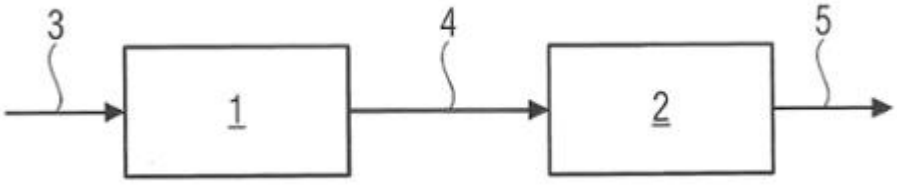
(12) BẢN MÔ TẢ SÁNG CHẾ THUỘC BẰNG ĐỘC QUYỀN SÁNG CHẾ
(19) **Cộng hòa xã hội chủ nghĩa Việt Nam (VN)**
CỤC SỞ HỮU TRÍ TUỆ
(11) **1-0034264**
(51)[7] **G10L 19/038**
(13) **B**

(21) 1-2018-04489
(22) 10/03/2017
(86) PCT/EP2017/055716 10/03/2017
(87) WO 2017/157800 21/09/2017
(30) 16160369.1 15/03/2016 EP; 16189893.7 21/09/2016 EP
(45) 26/12/2022 417
(43) 25/01/2019 370A
(73) FRAUNHOFER-GESELLSCHAFT ZUR FOERDERUNG DER ANGEWANDTEN FORSCHUNG E.V. (DE)
Hansastrasse 27c, 80686 Muenchen, Germany
(72) BAECKSTROEM, Tom (FI); GHIDO, Florin (RO); FISCHER, Johannes (DE).
(74) CÔNG TY LUẬT TRÁCH NHIỆM HỮU HẠN AMBYS HÀ NỘI (AMBYS HANOI)

(54) THIẾT BỊ MÃ HÓA, THIẾT BỊ GIẢI MÃ, PHƯƠNG PHÁP MÃ HÓA TÍN HIỆU ĐẦU VÀO, PHƯƠNG PHÁP GIẢI MÃ TÍN HIỆU ĐƯỢC MÃ HÓA, PHƯƠNG PHÁP MÃ HÓA ĐỂ XỬ LÝ TÍN HIỆU ĐẦU VÀO VÀ PHƯƠNG PHÁP GIẢI MÃ ĐỂ XỬ LÝ TÍN HIỆU ĐƯỢC MÃ HÓA

(57) Sáng chế đề cập đến thiết bị mã hóa, thiết bị giải mã, phương pháp mã hóa tín hiệu đầu vào, phương pháp giải mã tín hiệu được mã hóa, phương pháp mã hóa để xử lý tín hiệu đầu vào và phương pháp giải mã để xử lý tín hiệu được mã hóa. Thiết bị (1) xử lý tín hiệu đầu vào (3), bao gồm bộ gán trọng số theo cảm quan (10) và bộ lượng tử hóa (14). Bộ gán trọng số theo cảm quan (10) bao gồm bộ cấp mô hình (12) và bộ áp dụng mô hình (13). Bộ cấp mô hình (12) cung cấp mô hình được gán trọng số theo cảm quan (W) dựa trên tín hiệu đầu vào (3). Bộ áp dụng mô hình (13) cung cấp phổ được gán trọng số theo cảm quan (x) bằng cách áp dụng mô hình được gán trọng số theo cảm quan (W) cho phổ (s) dựa trên tín hiệu đầu vào (3). Bộ lượng tử hóa (14) được tạo cấu hình để lượng tử hóa phổ được gán trọng số theo cảm quan (x) và để cung cấp dòng bit ($\hat{u}$). Bộ lượng tử hóa (14) bao gồm bộ áp dụng ma trận ngẫu nhiên (16) và bộ tính toán hàm dấu (17). Bộ áp dụng ma trận ngẫu nhiên (16) được tạo cấu hình để áp dụng ma trận ngẫu nhiên (P) cho phổ được gán trọng số theo cảm quan (x) để cung cấp phổ được biến đổi (u). Bộ tính toán hàm dấu (17) được tạo cấu hình để tính toán hàm dấu của các thành phần của phổ biến đổi (u) để cung cấp dòng bit ($\hat{u}$). Sáng chế còn đề cập đến thiết bị (2) để xử lý tín hiệu được mã hóa (4) và các phương pháp tương ứng.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến thiết bị mã hóa để xử lý tín hiệu đầu vào và thiết bị giải mã để xử lý tín hiệu được mã hóa. Sáng chế cũng đề cập đến các phương pháp tương ứng và chương trình máy tính.

