nó tạo ra hệ số khuếch đại (và giá trị độ trễ) mà cần được áp dụng trên các tín hiệu đường vòng có thể có X_{BYPASS} .

Bộ phận "chứa đường vòng" 955 nhận thông tin này cùng với cảnh đầu ra được kết xuất và các tín hiệu đường vòng và tạo ra tín hiệu cảnh đầu ra đầy đủ. Bộ giải mã SAOC 945 cũng tạo ra tập hợp các giá trị độ khuếch đại siêu đối tượng, lượng giá trị này phụ thuộc vào việc nhóm siêu đối tượng và dạng thông tin độ to mong muốn.

Các giá trị độ khuếch đại được cung cấp cho bộ ước lượng độ to hỗn hợp 960 mà cũng nhận thông tin độ to siêu đối tượng từ bộ mã hóa.

Bộ ước lượng độ to hỗn hợp 960 sau đó có thể xác định thông tin độ to mong muốn, mà có thể bao gồm, nhưng không giới hạn, độ to của tín hiệu trộn giảm, độ to của cảnh đầu ra được kết xuất, và/hoặc sự chênh lệch độ to giữa tín hiệu trộn giảm và cảnh đầu ra được kết xuất.

Trong một số phương án, tự bản thân thông tin độ to là đủ, trong khi trong các phương án khác, mong muốn xử lý đầu ra đầy đủ phụ thuộc vào thông tin độ to được xác định. Việc xử lý này ví dụ có thể là sự bù của sự chênh lệch có thể có bất kỳ trong độ to giữa tín hiệu trộn giảm và cảnh đầu ra được kết xuất. Việc xử lý như vậy, ví dụ, bởi bộ phận xử lý độ to 970, sẽ tạo cảm giác trong kịch bản phát thanh, vì nó sẽ giảm sự thay đổi về độ to tín hiệu nhận thấy được không phụ thuộc vào tương tác người dùng (thiết lập của đầu vào "độ khuếch đại âm thanh hội thoại").

Việc xử lý liên quan đến độ to trong phương án cụ thể này bao gồm nhiều đặc tính mới. Ngoài ra, FGO, BGO và các kênh đường vòng có thể có được trộn trước thành cấu hình kênh cuối cùng để việc trộn giảm có thể được thực hiện bằng cách đơn

giảm bổ sung hai tín hiệu được trộn trước với nhau (ví dụ, các hệ số ma trận trộn giảm của 1) mà tạo thành đặc tính mới. Ngoài ra, như một đặc tính mới nữa, độ to trung bình của FGO và BGO được ước lượng, và sự chênh lệch được tính toán. Hơn nữa, các đối tượng được trộn thành tín hiệu trộn giảm mà được truyền dẫn. Ngoài ra, như một đặc tính mới nữa, thông tin sự chênh lệch độ to được chứa để thông tin phụ mà được truyền dẫn. Hơn nữa, bộ giải mã sử dụng thông tin phụ cho việc phân tách (ảo) các đối tượng và tổ hợp lại các đối tượng sử dụng thông tin kết xuất mà dựa trên thông tin trộn giảm và độ khuếch đại sự biến đổi đầu vào người dùng. Ngoài ra, như một đặc tính mới khác, bộ giải mã sử dụng độ khuếch đại biến đổi và thông tin độ to được truyền dẫn để ước lượng sự thay đổi về độ to trung bình của đầu ra hệ thống so với trộn giảm mặc định.

Sau đây, mô tả chính thức của các phương án được đưa ra.

Giả sử rằng các giá trị độ to đối tượng thể hiện tương tự như lôgarit của các giá trị năng lượng khi tính tổng các đối tượng, tức là, các giá trị độ to phải được biến đổi thành miền tuyến tính, được cộng vào ở đây, và cuối cùng được biến đổi trở lại miền lôgarit. Việc thúc đẩy điều này qua định nghĩa của số đo độ to BS.1770 ngay sau đây sẽ được trình bày (để đơn giản, số lượng của các kênh được đặt thành 1, nhưng nguyên tắc tương tự có thể được áp dụng trên các tín hiệu đa kênh với việc tính tổng tương thích trên các kênh).

Độ to của tín hiệu được lọc K z_i thứ i với năng lượng bình phương trung bình e_i được xác định là:

$$L_i = c + 10 \log_{10} e_i,$$

trong đó c là hằng số độ lệch. Ví dụ, c có thể bằng -0,691. Từ đây về sau năng lượng của tín hiệu có thể được xác định từ độ to với

$$e_i = 10^{(L_i - c)/10}$$

Năng lượng của tổng của N tín hiệu không tương quan $z_{SUM} = \sum_{i=1}^N z_i$ sau đó là:

$$e_{SUM} = \sum_{i=1}^N e_i = \sum_{i=1}^N 10^{(L_i - c)/10},$$

và độ to của tín hiệu tổng này sau đó là:

$$L_{SUM} = c + 10 \log_{10} e_{SUM} = c + 10 \log_{10} \sum_{i=1}^N 10^{(L_i - c)/10}$$

Nếu các tín hiệu không tương quan, các hệ số tương quan $C_{i,j}$ phải được xem xét khi tính gần đúng năng lượng của tín hiệu tổng là:

$$e_{SUM} = \sum_{i=1}^N \sum_{j=1}^N e_{i,j},$$

trong đó năng lượng chéo $e_{i,j}$ giữa các đối tượng thứ i và thứ j được xác định là:

$$\begin{aligned} e_{i,j} &= C_{i,j} \sqrt{e_i e_j} \\ &= C_{i,j} \sqrt{10^{(L_i - c)/10} 10^{(L_j - c)/10}} \\ &= C_{i,j} \sqrt{10^{(L_i + L_j - 2c)/10}} \end{aligned},$$

trong đó $-1 \leq C_{i,j} \leq 1$ là hệ số tương quan giữa hai đối tượng i và j . Khi hai đối tượng không tương quan, hệ số tương quan bằng 0, và khi hai đối tượng đồng nhất, hệ số tương quan bằng 1.

Hơn nữa việc mở rộng mô hình với việc trộn các trọng số g_i được áp dụng trên các tín hiệu trong quy trình trộn, tức là $z_{SUM} = \sum_{i=1}^N g_i z_i$, năng lượng của tín hiệu tổng sẽ là:

$$e_{SUM} = \sum_{i=1}^N \sum_{j=1}^N g_i g_j e_{i,j},$$

và độ to của tín hiệu hỗn hợp có thể thu được từ đây, như trước đó, với:

$$L_{SUM} = c + 10 \log_{10} e_{SUM}.$$

Sự chênh lệch giữa độ to của hai tín hiệu có thể được ước lượng là:

$$\Delta L(i, j) = L_i - L_j$$

Nếu định nghĩa độ to bây giờ được sử dụng như trước đây, có thể được viết là:

$$\begin{aligned}\Delta L(i, j) &= L_i - L_j \\ &= (c + 10 \log_{10} e_i) - (c + 10 \log_{10} e_j) \\ &= 10 \log_{10} \frac{e_i}{e_j}\end{aligned}$$

mà có thể quan sát được là hàm của các năng lượng tín hiệu. Nếu bây giờ mong muốn ước lượng sự chênh lệch độ to giữa hai hỗn hợp:

$$z_A = \sum_{i=1}^N g_i z_i \quad \text{và} \quad z_B = \sum_{i=1}^N h_i z_i$$

với các trọng số trộn chênh lệch có thể có g_i và h_i , điều đó có thể được ước lượng với:

$$\begin{aligned}\Delta L(A, B) &= 10 \log_{10} \frac{e_A}{e_B} \\ &= 10 \log_{10} \frac{\sum_{i=1}^N \sum_{j=1}^N g_i g_j e_{i,j}}{\sum_{i=1}^N \sum_{j=1}^N h_i h_j e_{i,j}} \\ &= 10 \log_{10} \frac{\sum_{i=1}^N \sum_{j=1}^N g_i g_j C_{i,j} \sqrt{10^{(L_i + L_j - 2c)/10}}}{\sum_{i=1}^N \sum_{j=1}^N h_i h_j C_{i,j} \sqrt{10^{(L_i + L_j - 2c)/10}}} \\ &= 10 \log_{10} \frac{\sum_{i=1}^N \sum_{j=1}^N g_i g_j C_{i,j} \sqrt{10^{(L_i + L_j)/10}}}{\sum_{i=1}^N \sum_{j=1}^N h_i h_j C_{i,j} \sqrt{10^{(L_i + L_j)/10}}}\end{aligned}$$

Trong trường hợp các đối tượng không tương quan ($C_{i,j} = 0, \forall i \neq j$ và $C_{i,j} = 1, \forall i = j$), ước lượng sự chênh lệch trở thành:

$$\begin{aligned}\Delta L(A, B) &= 10 \log_{10} \frac{\sum_{i=1}^N g_i^2 10^{(L_i - c)/10}}{\sum_{i=1}^N h_i^2 10^{(L_i - c)/10}} \\ &= 10 \log_{10} \frac{\sum_{i=1}^N g_i^2 10^{L_i/10}}{\sum_{i=1}^N h_i^2 10^{L_i/10}}\end{aligned}$$

Sau đây, việc mã hóa chênh lệch được xem xét.

Có thể mã hóa các giá trị độ to mỗi đối tượng làm sự chênh lệch với độ to của đối tượng tham chiếu được lựa chọn:

$$K_i = L_i - L_{REF},$$

trong đó L_{REF} là độ to của đối tượng tham chiếu. Việc mã hóa này là có lợi nếu kết quả là không cần giá trị độ to tuyệt đối nào, vì bây giờ cần phải truyền bớt một giá trị, và ước lượng sự chênh lệch độ to có thể được viết là:

$$\Delta L(A, B) = 10 \log_{10} \frac{\sum_{i=1}^N \sum_{j=1}^N g_i g_j C_{i,j} \sqrt{10^{(K_i + K_j)/10}}}{\sum_{i=1}^N \sum_{j=1}^N h_i h_j C_{i,j} \sqrt{10^{(K_i + K_j)/10}}},$$

hoặc trong trường hợp các đối tượng không tương quan:

$$\Delta L(A, B) = 10 \log_{10} \frac{\sum_{i=1}^N g_i^2 10^{K_i/10}}{\sum_{i=1}^N h_i^2 10^{K_i/10}}.$$

Sau đây, kịch bản cải thiện âm thanh hội thoại (DE) sẽ được xem xét.

Xem xét một lần nữa kịch bản ứng dụng cải thiện âm thanh hội thoại. Việc tự do định nghĩa thông tin kết xuất trong bộ giải mã bị giới hạn chỉ thành thay đổi mức của hai siêu đối tượng. Giả sử thêm là hai siêu đối tượng không tương quan, tức là $C_{FGO, BGO} = 0$. Nếu các trọng số trộn giảm của các siêu đối tượng là h_{FGO} và h_{BGO} , và

chúng được kết xuất với độ khuếch đại f_{FGO} và f_{BGO} , độ to của đầu ra liên quan đến trộn giảm mặc định là:

$$\begin{aligned}\Delta L(A, B) &= 10 \log_{10} \frac{f_{FGO}^2 10^{(L_{FGO}-c)/10} + f_{BGO}^2 10^{(L_{BGO}-c)/10}}{h_{FGO}^2 10^{(L_{FGO}-c)/10} + h_{BGO}^2 10^{(L_{BGO}-c)/10}} \\ &= 10 \log_{10} \frac{f_{FGO}^2 10^{L_{FGO}/10} + f_{BGO}^2 10^{L_{BGO}/10}}{h_{FGO}^2 10^{L_{FGO}/10} + h_{BGO}^2 10^{L_{BGO}/10}}\end{aligned}$$

Sau đó, đây cũng là mức bù bắt buộc nếu muốn có cùng độ to tương tự ở đầu ra như trong trộn giảm mặc định.

$\Delta L(A, B)$ có thể được xem xét làm giá trị bù độ to, mà có thể được truyền dẫn bởi bộ xử lý tín hiệu 120 của bộ giải mã. $\Delta L(A, B)$ có thể cũng được đặt tên là giá trị thay đổi độ to và do giá trị bù thực tế có thể là giá trị nghịch đảo. Hoặc cũng có thể sử dụng tên "hệ số bù độ to" cho nó? Do đó, giá trị bù độ to lcv được đề cập đến trước đây trong bản mô tả này sẽ tương ứng với giá trị g_{Delta} dưới đây.