Tình trạng kỹ thuật của sáng chế

Phần trung tâm của các bộ mã hóa-giải mã âm thanh và tiếng nói là các mô hình cảm quan, mà mô tả sự quan trọng theo cảm quan của các lỗi trong các phần tử khác nhau của sự biểu diễn tín hiệu. Thực tế, các mô hình cảm quan bao gồm các thừa số có trọng số phụ thuộc vào tín hiệu mà được sử dụng trong phép lượng tử hóa của từng phần tử. Với hiệu quả tối ưu, sẽ mong muốn sử dụng cùng mô hình cảm quan tại bộ giải mã. Trong khi mô hình cảm quan là phụ thuộc tín hiệu, tuy nhiên, nó không được biết trước tại bộ giải mã, do đó các bộ mã hóa-giải mã âm thanh thường truyền mô hình này một cách rõ ràng, với chi phí tiêu thụ bit tăng.

Thời đại internet vạn vật (Internet of Things - IoT) đang tới gần, do đó thế hệ kế tiếp của các bộ mã hóa âm thanh và tiếng nói cần đi theo nó. Tuy nhiên, các mục đích thiết kế các hệ thống IoT ít thích hợp với thiết kế cổ điển của các bộ mã hóa âm thanh và tiếng nói, do đó yêu cầu việc thiết kế lại các bộ mã hóa lớn hơn.

Trước hết, trong khi bộ mã hóa âm thanh và tiếng nói trong tình trạng kỹ thuật đã biết như AMR-WB, EVS, USAC và AAC bao gồm các bộ mã hóa phức hợp và thông minh và các bộ giải mã tương đối đơn giản [1 – 4], vì IoT cần hỗ trợ các nút cảm biến phức tạp thấp được phân phối, các bộ mã hóa tốt hơn là đơn giản.

Thứ hai, vì các nút cảm biến đang mã hóa cùng tín hiệu nguồn nên ứng dụng của cùng phép lượng tử hóa ở từng nút cảm biến sẽ thể hiện việc mã hóa quá mức và có khả năng tổn hao nghiêm trọng về hiệu quả. Đặc biệt là, vì mô hình cảm quan cần giống nhiều hơn hoặc ít hơn ở mọi nút, nên việc truyền nó từ mọi nút hầu như là mã hóa quá mức thuần túy.

Các phương pháp mã hóa âm thanh và tiếng nói thông thường bao gồm ba phần:

1. mô hình cảm quan mà định rõ tác động tương đối của các lỗi trong các tham số khác nhau của bộ mã hóa-giải mã,
2. mô hình nguồn mà mô tả phạm vi và khả năng của các đầu vào khác nhau và
3. bộ mã hóa entropy mà sử dụng mô hình nguồn để giảm thiểu sự biến dạng theo cảm quan [5].

Hơn nữa, mô hình cảm quan có thể được áp dụng theo một trong hai cách:

1. Các tham số tín hiệu có thể được gán trọng số theo mô hình cảm quan, sao cho các tham số sau đó có thể được lượng tử hóa với cùng độ chính xác. Mô hình cảm quan sau đó phải được truyền đến bộ giải mã sao cho việc gán trọng số có thể không được thực hiện.
2. Mô hình cảm quan ngoài ra có thể được áp dụng như mô hình đánh giá, sao cho đầu ra tổng hợp của các phép lượng tử hóa khác nhau được so sánh, được gán trọng số bởi mô hình cảm quan, theo phép lặp phân tích bằng tổng hợp. Dù ở đây mô hình cảm quan không phải được truyền, phương thức này có bất lợi là hình dạng các ô lượng tử hóa không được tạo dạng đồng đều mà giảm hiệu quả mã hóa. Tuy nhiên, nghiêm trọng hơn, để thấy phép lượng tử hóa tối ưu, thuật toán tìm kiếm brute-force phức tạp tính toán của các phép lượng tử hóa khác nhau phải được sử dụng.

Vì phương pháp phân tích bằng tổng hợp dẫn đến bộ mã hóa phức tạp tính toán nên không phải là thay thế khả thi cho IoT. Do đó, bộ giải mã phải được truy cập tới mô hình cảm quan. Tuy nhiên, như được lưu ý ở trên, việc truyền rõ ràng của mô hình cảm quan (hoặc tương đương là, mô hình đường bao của phổ tín hiệu) là không mong muốn vì nó làm giảm hiệu quả mã hóa.

Tài liệu tham khảo

[1] TS 26.445, EVS Codec Detailed Algorithmic Description; 3GPP Technical Specification (Release 12), 3GPP, 2014.

[2] TS 26.190, Adaptive Multi-Rate (AMR-WB) speech codec, 3GPP, 2007.

[3] ISO/IEC 23003–3:2012, "MPEG-D (MPEG audio technologies), Part 3: Unified speech and audio coding," 2012.

[4] M. Bosi, K. Brandenburg, S. Quackenbush, L. Fielder, K. Akagiri, H. Fuchs, and M. Dietz, "ISO/IEC MPEG-2 advanced audiocoding," Journal of the Audio engineering society, vol. 45, no. 10, pp. 789–814, 1997.