Ví dụ, $g_{\Delta} = 10^{-\Delta L(A, B)/20}$ $1 / \Delta L(A, B)$ có thể được áp dụng làm hệ số nhân trên mỗi kênh của tín hiệu âm thanh được biến đổi mà thu được từ việc áp dụng thông tin kết xuất trên tín hiệu đầu vào âm thanh. Phương trình này để g_{Delta} hoạt động trong miền tuyến tính. Trong miền lôgarit, phương trình sẽ là sự chênh lệch chẳng hạn như $1 / \Delta L(A, B)$ và theo đó được áp dụng.

Nếu quy trình trộn giảm được đơn giản hóa sao cho hai siêu đối tượng có thể được trộn với các trọng số phần tử đơn vị để thu được tín hiệu trộn giảm, tức là $h_{FGO} = h_{BGO} = 1$, và bây giờ độ khuếch đại kết xuất cho hai đối tượng này được kí hiệu là g_{FGO} và g_{BGO} . Điều này đơn giản hóa phương trình cho việc thay đổi độ to thành:

$$\begin{aligned}\Delta L(A, B) &= 10 \log_{10} \frac{g_{FGO}^2 10^{(L_{FGO}-c)/10} + g_{BGO}^2 10^{(L_{BGO}-c)/10}}{10^{(L_{FGO}-c)/10} + 10^{(L_{BGO}-c)/10}} \\ &= 10 \log_{10} \frac{g_{FGO}^2 10^{L_{FGO}/10} + g_{BGO}^2 10^{L_{BGO}/10}}{10^{L_{FGO}/10} + 10^{L_{BGO}/10}}\end{aligned}$$

Một lần nữa, $\Delta L(A, B)$ có thể được xem xét làm giá trị bù độ to mà được xác định bởi bộ xử lý tín hiệu 120.

Nói chung, g_{FGO} có thể được xem xét làm độ khuếch đại kết xuất cho đối tượng nổi FGO (nhóm đối tượng nổi), và g_{BGO} có thể được xem xét làm độ khuếch đại kết xuất cho đối tượng nền BGO (nhóm đối tượng nền).

Như đã được đề cập trước đây, có thể truyền dẫn sự chênh lệch độ to thay vì độ to tuyệt đối. Để định nghĩa độ to tham chiếu như độ to của siêu đối tượng FGO $L_{REF} = L_{FGO}$, tức là $K_{FGO} = L_{FGO} - L_{REF} = 0$ và $K_{BGO} = L_{BGO} - L_{REF} = L_{BGO} - L_{FGO}$. Bây giờ, mức độ thay đổi độ to là:

$$\Delta L(A, B) = 10 \log_{10} \frac{g_{FGO}^2 + g_{BGO}^2 10^{K_{BGO}/10}}{1 + 10^{K_{BGO}/10}}$$

Cũng có thể, như trường hợp trong SAOC-DE mà hai siêu đối tượng không có các hệ số tỷ lệ riêng, nhưng một trong các đối tượng không được biến đổi bên trái, trong khi đối tượng còn lại được làm suy giảm để thu được tỷ lệ trộn chính xác giữa các đối tượng. Trong thiết lập kết xuất này, đầu ra sẽ thấp hơn về độ to so với hỗn hợp mặc định, và sự thay đổi về độ to là:

$$\Delta L(A, B) = 10 \log_{10} \frac{\hat{g}_{FGO}^2 + \hat{g}_{BGO}^2 10^{K_{BGO}/10}}{1 + 10^{K_{BGO}/10}},$$

với

$$\hat{g}_{FGO} = \begin{cases} 1 & , \text{ if } g_{FGO} \geq g_{BGO} \\ \frac{g_{FGO}}{g_{BGO}} & , \text{ if } g_{FGO} < g_{BGO} \end{cases}, \text{ và } \hat{g}_{BGO} = \begin{cases} \frac{g_{BGO}}{g_{FGO}} & , \text{ if } g_{BGO} < g_{FGO} \\ 1 & , \text{ if } g_{BGO} \geq g_{FGO} \end{cases}.$$

Dạng này đã là khá đơn giản, và khá bất khả thi đối với phép đo độ to được sử dụng. Yêu cầu thực tế duy nhất là các giá trị độ to cần tính tổng trong miền số mũ. Có thể truyền dẫn/lưu trữ các giá trị của các năng lượng tín hiệu thay vì các giá trị độ to, vì cả hai có kết nối chặt chẽ.

Trong mỗi công thức trên, $\Delta L(A, B)$ có thể được xem là giá trị bù độ to, mà có thể được truyền dẫn bởi bộ xử lý tín hiệu 120 của bộ giải mã.

Sau đây, các trường hợp ví dụ được xem xét. Độ chính xác của các khái niệm đã đề xuất được minh họa qua hai tín hiệu ví dụ. Cả hai tín hiệu có trộn giảm 5.1 với các kênh xung quanh và kênh LFE được tạo đường vòng từ xử lý SAOC.

Hai phép tính gần đúng chính được sử dụng: một ("3 số hạng") với ba siêu đối tượng: FGO, BGO và các kênh đường vòng, ví dụ:

$$\mathbf{X} = \mathbf{X}_{FGO} + \mathbf{X}_{BGO} + \mathbf{X}_{BYPASS},$$

Và một phép tính gần đúng khác khác ("2 số hạng") với hai siêu đối tượng, ví dụ:

$$\mathbf{X} = \mathbf{X}_{FGO} + \mathbf{X}_{BGO}.$$

Trong phép tính gần đúng 2 số hạng, các kênh đường vòng có thể, ví dụ, được trộn với nhau với BGO cho ước lượng độ to siêu đối tượng. Độ to của cả hai đối tượng (hoặc cả ba) cũng như độ to của tín hiệu trộn giảm được ước lượng, và các giá trị được lưu trữ.

Các lệnh kết xuất là dạng:

$$\mathbf{Y} = \hat{g}_{FGO} \mathbf{X}_{FGO} + \hat{g}_{BGO} \mathbf{X}_{BGO} + \hat{g}_{BGO} \mathbf{X}_{BYPASS}$$

và

$$\mathbf{Y} = \hat{g}_{FGO} \mathbf{X}_{FGO} + \hat{g}_{BGO} \mathbf{X}_{BGO}$$

lần lượt cho hai phép tính gần đúng.

Các giá trị độ khuếch đại ví dụ được xác định theo:

$$\hat{g}_{FGO} = \begin{cases} 1 & , \text{ nếu } g_{FGO} > 1 \\ g_{FGO} & , \text{ ngược lại} \end{cases}, \quad \text{và} \quad \hat{g}_{BGO} = \begin{cases} \frac{1}{g_{FGO}} & , \text{ nếu } g_{FGO} > 1 \\ 1 & , \text{ ngược lại} \end{cases},$$

trong đó độ khuếch đại FGO g_{FGO} được biến thiên từ -24 đến +24dB (dềxiben).

Kịch bản đầu ra được kết xuất, độ to được đo và sự suy giảm từ độ to của tín hiệu trộn giảm được tính toán.

Kết quả này được thể hiện trên Fig.10 và Fig.11 với đường màu xanh da trời với các dấu hiệu tròn. Fig.10 mô tả minh họa thứ nhất và Fig.11 mô tả minh họa thứ hai của thay đổi độ to được đo và kết quả của việc sử dụng các khái niệm được cung cấp để ước lượng sự thay đổi về độ to và theo cách tham số hoàn toàn,

Tiếp theo, sự suy giảm từ trộn giảm được ước lượng theo tham số bằng cách sử dụng các giá trị độ to siêu đối tượng được lưu trữ và thông tin kết xuất và trộn giảm.

Việc ước lượng sử dụng độ to của ba siêu đối tượng được minh họa với đường màu xanh lá cây với các dấu hiệu hình vuông, và việc ước lượng sử dụng độ to của hai siêu đối tượng được minh họa với đường màu đỏ với các dấu hiệu hình ngôi sao.

Có thể thấy được từ các hình, mà phép tính gần đúng 2 số hạng cung cấp kết quả thực tế giống hệt nhau và cả hai đều xấp xỉ giá trị đo được khá tốt.

Các khái niệm được cung cấp thể hiện được nhiều lợi thế. Ví dụ, các khái niệm được cung cấp cho phép ước lượng độ to của tín hiệu hỗn hợp từ độ to của các tín hiệu thành phần tạo nên hỗn hợp. Lợi ích của việc này là độ to tín hiệu thành phần có thể được ước lượng một lần, và ước lượng độ to của tín hiệu hỗn hợp của thể thu được theo tham số cho hỗn hợp bất kỳ không bao gồm ước lượng độ to dựa trên tín hiệu thực. Điều này cung cấp một cải thiện đáng kể trong hiệu quả tính toán của hệ thống tổng thể trong đó việc ước lượng độ to của các hỗn hợp khác nhau là cần thiết. Ví dụ, khi người dùng cuối thay đổi thiết lập kết xuất, việc ước lượng độ to của đầu ra là sẵn có ngay lập tức.

Trong một số ứng dụng, chẳng hạn như khi tuân theo khuyến nghị EBU R128, độ to trung bình trong toàn bộ chương trình là rất quan trọng. Nếu việc ước lượng độ to trong bộ nhận, ví dụ trong kịch bản phát thanh, được thực hiện dựa trên tín hiệu nhận được, thì ước lượng chỉ hội tụ thành độ to trung bình sau khi đã nhận được toàn bộ chương trình. Do đó, bất kỳ sự bù nào của độ to sẽ có sai số hoặc thể hiện các biến thể theo thời gian. Khi ước lượng độ to của các đối tượng thành phần như đã đề xuất và truyền thông tin về độ to, có thể ước lượng độ to hỗn hợp trung bình trong bộ nhận mà không có độ trễ.

Nếu muốn độ to trung bình của tín hiệu đầu ra không đổi (gần đúng) bất kể những thay đổi trong thông tin kết xuất, các khái niệm được cung cấp cho phép xác định hệ số bù vì lý do này. Các tính toán cần thiết cho việc này trong bộ giải mã là do độ phức tạp tính toán của chúng không đáng kể và do đó, chức năng này có thể được thêm vào bất kỳ bộ giải mã nào.

Có những trường hợp mà trong đó mức độ to tuyệt đối của đầu ra là không quan trọng, nhưng điều quan trọng nằm trong việc xác định sự thay đổi về độ to từ cảnh tham chiếu. Trong các trường hợp như vậy mức tuyệt đối của các đối tượng là không

quan trọng, nhưng mức tương đối của chúng thì quan trọng. Điều này cho phép xác định một trong số các đối tượng như đối tượng tham chiếu và biểu diễn độ to của các đối tượng khác trong môi liên hệ với độ to của đối tượng tham chiếu này. Điều này có một số lợi ích trong việc xem xét sự truyền dẫn và/hoặc lưu trữ thông tin độ to.

Trước hết, không cần thiết phải truyền dẫn mức độ to tham chiếu. Trong trường hợp ứng dụng hai siêu đối tượng, điều này làm giảm một nửa lượng dữ liệu được truyền. Lợi ích thứ hai liên quan đến khả năng lượng tử hóa và biểu diễn các giá trị độ to. Vì các mức tuyệt đối của các đối tượng có thể là hầu hết mọi thứ, nên giá trị độ to tuyệt đối cũng có thể là hầu hết mọi thứ. Mặt khác, các giá trị độ to tương đối được giả định có giá trị trung bình 0 và một phân bố được hình thành khá độc đáo xung quanh giá trị trung bình. Sự khác biệt giữa các phép biểu diễn cho phép xác định lưới lượng tử hóa của phép biểu diễn tương đối theo cách có khả năng chính xác cao hơn với cùng số bit được sử dụng cho phép biểu diễn lượng tử hóa.

Fig.12 minh họa phương án khác để thực hiện bù độ to. Trên Fig.12, sự bù độ to có thể được thực hiện, ví dụ để bù những mất mát độ to, với mục đích này, ví dụ các giá trị **DE_loudness_diff_dialogue** (= K_{FGO}) và **DE_loudness_diff_background** (= K_{BGO}) từ **DE_control_info** có thể được sử dụng. Ở đây **DE_control_info** có thể xác định thông tin kiểm soát "cải thiện âm thanh hội thoại" (DE) âm thanh sạch được cải tiến.

Sự bù độ to đạt được bằng cách áp dụng giá trị khuếch đại "g" trên tín hiệu đầu ra SAOC-DE và các kênh đường vòng (trong trường hợp tín hiệu đa kênh).

Trong phương án của Fig.12, điều đó được thực hiện như sau:

Giá trị khuếch đại biến đổi hội thoại bị giới hạn m_G được sử dụng để xác định độ khuếch đại hiệu quả cho đối tượng nổi (FGO, ví dụ, hội thoại) và cho đối tượng nền (BGO, ví dụ môi trường xung quanh). Điều này được thực hiện bởi khối "ánh xạ độ khuếch đại" 1220 mà tạo ra các giá trị độ khuếch đại m_{FGO} và m_{BGO} .

Khối "bộ ước lượng độ to đầu ra" 1230 sử dụng thông tin độ to K_{FGO} và K_{BGO} , và các giá trị độ khuếch đại hiệu quả m_{FGO} và m_{BGO} để ước lượng sự thay đổi có thể có này trong độ to so với trường hợp trộn giảm mặc định. Sự thay đổi sau đó được ánh xạ

thành "hệ số bù độ to" mà được áp dụng trên các kênh đầu ra để tạo ra "các tín hiệu đầu ra" cuối cùng.

Các bước sau được áp dụng cho việc bù độ to:

- Nhận giá trị độ khuếch đại bị giới hạn m_G từ bộ giải mã SAOC-DE (như được định nghĩa trong cụm từ 12.8 "kiểm soát khoảng biến đổi cho SAOC-DE" [DE]), và xác định các độ khuếch đại FGO/BGO được áp dụng:

$$\begin{cases} m_{FGO} = m_G, \text{ and } m_{BGO} = 1 & \text{if } m_G \leq 1 \\ m_{FGO} = 1, \text{ and } m_{BGO} = m_G^{-1} & \text{if } m_G > 1 \end{cases}$$

- Thu được thông tin độ to siêu đối tượng K_{FGO} và K_{BGO} .

- Tính toán sự thay đổi độ to đầu ra so với trộn giảm mặc định với:

$$\Delta L = 10 \log_{10} \frac{m_{FGO}^2 10^{K_{FGO}/10} + m_{BGO}^2 10^{K_{BGO}/10}}{10^{K_{FGO}/10} + 10^{K_{BGO}/10}}$$

- Tính toán độ khuếch đại sự bù độ to $g_{\Delta} = 10^{-0.05\Delta L}$.

- Tính toán các hệ số định tỷ lệ $\mathbf{g} = \begin{bmatrix} g_1 \\ \vdots \\ g_N \end{bmatrix}$, trong đó

$$g_i = \begin{cases} g_{\Delta} & \text{nếu kênh } i \text{ thuộc về đầu ra SAOC-DE} \\ m_{BGO} g_{\Delta} & \text{nếu kênh } i \text{ là kênh đường vòng} \end{cases}, \text{ và } N \text{ là tổng số các kênh đầu}$$

ra. Trên Fig.12, việc điều chỉnh độ khuếch đại được chia thành hai bước: độ khuếch đại của "các kênh đường vòng" có thể có được điều chỉnh với m_{BGO} trước khi kết hợp chúng với "các kênh đầu ra SAOC-DE", và sau đó độ khuếch đại chung g_{Δ} được áp dụng trên tất cả các kênh được kết hợp. Điều này chỉ là sắp xếp lại thứ tự có thể có của các vận hành điều chỉnh độ khuếch đại, trong khi $\mathbf{g}$ ở đây được kết hợp cả hai bước điều chỉnh độ khuếch đại thành một điều chỉnh độ khuếch đại.

- Áp dụng các giá trị định tỷ lệ g trên các kênh âm thanh Y_{FULL} bao gồm "các kênh đầu ra SAOC-DE" Y_{SAOC} và "các kênh đường vòng" được sắp hàng theo thời gian có thể có Y_{BYPASS} : $Y_{FULL} = Y_{SAOC} \cup Y_{BYPASS}$

Việc áp dụng các giá trị định tỷ lệ g trên các kênh âm thanh Y_{FULL} được thực hiện bởi bộ phận điều chỉnh độ khuếch đại 1240.

ΔL như được tính toán ở trên có thể được xem là giá trị bù độ to. Nói chung, $mFGO$ có thể cho biết độ khuếch đại kết xuất cho đối tượng nổi FGO (nhóm đối tượng nổi), và $mBGO$ cho biết độ khuếch đại kết xuất cho đối tượng nền BGO (nhóm đối tượng nền).

Mặc dù một vài khía cạnh đã được mô tả trong phạm vi của thiết bị, rõ ràng rằng các khía cạnh này cũng biểu diễn sự mô tả của phương pháp tương ứng, mà khối hoặc thiết bị tương ứng với bước phương pháp hoặc đặc điểm của bước 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 đặc điểm tương ứng của thiết bị tương ứng.

Tín hiệu đã được tách ra theo sáng chế có thể được lưu trữ trên vật ghi kỹ thuật số hoặc có thể được truyền dẫn trên phương tiện truyền dẫn chẳng hạn như phương tiện truyền dẫn không dây hoặc phương tiện truyền dẫn có dây như Internet.

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

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

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

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

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

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 các phương pháp đã được mô tả ở đây. Ví dụ, có thể tạo cấu hình dòng dữ liệu hoặc chuỗi các tín hiệu để được truyền thông qua sự kết nối truyền thông dữ liệu, ví dụ thông qua Internet.

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

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

Theo một số phương án, thiết bị logic có thể 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. Theo 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 ưu tiên được thực hiện bởi thiết bị phần cứng bất kỳ.

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

YÊU CẦU BẢO HỘ

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

giao diện nhận (110) để nhận tín hiệu đầu vào âm thanh bao gồm nhiều tín hiệu đối tượng âm thanh, để nhận thông tin độ to trên các tín hiệu đối tượng âm thanh, và để nhận thông tin kết xuất cho biết một hoặc nhiều tín hiệu đối tượng âm thanh sẽ được khuếch đại hoặc được làm suy giảm như thế nào, và

bộ xử lý tín hiệu (120) để tạo ra một hoặc nhiều kênh đầu ra âm thanh của tín hiệu đầu ra âm thanh,

trong đó bộ xử lý tín hiệu (120) được tạo cấu hình để xác định giá trị bù độ to phụ thuộc vào thông tin độ to và phụ thuộc vào thông tin kết xuất, và

trong đó bộ xử lý tín hiệu (120) được tạo cấu hình để tạo ra một hoặc nhiều kênh đầu ra âm thanh của tín hiệu đầu ra âm thanh từ tín hiệu đầu vào âm thanh phụ thuộc vào thông tin kết xuất và phụ thuộc vào giá trị bù độ to,

trong đó bộ xử lý tín hiệu (120) được tạo cấu hình để tạo ra một hoặc nhiều kênh đầu ra âm thanh của tín hiệu đầu ra âm thanh từ tín hiệu đầu vào âm thanh phụ thuộc vào thông tin kết xuất và phụ thuộc vào giá trị bù độ to, sao cho độ to của tín hiệu đầu ra âm thanh bằng với độ to của tín hiệu đầu vào âm thanh, hoặc sao cho độ to của tín hiệu đầu ra âm thanh gần hơn với độ to của tín hiệu đầu vào âm thanh so với độ to của tín hiệu âm thanh được biến đổi mà sẽ có được từ việc biến đổi tín hiệu đầu vào âm thanh bằng cách khuếch đại hoặc làm suy giảm các tín hiệu đối tượng âm thanh của tín hiệu đầu vào âm thanh theo thông tin kết xuất.

2. Bộ giải mã theo điểm 1,

trong đó bộ xử lý tín hiệu (120) được tạo cấu hình để tạo ra tín hiệu âm thanh được biến đổi bằng việc biến đổi tín hiệu đầu vào âm thanh bằng cách khuếch đại hoặc làm suy giảm các tín hiệu đối tượng âm thanh của tín hiệu đầu vào âm thanh theo thông tin kết xuất, và

trong đó bộ xử lý tín hiệu (120) được tạo cấu hình để tạo ra tín hiệu đầu ra âm thanh bằng cách áp dụng giá trị bù độ to trên tín hiệu âm thanh được biến đổi, sao cho độ to của tín hiệu đầu ra âm thanh bằng với độ to của tín hiệu đầu vào âm thanh, hoặc sao cho độ to của tín hiệu đầu ra âm thanh gần hơn với độ to của tín hiệu đầu vào âm thanh so với độ to của tín hiệu âm thanh được biến đổi.

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

trong đó mỗi tín hiệu trong số các tín hiệu đối tượng âm thanh của tín hiệu đầu vào âm thanh được ấn định cho chính xác một nhóm trong số hai hoặc nhiều hơn hai nhóm, trong đó mỗi nhóm trong số hai hoặc nhiều hơn hai nhóm bao gồm một hoặc nhiều tín hiệu đối tượng âm thanh của tín hiệu đầu vào âm thanh,

trong đó giao diện nhận (110) được tạo cấu hình để nhận giá trị độ to cho mỗi nhóm trong số hai hoặc nhiều hơn hai nhóm làm thông tin độ to,

trong đó bộ xử lý tín hiệu (120) được tạo cấu hình để xác định giá trị bù độ to phụ thuộc vào giá trị độ to của mỗi nhóm trong số hai hoặc nhiều hơn hai nhóm, và

trong đó bộ xử lý tín hiệu (120) được tạo cấu hình để tạo ra một hoặc nhiều kênh đầu ra âm thanh của tín hiệu đầu ra âm thanh từ tín hiệu đầu vào âm thanh phụ thuộc vào giá trị bù độ to.

4. Bộ giải mã theo điểm 3, trong đó ít nhất một nhóm trong số hai hoặc nhiều hơn hai nhóm bao gồm hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh.

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

trong đó mỗi tín hiệu trong số các tín hiệu đối tượng âm thanh của tín hiệu đầu vào âm thanh được ấn định cho chính xác một nhóm trong số nhiều hơn hai nhóm, trong đó mỗi nhóm trong số nhiều hơn hai nhóm bao gồm một hoặc nhiều tín hiệu đối tượng âm thanh của tín hiệu đầu vào âm thanh,

trong đó giao diện nhận (110) được tạo cấu hình để nhận giá trị độ to từ mỗi nhóm trong số nhiều hơn hai nhóm như thông tin độ to,

trong đó bộ xử lý tín hiệu (120) được tạo cấu hình để xác định giá trị bù độ to phụ thuộc vào giá trị độ to của mỗi nhóm trong số nhiều hơn hai nhóm, và

trong đó bộ xử lý tín hiệu (120) được tạo cấu hình để tạo ra một hoặc nhiều kênh đầu ra âm thanh của tín hiệu đầu ra âm thanh từ tín hiệu đầu vào âm thanh phụ thuộc vào giá trị bù độ to.

6. Bộ giải mã theo điểm 5, trong đó ít nhất một nhóm trong số nhiều hơn hai nhóm bao gồm hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh.

7. Bộ giải mã theo điểm bất kỳ trong số các điểm từ 3 đến 6,

trong đó bộ xử lý tín hiệu (120) được tạo cấu hình để xác định giá trị bù độ to theo công thức:

$$\Delta L = 10 \log_{10} \frac{\sum_{i=1}^N g_i^2 10^{(L_i-c)/10}}{\sum_{i=1}^N h_i^2 10^{(L_i-c)/10}}$$

hoặc theo công thức:

$$\Delta L = 10 \log_{10} \frac{\sum_{i=1}^N g_i^2 10^{L_i/10}}{\sum_{i=1}^N h_i^2 10^{L_i/10}}$$

trong đó ΔL là giá trị bù độ to,

trong đó i biểu thị tín hiệu đối tượng âm thanh thứ i của các tín hiệu đối tượng âm thanh,

trong đó L_i là độ to của tín hiệu đối tượng âm thanh thứ i ,

trong đó g_i là trọng số trộn thứ nhất cho tín hiệu đối tượng âm thanh thứ i ,

trong đó h_i là trọng số trộn thứ hai cho tín hiệu đối tượng âm thanh thứ i ,

trong đó c là hằng số, và

trong đó N là số lượng.

8. Bộ giải mã theo điểm bất kỳ trong số các điểm từ 3 đến 6,

trong đó bộ xử lý tín hiệu (120) được tạo cấu hình để xác định giá trị bù độ to theo công thức

$$\Delta L = 10 \log_{10} \frac{\sum_{i=1}^N g_i^2 10^{K_i/10}}{\sum_{i=1}^N h_i^2 10^{K_i/10}}$$

trong đó ΔL là giá trị bù độ to,

trong đó i biểu thị tín hiệu đối tượng âm thanh thứ i của các tín hiệu đối tượng âm thanh,

trong đó g_i là trọng số trộn thứ nhất cho tín hiệu đối tượng âm thanh thứ i ,

trong đó h_i là trọng số trộn thứ hai cho tín hiệu đối tượng âm thanh thứ i ,

trong đó N là số lượng, và

trong đó K_i được xác định theo:

$$K_i = L_i - L_{REF},$$

trong đó L_i là độ to của tín hiệu đối tượng âm thanh thứ i , và

trong đó L_{REF} là độ to của đối tượng tham chiếu.

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

trong đó mỗi tín hiệu trong số các tín hiệu đối tượng âm thanh của tín hiệu đầu vào âm thanh được ấn định cho chính xác một nhóm trong số chính xác hai nhóm như hai hoặc nhiều hơn hai nhóm,

trong đó mỗi tín hiệu trong số các tín hiệu đối tượng âm thanh của tín hiệu đầu vào âm thanh hoặc được ấn định cho nhóm đối tượng nổi của chính xác hai nhóm hoặc cho nhóm đối tượng nền của chính xác hai nhóm,

trong đó giao diện nhận (110) được tạo cấu hình để nhận giá trị độ to của nhóm đối tượng nổi,

trong đó giao diện nhận (110) được tạo cấu hình để nhận giá trị độ to của nhóm đối tượng nền,

trong đó bộ xử lý tín hiệu (120) được tạo cấu hình để xác định giá trị bù độ to phụ thuộc vào giá trị độ to của nhóm đối tượng nổi, và phụ thuộc vào giá trị độ to của nhóm đối tượng nền, và

trong đó bộ xử lý tín hiệu (120) được tạo cấu hình để tạo ra một hoặc nhiều kênh đầu ra âm thanh của tín hiệu đầu ra âm thanh từ tín hiệu đầu vào âm thanh phụ thuộc vào giá trị bù độ to.

10. Bộ giải mã theo điểm 9,

trong đó bộ xử lý tín hiệu (120) được tạo cấu hình để xác định giá trị bù độ to theo công thức:

$$\Delta L = 10 \log_{10} \frac{m_{FGO}^2 10^{K_{FGO}/10} + m_{BGO}^2 10^{K_{BGO}/10}}{10^{K_{FGO}/10} + 10^{K_{BGO}/10}}$$

trong đó ΔL là giá trị bù độ to,

trong đó K_{FGO} biểu thị giá trị độ to của nhóm đối tượng nổi,

trong đó K_{BGO} biểu thị giá trị độ to của nhóm đối tượng nền,

trong đó m_{FGO} biểu thị độ khuếch đại kết xuất của nhóm đối tượng nổi, và

trong đó m_{BGO} biểu thị độ khuếch đại kết xuất của nhóm đối tượng nền.

11. Bộ giải mã theo điểm 9,

trong đó bộ xử lý tín hiệu (120) được tạo cấu hình để xác định giá trị bù độ to theo công thức:

$$\Delta L(A, B) = 10 \log_{10} \frac{g_{FGO}^2 10^{L_{FGO}/10} + g_{BGO}^2 10^{L_{BGO}/10}}{10^{L_{FGO}/10} + 10^{L_{BGO}/10}}$$

trong đó ΔL là giá trị bù độ to,

trong đó L_{FGO} biểu thị giá trị độ to của nhóm đối tượng nổi,

trong đó L_{BGO} biểu thị giá trị độ to của nhóm đối tượng nền,

trong đó g_{FGO} biểu thị độ khuếch đại kết xuất của nhóm đối tượng nổi, và

trong đó g_{BGO} biểu thị độ khuếch đại kết xuất của nhóm đối tượng nền.

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

trong đó giao diện nhận (110) được tạo cấu hình để nhận tín hiệu trộn giảm bao gồm một hoặc nhiều kênh trộn giảm làm tín hiệu đầu vào âm thanh, trong đó một hoặc nhiều kênh trộn giảm bao gồm các tín hiệu đối tượng âm thanh, và trong đó số lượng của một hoặc nhiều kênh trộn giảm nhỏ hơn số lượng của các tín hiệu đối tượng âm thanh,

trong đó giao diện nhận (110) được tạo cấu hình để nhận thông tin trộn giảm cho biết các tín hiệu đối tượng âm thanh được trộn trong một hoặc nhiều kênh trộn giảm như thế nào, và

trong đó bộ xử lý tín hiệu (120) được tạo cấu hình để tạo ra một hoặc nhiều kênh đầu ra âm thanh của tín hiệu đầu ra âm thanh từ tín hiệu đầu vào âm thanh phụ thuộc vào thông tin trộn giảm, phụ thuộc vào thông tin kết xuất và phụ thuộc vào giá trị bù độ to.

13. Bộ giải mã theo điểm 12,

trong đó giao diện nhận (110) được tạo cấu hình để nhận một hoặc nhiều tín hiệu đối tượng âm thanh đường vòng thêm, trong đó một hoặc nhiều tín hiệu đối tượng âm thanh đường vòng thêm không được trộn trong tín hiệu trộn giảm,

trong đó giao diện nhận (110) được tạo cấu hình để nhận thông tin độ to cho biết thông tin về độ to của các tín hiệu đối tượng âm thanh mà được trộn trong tín hiệu trộn giảm và cho biết thông tin về độ to của một hoặc nhiều tín hiệu đối tượng âm thanh đường vòng thêm mà không được trộn trong tín hiệu trộn giảm, và

trong đó bộ xử lý tín hiệu (120) được tạo cấu hình để xác định giá trị bù độ to phụ thuộc vào thông tin về độ to của các tín hiệu đối tượng âm thanh mà được trộn trong tín hiệu trộn giảm, và phụ thuộc vào thông tin về độ to của một hoặc nhiều tín hiệu đối tượng âm thanh đường vòng thêm mà không được trộn trong tín hiệu trộn giảm.

14. Bộ giải mã tạo ra tín hiệu đầu ra âm thanh bao gồm một hoặc nhiều kênh đầu ra âm thanh, trong đó bộ giải mã bao gồm:

giao diện nhận (110) để nhận tín hiệu đầu vào âm thanh bao gồm nhiều tín hiệu đối tượng âm thanh, để nhận thông tin độ to trên các tín hiệu đối tượng âm thanh, và để nhận thông tin kết xuất cho biết một hoặc nhiều tín hiệu đối tượng âm thanh sẽ được khuếch đại hoặc được làm suy giảm như thế nào, và

bộ xử lý tín hiệu (120) để tạo ra một hoặc nhiều kênh đầu ra âm thanh của tín hiệu đầu ra âm thanh,

trong đó bộ xử lý tín hiệu (120) được tạo cấu hình để xác định giá trị bù độ to phụ thuộc vào thông tin độ to và phụ thuộc vào thông tin kết xuất, và

trong đó bộ xử lý tín hiệu (120) được tạo cấu hình để tạo ra một hoặc nhiều kênh đầu ra âm thanh của tín hiệu đầu ra âm thanh từ tín hiệu đầu vào âm thanh phụ thuộc vào thông tin kết xuất và phụ thuộc vào giá trị bù độ to,

trong đó giao diện nhận (110) được tạo cấu hình để nhận tín hiệu trộn giảm bao gồm một hoặc nhiều kênh trộn giảm làm tín hiệu đầu vào âm thanh, trong đó một hoặc nhiều kênh trộn giảm bao gồm các tín hiệu đối tượng âm thanh, và trong đó số lượng của một hoặc nhiều kênh trộn giảm nhỏ hơn số lượng của các tín hiệu đối tượng âm thanh,

trong đó giao diện nhận (110) được tạo cấu hình để nhận thông tin trộn giảm cho biết các tín hiệu đối tượng âm thanh được trộn như thế nào trong một hoặc nhiều kênh trộn giảm, và

trong đó bộ xử lý tín hiệu (120) được tạo cấu hình để tạo ra một hoặc nhiều kênh đầu ra âm thanh của tín hiệu đầu ra âm thanh từ tín hiệu đầu vào âm thanh phụ thuộc vào thông tin trộn giảm, phụ thuộc vào thông tin kết xuất và phụ thuộc vào giá trị bù độ to,

trong đó giao diện nhận (110) được tạo cấu hình để nhận một hoặc nhiều tín hiệu đối tượng âm thanh đường vòng thêm, trong đó một hoặc nhiều tín hiệu đối tượng âm thanh đường vòng thêm không được trộn trong tín hiệu trộn giảm,

trong đó giao diện nhận (110) được tạo cấu hình để nhận thông tin độ to cho biết thông tin về độ to của các tín hiệu đối tượng âm thanh mà được trộn trong tín hiệu

trộn giảm và cho biết thông tin về độ to của một hoặc nhiều tín hiệu đối tượng âm thanh đường vòng thêm mà không được trộn trong tín hiệu trộn giảm, và

trong đó bộ xử lý tín hiệu (120) được tạo cấu hình để xác định giá trị bù độ to phụ thuộc vào thông tin về độ to của các tín hiệu đối tượng âm thanh mà được trộn trong tín hiệu trộn giảm, và phụ thuộc vào thông tin về độ to của một hoặc nhiều tín hiệu đối tượng âm thanh đường vòng thêm mà không được trộn trong tín hiệu trộn giảm.

15. Bộ mã hóa bao gồm:

bộ phận mã hóa dựa trên đối tượng (210; 710) để mã hóa nhiều tín hiệu đối tượng âm thanh để thu được tín hiệu âm thanh được mã hóa bao gồm nhiều tín hiệu đối tượng âm thanh, và

bộ phận mã hóa độ to đối tượng (220; 720; 820) để mã hóa thông tin độ to trên các tín hiệu đối tượng âm thanh,

trong đó thông tin độ to bao gồm một hoặc nhiều giá trị độ to, trong đó mỗi giá trị trong số một hoặc nhiều giá trị độ to phụ thuộc vào một hoặc nhiều tín hiệu đối tượng âm thanh,

trong đó mỗi tín hiệu trong số các tín hiệu đối tượng âm thanh của tín hiệu âm thanh được mã hóa được ấn định cho chính xác một nhóm trong số hai hoặc nhiều hơn hai nhóm, trong đó mỗi nhóm trong số hai hoặc nhiều hơn hai nhóm bao gồm một hoặc nhiều tín hiệu đối tượng âm thanh của tín hiệu âm thanh được mã hóa, trong đó ít nhất một nhóm trong số hai hoặc nhiều hơn hai nhóm bao gồm hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh,

trong đó bộ phận mã hóa độ to đối tượng (220; 720; 820) được tạo cấu hình để xác định một hoặc nhiều giá trị độ to của thông tin độ to bằng cách xác định giá trị độ to cho mỗi nhóm trong số hai hoặc nhiều hơn hai nhóm, trong đó giá trị độ to đã nêu thuộc nhóm đã nêu cho biết tổng độ to của một hoặc nhiều tín hiệu đối tượng âm thanh của nhóm đã nêu.

16. Bộ mã hóa bao gồm:

bộ phận mã hóa dựa trên đối tượng (210; 710) để mã hóa nhiều tín hiệu đối tượng âm thanh để thu được tín hiệu âm thanh được mã hóa bao gồm nhiều tín hiệu đối tượng âm thanh, và

bộ phận mã hóa độ to đối tượng (220; 720; 820) để mã hóa thông tin độ to trên các tín hiệu đối tượng âm thanh,

trong đó thông tin độ to bao gồm một hoặc nhiều giá trị độ to, trong đó mỗi giá trị trong số một hoặc nhiều giá trị độ to phụ thuộc vào một hoặc nhiều tín hiệu đối tượng âm thanh,

trong đó bộ phận mã hóa dựa trên đối tượng (210; 710) được tạo cấu hình để nhận các tín hiệu đối tượng âm thanh, trong đó mỗi tín hiệu trong số các tín hiệu đối tượng âm thanh được ấn định cho chính xác một trong số chính xác hai nhóm, trong đó mỗi nhóm trong số chính xác hai nhóm bao gồm một hoặc nhiều tín hiệu đối tượng âm thanh, trong đó ít nhất một nhóm trong số chính xác hai nhóm bao gồm hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh,

trong đó bộ phận mã hóa dựa trên đối tượng (210; 710) được tạo cấu hình để trộn giảm các tín hiệu đối tượng âm thanh, có trong chính xác hai nhóm, để thu được tín hiệu trộn giảm bao gồm một hoặc nhiều kênh âm thanh trộn giảm làm tín hiệu âm thanh được mã hóa, trong đó số lượng của một hoặc nhiều kênh trộn giảm nhỏ hơn số lượng của các tín hiệu đối tượng âm thanh có trong chính xác hai nhóm,

trong đó bộ phận mã hóa độ to đối tượng (220; 720; 820) được tạo cấu hình để nhận một hoặc nhiều tín hiệu đối tượng âm thanh đường vòng thêm, trong đó mỗi tín hiệu trong số một hoặc nhiều tín hiệu đối tượng âm thanh đường vòng thêm được ấn định cho nhóm thứ ba, trong đó mỗi tín hiệu trong số một hoặc nhiều tín hiệu đối tượng âm thanh đường vòng thêm không có trong nhóm thứ nhất và không có trong nhóm thứ hai, trong đó bộ phận mã hóa dựa trên đối tượng (210; 710) được tạo cấu hình để không trộn giảm một hoặc nhiều tín hiệu đối tượng âm thanh đường vòng thêm trong tín hiệu trộn giảm, và

trong đó bộ phận mã hóa độ to đối tượng (220; 720; 820) được tạo cấu hình để xác định giá trị độ to thứ nhất, giá trị độ to thứ hai và giá trị độ to thứ ba của thông tin

độ to, giá trị độ to thứ nhất cho biết tổng độ to của một hoặc nhiều tín hiệu đối tượng âm thanh của nhóm thứ nhất, giá trị độ to thứ hai cho biết tổng độ to của một hoặc nhiều tín hiệu đối tượng âm thanh của nhóm thứ hai, và giá trị độ to thứ ba cho biết tổng độ to của một hoặc nhiều tín hiệu đối tượng âm thanh đường vòng thêm của nhóm thứ ba, hoặc được tạo cấu hình để xác định giá trị độ to thứ nhất và giá trị độ to thứ hai của thông tin độ to, giá trị độ to thứ nhất cho biết tổng độ to của một hoặc nhiều tín hiệu đối tượng âm thanh của nhóm thứ nhất, và giá trị độ to thứ hai cho biết tổng độ to của một hoặc nhiều tín hiệu đối tượng âm thanh của nhóm thứ hai và của một hoặc nhiều tín hiệu đối tượng âm thanh đường vòng thêm của nhóm thứ ba.

17. Hệ thống mã hóa tín hiệu đối tượng âm thanh, mã hóa thông tin độ to và tạo ra tín hiệu đầu ra âm thanh, hệ thống bao gồm:

bộ mã hóa (310) bao gồm:

bộ phận mã hóa dựa trên đối tượng (210; 710) để mã hóa nhiều tín hiệu đối tượng âm thanh để thu được tín hiệu âm thanh được mã hóa bao gồm nhiều tín hiệu đối tượng âm thanh, và

bộ phận mã hóa độ to đối tượng (220; 720; 820) để mã hóa thông tin độ to trên các tín hiệu đối tượng âm thanh,

trong đó thông tin độ to bao gồm một hoặc nhiều giá trị độ to, trong đó mỗi giá trị trong số một hoặc nhiều giá trị độ to phụ thuộc vào một hoặc nhiều tín hiệu đối tượng âm thanh,

bộ giải mã (320) theo một trong số các điểm từ 1 đến 14 để tạo ra tín hiệu đầu ra âm thanh bao gồm một hoặc nhiều kênh đầu ra âm thanh,

trong đó bộ giải mã (320) được tạo cấu hình để nhận tín hiệu âm thanh được mã hóa làm tín hiệu đầu vào âm thanh và để nhận thông tin độ to

trong đó bộ giải mã (320) được tạo cấu hình để nhận thêm thông tin kết xuất,

trong đó bộ giải mã (320) được tạo cấu hình để xác định giá trị bù độ to phụ thuộc vào thông tin độ to và phụ thuộc vào thông tin kết xuất, và

trong đó bộ giải mã (320) được tạo cấu hình để tạo ra một hoặc nhiều kênh đầu ra âm thanh của tín hiệu đầu ra âm thanh từ tín hiệu đầu vào âm thanh phụ thuộc vào thông tin kết xuất và phụ thuộc vào giá trị bù độ to.

18. Phương pháp tạo ra tín hiệu đầu ra âm thanh bao gồm một hoặc nhiều kênh đầu ra âm thanh, trong đó phương pháp bao gồm các bước:

nhận tín hiệu đầu vào âm thanh bao gồm nhiều tín hiệu đối tượng âm thanh,

nhận thông tin độ to trên các tín hiệu đối tượng âm thanh,

nhận thông tin kết xuất cho biết một hoặc nhiều tín hiệu đối tượng âm thanh sẽ được khuếch đại hoặc được làm suy giảm như thế nào,

xác định giá trị bù độ to phụ thuộc vào thông tin độ to và phụ thuộc vào thông tin kết xuất, và

tạo ra một hoặc nhiều kênh đầu ra âm thanh của tín hiệu đầu ra âm thanh từ tín hiệu đầu vào âm thanh phụ thuộc vào thông tin kết xuất và phụ thuộc vào giá trị bù độ to,

trong đó việc tạo ra một hoặc nhiều kênh đầu ra âm thanh của tín hiệu đầu ra âm thanh từ tín hiệu đầu vào âm thanh được thực hiện phụ thuộc vào thông tin kết xuất và phụ thuộc vào giá trị bù độ to, sao cho độ to của tín hiệu đầu ra âm thanh bằng với độ to của tín hiệu đầu vào âm thanh, hoặc sao cho độ to của tín hiệu đầu ra âm thanh gần hơn độ to của tín hiệu đầu vào âm thanh so với độ to của tín hiệu âm thanh được biến đổi mà sẽ có được từ việc biến đổi tín hiệu đầu vào âm thanh bằng cách khuếch đại hoặc làm suy giảm các tín hiệu đối tượng âm thanh của tín hiệu đầu vào âm thanh theo thông tin kết xuất.

19. Phương pháp tạo ra tín hiệu đầu ra âm thanh bao gồm một hoặc nhiều kênh đầu ra âm thanh, trong đó phương pháp bao gồm:

nhận tín hiệu đầu vào âm thanh bao gồm nhiều tín hiệu đối tượng âm thanh, trong đó việc nhận tín hiệu đầu vào âm thanh được thực hiện bằng cách nhận tín hiệu trộn giảm bao gồm một hoặc nhiều kênh trộn giảm làm tín hiệu đầu vào âm thanh, trong đó một hoặc nhiều kênh trộn giảm bao gồm các tín hiệu đối tượng âm thanh, và

trong đó số lượng của một hoặc nhiều kênh trộn giảm nhỏ hơn số lượng của các tín hiệu đối tượng âm thanh,

nhận thông tin kết xuất cho biết hoặc một hoặc nhiều tín hiệu đối tượng âm thanh sẽ được khuếch đại hoặc được làm suy giảm,

nhận thông tin trộn giảm cho biết các tín hiệu đối tượng âm thanh được trộn như thế nào trong một hoặc nhiều kênh trộn giảm,

nhận một hoặc nhiều tín hiệu đối tượng âm thanh đường vòng thêm, trong đó một hoặc nhiều tín hiệu đối tượng âm thanh đường vòng thêm không được trộn trong tín hiệu trộn giảm,

nhận thông tin độ to trên các tín hiệu đối tượng âm thanh, trong đó thông tin độ to cho biết thông tin về độ to của các tín hiệu đối tượng âm thanh mà được trộn trong tín hiệu trộn giảm và cho biết thông tin về độ to của một hoặc nhiều tín hiệu đối tượng âm thanh đường vòng thêm mà không được trộn trong tín hiệu trộn giảm, và

xác định giá trị bù độ to phụ thuộc vào thông tin độ to và phụ thuộc vào thông tin kết xuất, trong đó việc xác định giá trị bù độ to được thực hiện phụ thuộc vào thông tin về độ to của các tín hiệu đối tượng âm thanh mà được trộn trong tín hiệu trộn giảm, và phụ thuộc vào thông tin về độ to của một hoặc nhiều tín hiệu đối tượng âm thanh đường vòng thêm mà không được trộn trong tín hiệu trộn giảm, và

tạo ra một hoặc nhiều kênh đầu ra âm thanh của tín hiệu đầu ra âm thanh từ tín hiệu đầu vào âm thanh phụ thuộc vào thông tin trộn giảm, phụ thuộc vào thông tin kết xuất và phụ thuộc vào giá trị bù độ to.

20. Phương pháp mã hóa bao gồm các bước:

mã hóa nhiều tín hiệu đối tượng âm thanh để thu được tín hiệu âm thanh được mã hóa bao gồm nhiều tín hiệu đối tượng âm thanh, và

xác định thông tin độ to trên các tín hiệu đối tượng âm thanh, trong đó thông tin độ to bao gồm một hoặc nhiều giá trị độ to, trong đó mỗi giá trị trong số một hoặc nhiều giá trị độ to phụ thuộc vào một hoặc nhiều tín hiệu đối tượng âm thanh

trong đó việc xác định một hoặc nhiều giá trị độ to của thông tin độ to được thực hiện bằng cách xác định giá trị độ to cho mỗi nhóm trong số hai hoặc nhiều hơn hai nhóm, trong đó giá trị độ to đã nêu của nhóm đã nêu cho biết tổng độ to của một hoặc nhiều tín hiệu đối tượng âm thanh của nhóm đã nêu

mã hóa thông tin độ to trên các tín hiệu đối tượng âm thanh,

trong đó mỗi tín hiệu trong số các tín hiệu đối tượng âm thanh của tín hiệu âm thanh được mã hóa được ấn định cho chính xác một nhóm trong số hai hoặc nhiều hơn hai nhóm, trong đó mỗi nhóm trong số hai hoặc nhiều hơn hai nhóm bao gồm một hoặc nhiều tín hiệu đối tượng âm thanh của tín hiệu âm thanh được mã hóa, trong đó ít nhất một nhóm trong số hai hoặc nhiều hơn hai nhóm bao gồm hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh.

21. Phương pháp mã hóa bao gồm các bước:

nhận các tín hiệu đối tượng âm thanh, trong đó mỗi tín hiệu trong số các tín hiệu đối tượng âm thanh được ấn định cho chính xác một nhóm trong số chính xác hai nhóm, trong đó mỗi nhóm trong số chính xác hai nhóm bao gồm một hoặc nhiều tín hiệu đối tượng âm thanh, trong đó ít nhất một nhóm trong số chính xác hai nhóm bao gồm hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh,

mã hóa nhiều tín hiệu đối tượng âm thanh để thu được tín hiệu âm thanh được mã hóa bao gồm nhiều tín hiệu đối tượng âm thanh bằng cách trộn giảm các tín hiệu đối tượng âm thanh, có trong chính xác hai nhóm, để thu được tín hiệu trộn giảm bao gồm một hoặc nhiều kênh âm thanh trộn giảm làm tín hiệu âm thanh được mã hóa, trong đó số lượng của một hoặc nhiều kênh trộn giảm nhỏ hơn số lượng của các tín hiệu đối tượng âm thanh có trong chính xác hai nhóm,

xác định thông tin độ to trên các tín hiệu đối tượng âm thanh, trong đó thông tin độ to bao gồm một hoặc nhiều giá trị độ to, trong đó mỗi giá trị trong số một hoặc nhiều giá trị độ to phụ thuộc vào một hoặc nhiều tín hiệu đối tượng âm thanh,

bằng cách xác định giá trị độ to thứ nhất, giá trị độ to thứ hai và giá trị độ to thứ ba của thông tin độ to, giá trị độ to thứ nhất cho biết tổng độ to của một hoặc nhiều tín hiệu đối tượng âm thanh của nhóm thứ nhất, giá trị độ to thứ hai

cho biết tổng độ to của một hoặc nhiều tín hiệu đối tượng âm thanh của nhóm thứ hai, và giá trị độ to thứ ba cho biết tổng độ to của một hoặc nhiều tín hiệu đối tượng âm thanh đường vòng thêm của nhóm thứ ba, hoặc

bằng cách xác định giá trị độ to thứ nhất và giá trị độ to thứ hai của thông tin độ to, giá trị độ to thứ nhất cho biết tổng độ to của một hoặc nhiều tín hiệu đối tượng âm thanh của nhóm thứ nhất, giá trị độ to thứ hai cho biết tổng độ to của một hoặc nhiều tín hiệu đối tượng âm thanh của nhóm thứ hai và của một hoặc nhiều tín hiệu đối tượng âm thanh đường vòng thêm của nhóm thứ ba,

mã hóa thông tin độ to trên các tín hiệu đối tượng âm thanh,

nhận một hoặc nhiều tín hiệu đối tượng âm thanh đường vòng thêm, trong đó mỗi tín hiệu trong số một hoặc nhiều tín hiệu đối tượng âm thanh đường vòng thêm được ấn định cho nhóm thứ ba, trong đó mỗi tín hiệu trong số một hoặc nhiều tín hiệu đối tượng âm thanh đường vòng thêm không có trong nhóm thứ nhất và không có trong nhóm thứ hai, và

không trộn giảm một hoặc nhiều tín hiệu đối tượng âm thanh đường vòng thêm trong tín hiệu trộn giảm.

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

23. Vật ghi đọc được bằng máy tính chứa chương trình máy tính để thực hiện phương pháp theo điểm 19, khi được thực hiện trên máy tính hoặc bộ xử lý tín hiệu.

24. Vật ghi đọc được bằng máy tính chứa chương trình máy tính để thực hiện phương pháp theo điểm 20, khi được thực hiện trên máy tính hoặc bộ xử lý tín hiệu.

25. Vật ghi đọc được bằng máy tính chứa chương trình máy tính để thực hiện phương pháp theo điểm 21, khi được thực hiện trên máy tính hoặc bộ xử lý tín hiệu.

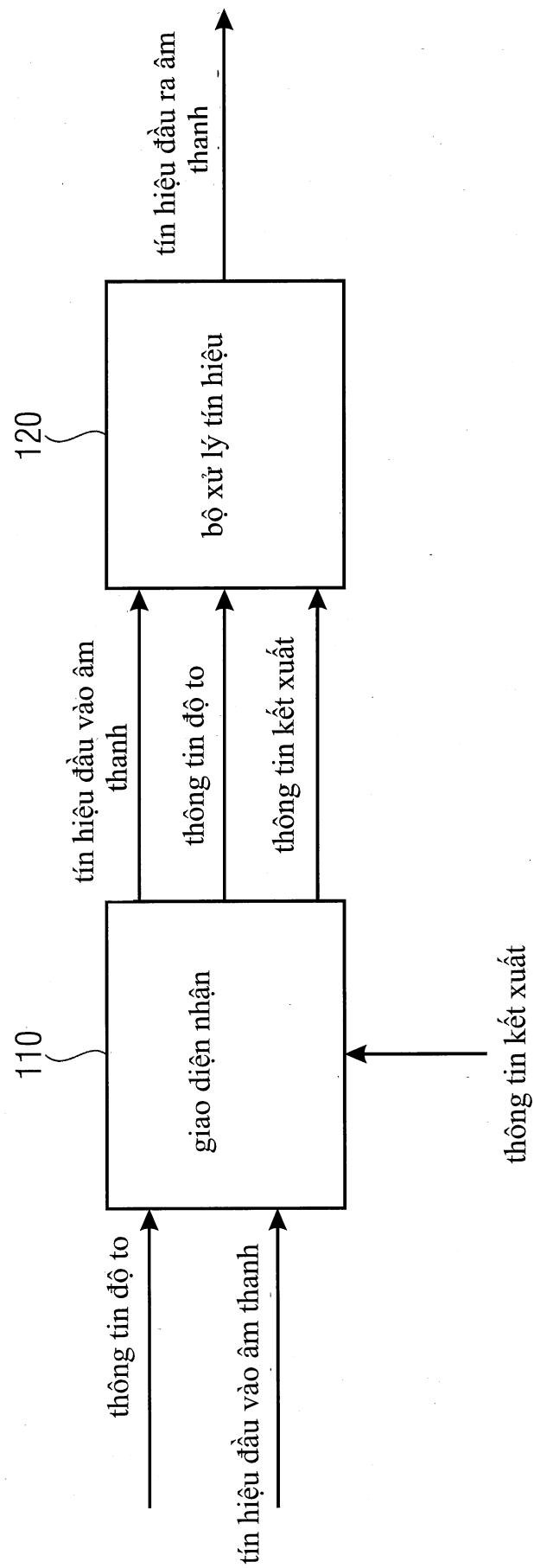


FIG 1

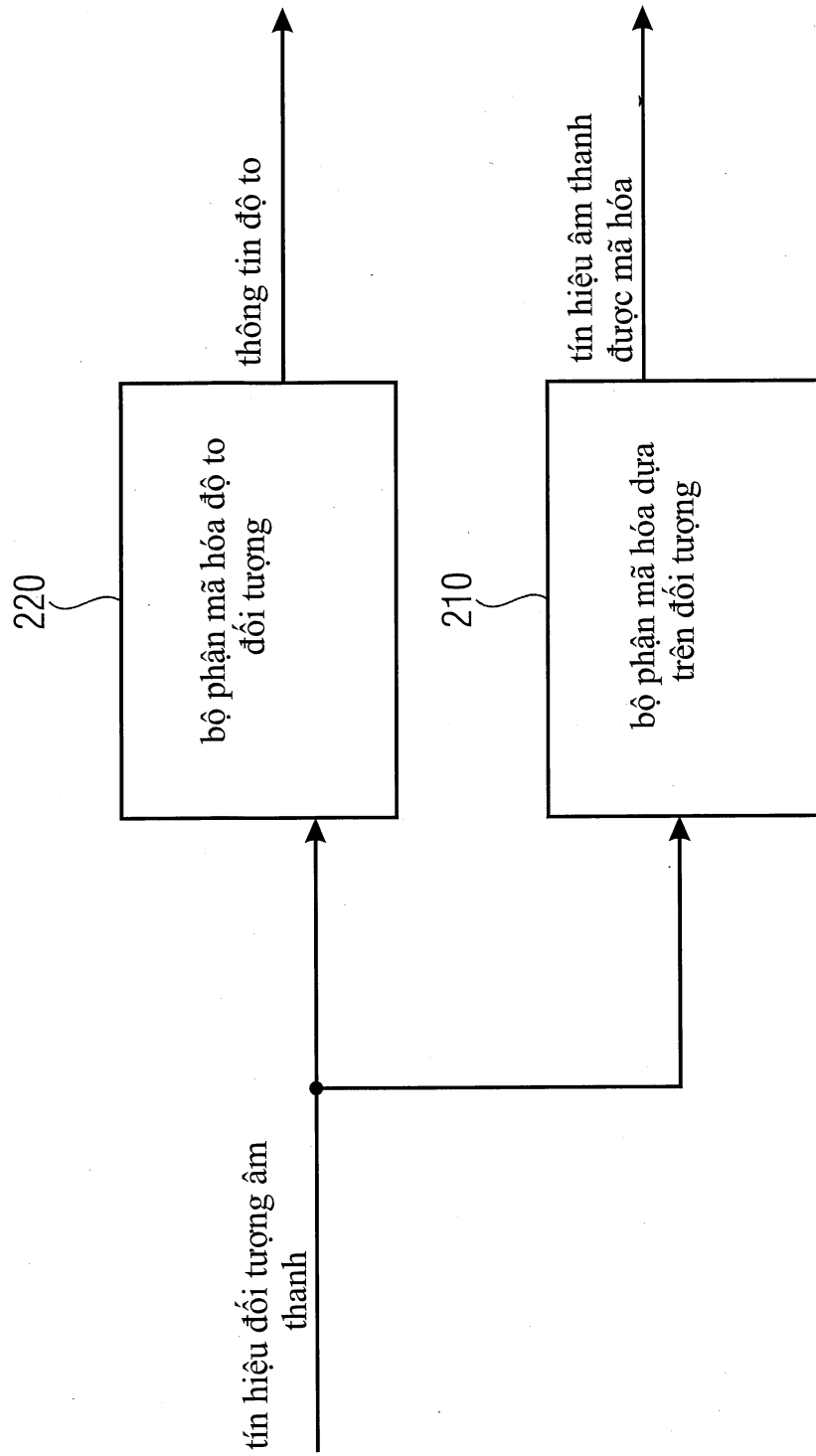


FIG 2

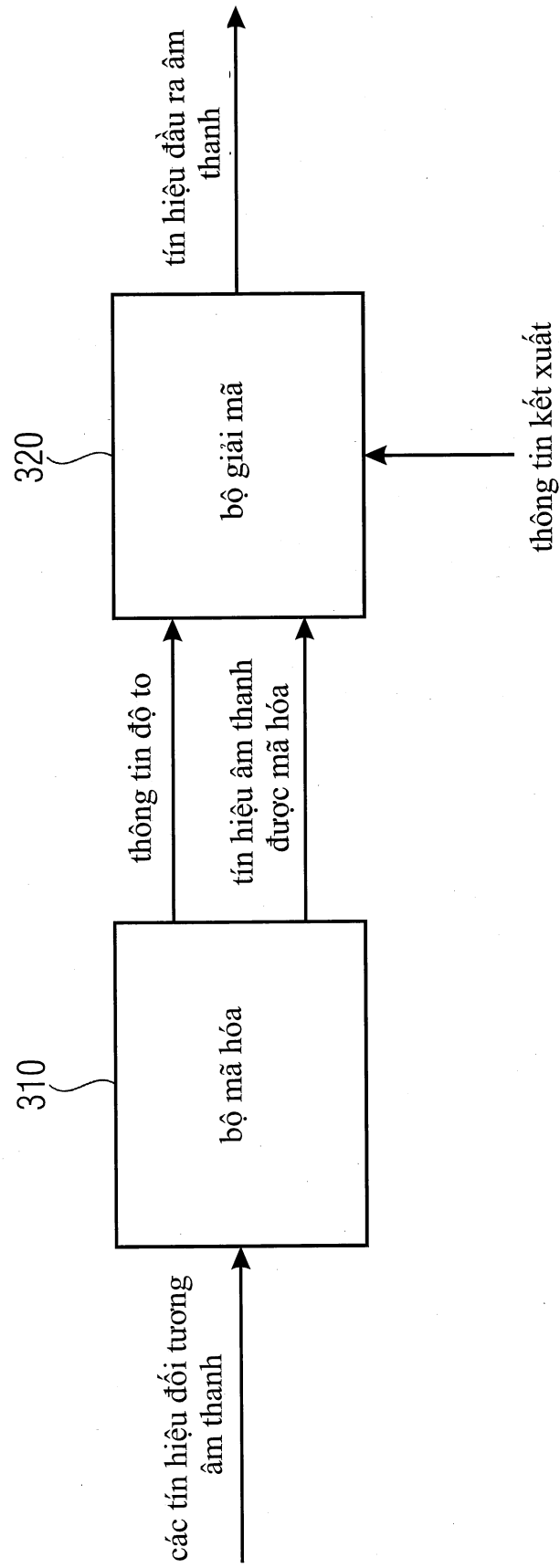


FIG 3

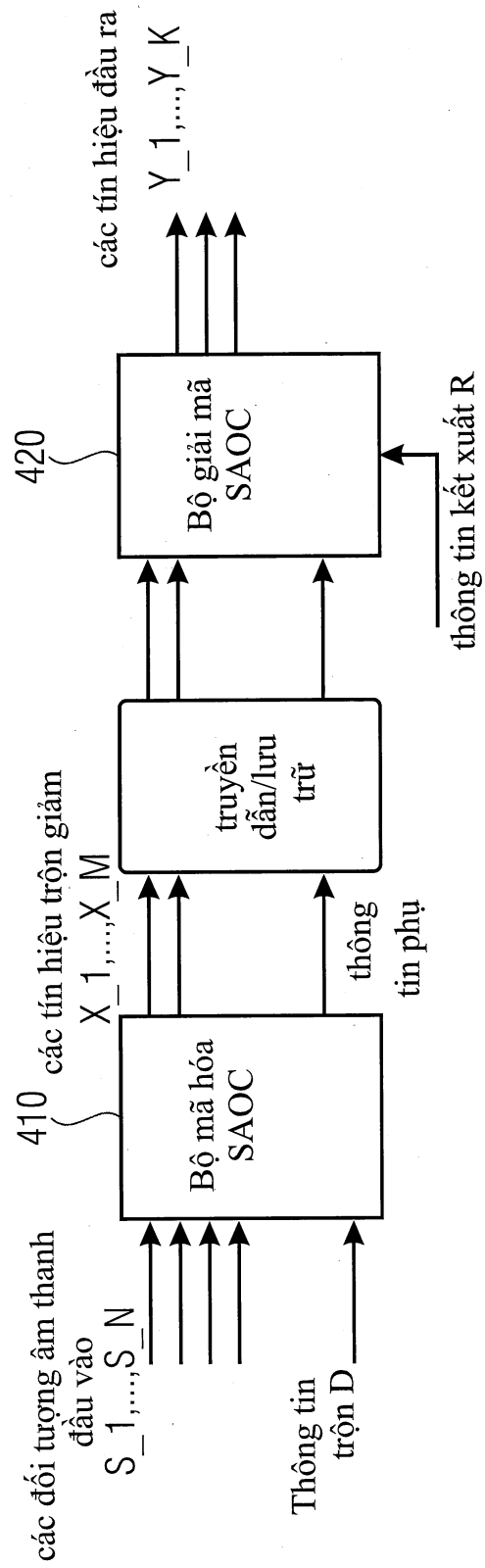


FIG 4

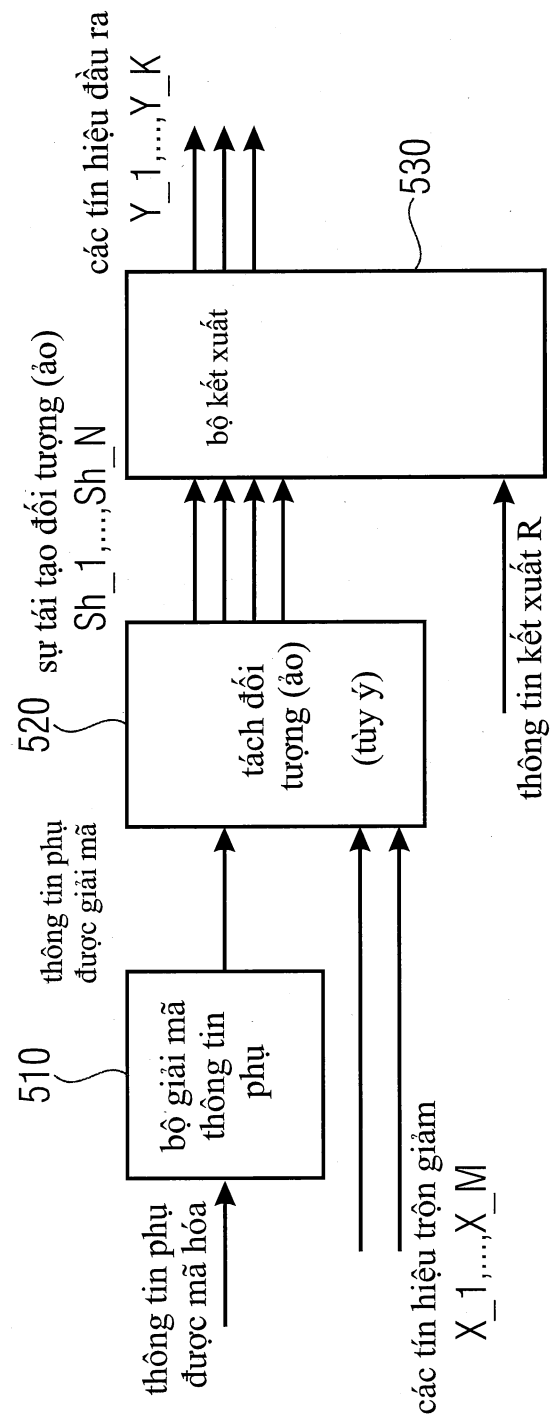


FIG 5

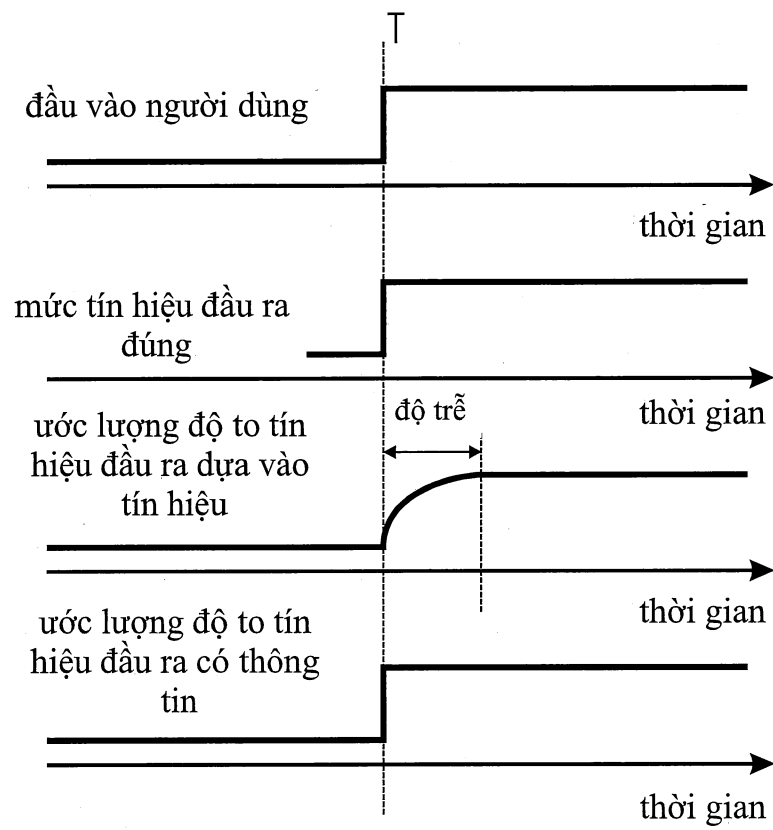


FIG 6

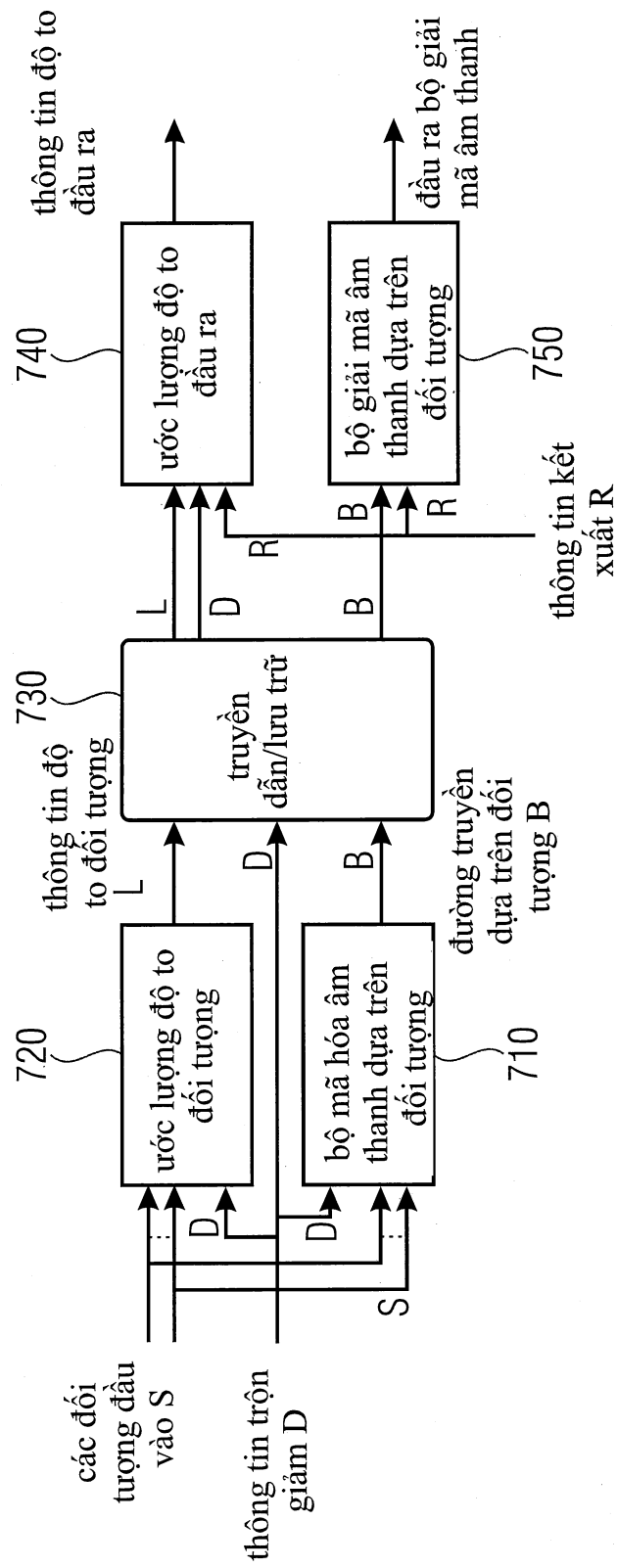


FIG 7

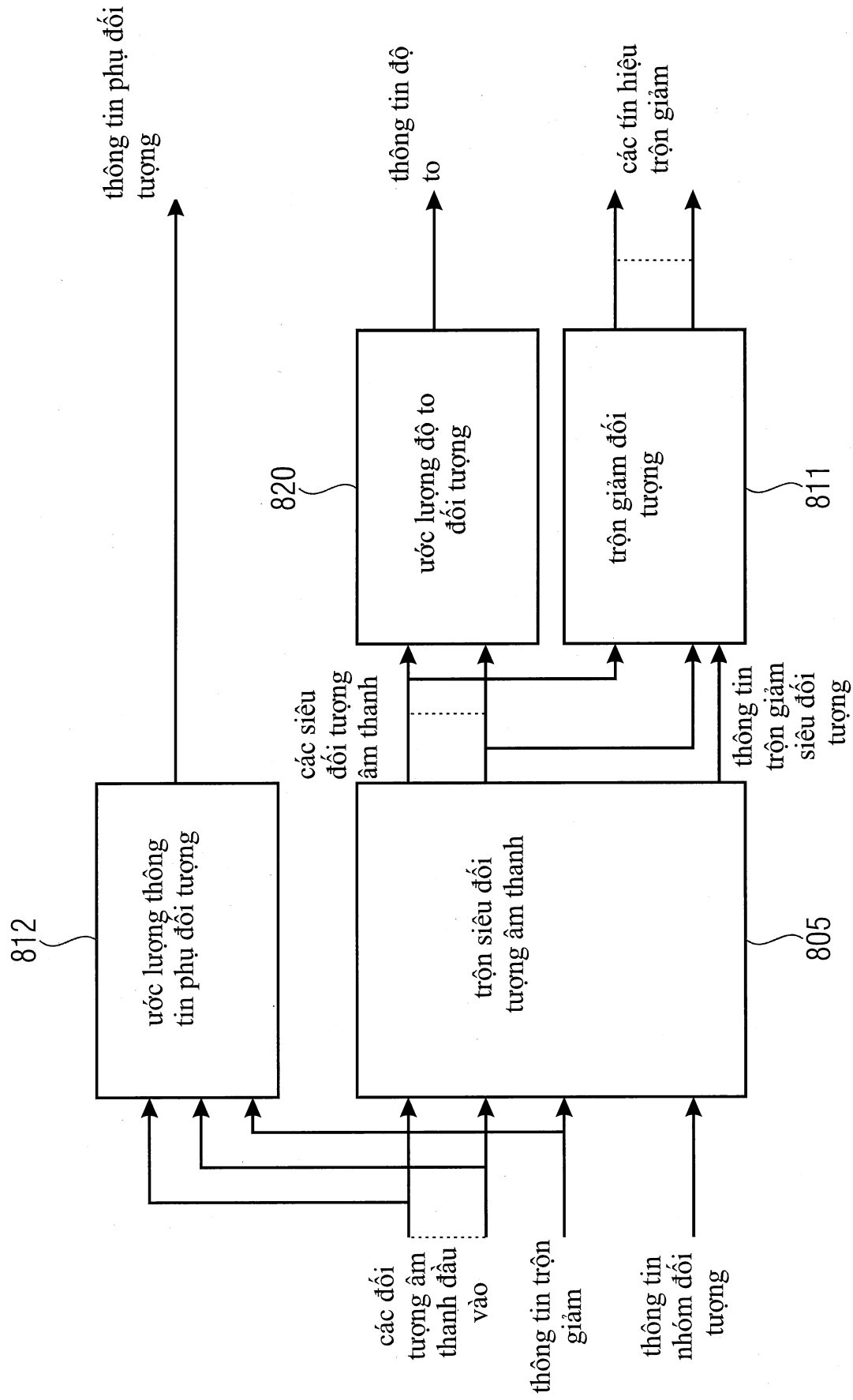


FIG 8

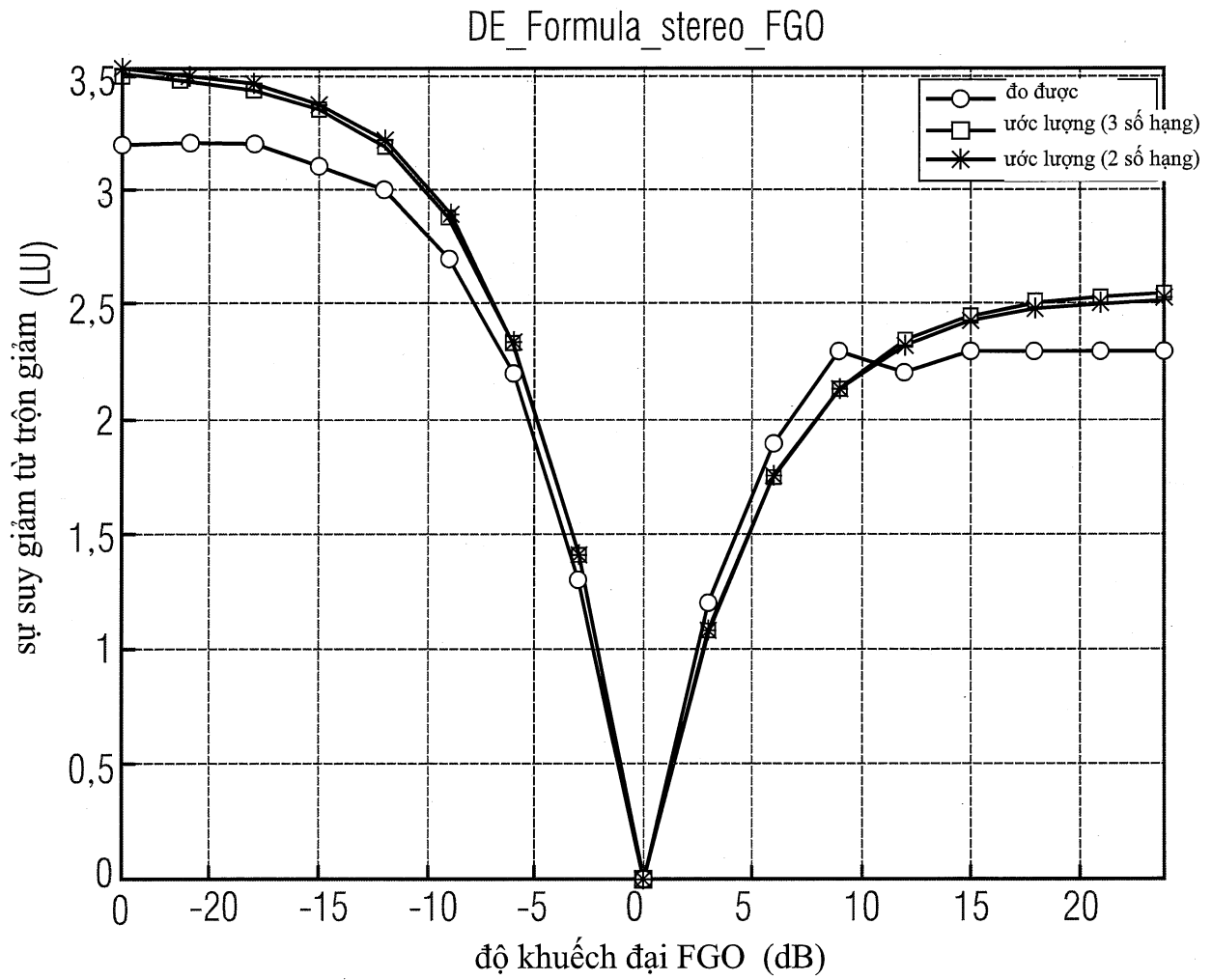


FIG 10

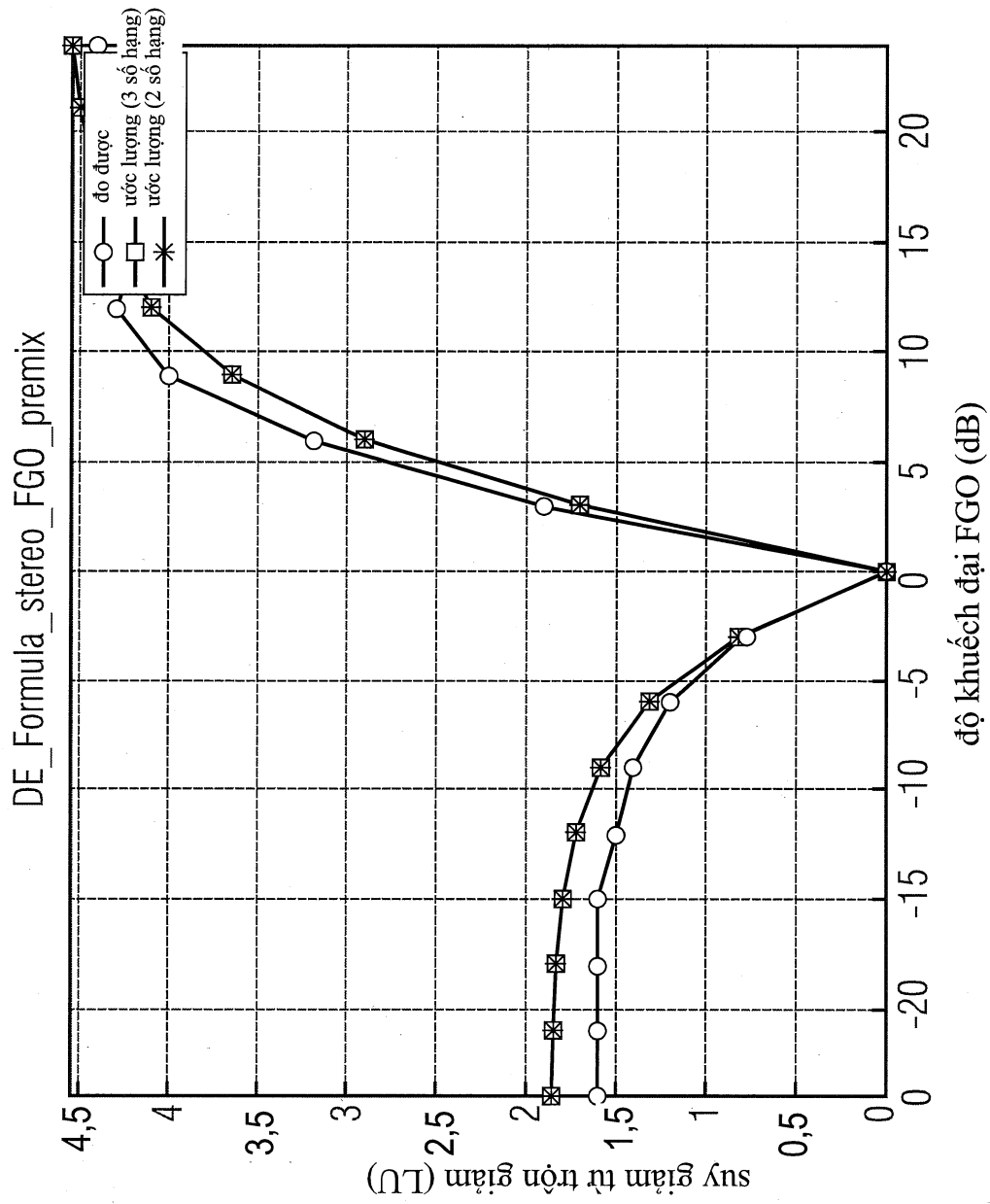


FIG 11

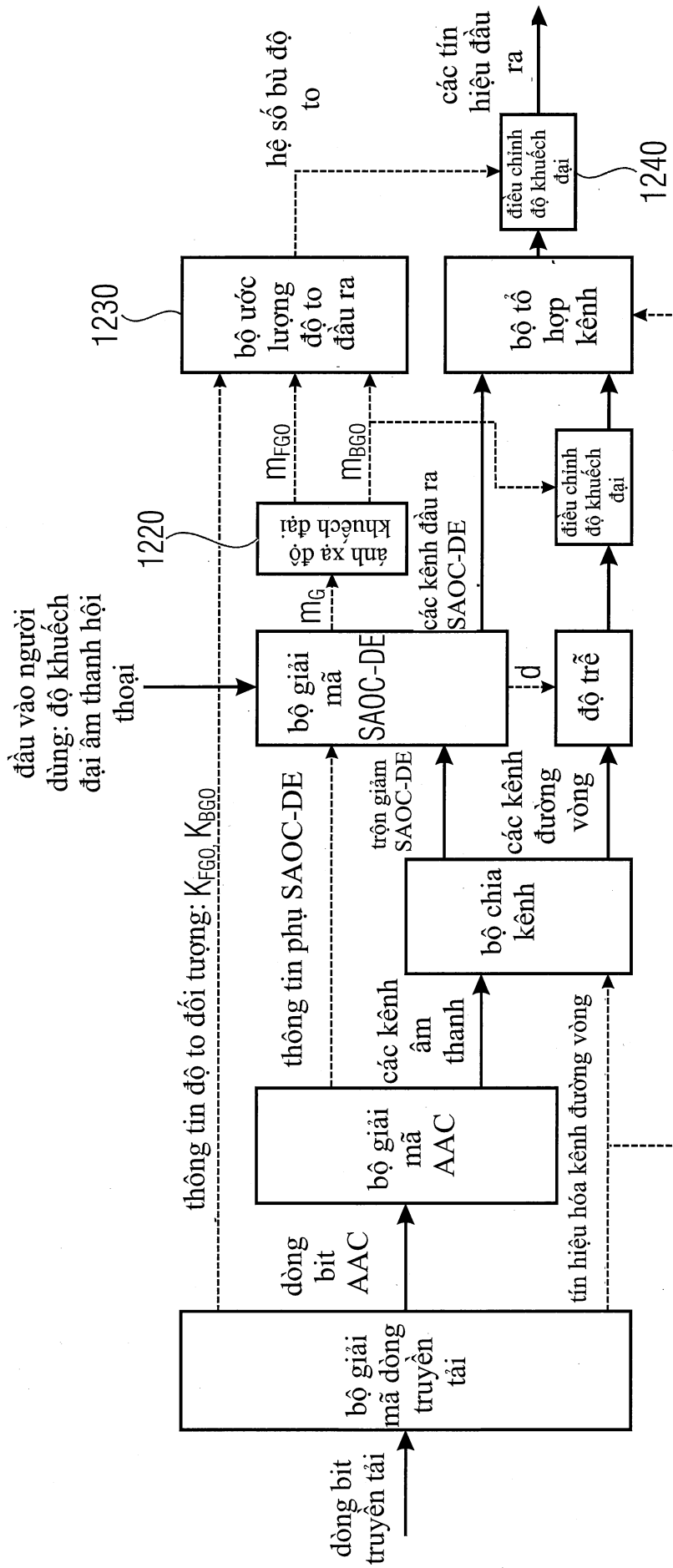


FIG 12