[5] M. Bosi and R. E. Goldberg, Introduction to Digital Audio Coding and Standards. Dordrecht, The Netherlands: Kluwer Academic Publishers, 2003.

[6] P. T. Boufounos and R. G. Baraniuk, "1-bit compressive sensing," in Information Sciences and Systems, 2008. CISS 2008. 42nd Annual Conference on. IEEE, 2008, pp. 16–21.

[7] Z. Xiong, A. D. Liveris, and S. Cheng, "Distributed source coding for sensor networks," IEEE Signal Process. Mag., vol. 21, no. 5, pp. 80–94, 2004.

[8] Z. Xiong, A. D. Liveris, and Y. Yang, "Distributed source coding," Handbook on Array Processing and Sensor Networks, pp. 609– 643, 2009.

[9] B. Girod, A. M. Aaron, S. Rane, and D. Rebollo-Monedero, "Distributed video coding," Proc. IEEE, vol. 93, no. 1, pp. 71–83, 2005.

[10] A. Majumdar, K. Ramchandran, and L. Kozintsev, "Distributed coding for wireless audio sensors," in Applications of Signal Processing to Audio and Acoustics, 2003 IEEE Workshop on. IEEE, 2003, pp. 209–212.

[11] H. Dong, J. Lu, and Y. Sun, "Distributed audio coding in wireless sensor networks," in Computational Intelligence and Security, 2006 International Conference on, vol. 2. IEEE, 2006, pp. 1695–1699.

[12] A. Zahedi, J. Østergaard, S. H. Jensen, P. Naylor, and S. Bech, "Coding and enhancement in wireless acoustic sensor networks," in Data Compression Conference (DCC), 2015. IEEE, 2015, pp. 293–302.

[13] A. Zahedi, J. Østergaard, S. H. Jensen, S. Bech, and P. Naylor, "Audio coding in wireless acoustic sensor networks," Signal Processing, vol. 107, pp. 141–152, 2015.

[14] US 7,835,904.

[15] G. Kubin and W. B. Kleijn, "Multiple-description coding (MDC) of speech with an invertible auditory model," in Speech Coding, IEEE Workshop on,

1999, pp. 81–83.

[16] V. K. Goyal, "Multiple description coding: Compression meets the network," IEEE Signal Process. Mag., vol. 18, no. 5, pp. 74– 93, 2001.

[17] J. O. Smith III and J. S. Abel, "Bark and ERB bilinear transforms," IEEE Trans. Speech Audio Process., vol. 7, no. 6, pp. 697–708, 1999.

[18] T. Bäckström, "Vandermonde factorization of Toeplitz matrices and applications in filtering and warping," IEEE Trans. Signal Process., vol. 61, no. 24, pp. 6257–6263, Dec. 2013.

[19] F. Zheng, G. Zhang, and Z. Song, "Comparison of different implementations of MFCC," Journal of Computer Science and Technology, vol. 16, no. 6, pp. 582–589, 2001.

[20] H. Fastl and E. Zwicker, Psychoacoustics: Facts and models. Springer, 2006, vol. 22.

[21] NTT-AT, "Super wideband stereo speech database," http://www.ntt-at.com/product/widebandspeech, accessed: 09.09.2014. [Online]. Available: http://www.ntt-at.com/product/ widebandspeech

[22] S. Korse, T. Jähnel, and T. Bäckström, "Entropy coding of spectral envelopes for speech and audio coding using distribution quantization," in Proc. Interspeech, 2016.

Bản chất kỹ thuật của sáng chế

Mục đích của sáng chế là thể hiện cách để phục hồi mô hình cảm quan ở bộ giải mã từ tín hiệu được truyền mà không có thông tin phụ liên quan đến mô hình cảm quan.

Mục đích của sáng chế đạt được bởi thiết bị mã hóa để xử lý tín hiệu đầu vào, cũng như thiết bị giải mã để xử lý tín hiệu được mã hóa. Mục đích này cũng đạt được bởi các phương pháp và hệ thống tương ứng bao gồm thiết bị mã hóa và thiết bị giải mã.

Sáng chế đặc biệt hữu ích trong các mạng cảm biến được phân phối và internet vạn vật, trong đó chi phí tiêu thụ bit bổ sung từ việc truyền mô hình cảm quan tăng với

số lượng cảm biến.

Theo một phương án, sáng chế cũng có thể được ký hiệu như sự phục hồi mù của các mô hình cảm quan trong mã hóa âm thanh và tiếng nói được phân phối.

Tín hiệu đầu vào là tín hiệu âm thanh và/hoặc tiếng nói sẽ được mã hóa bởi thiết bị mã hóa.

Mục đích của sáng chế đạt được bởi thiết bị mã hóa để xử lý tín hiệu đầu vào. Thiết bị mã hóa tốt hơn là cung cấp tín hiệu được mã hóa.

Thiết bị mã hóa bao gồm bộ gán trọng số theo cảm quan và bộ lượng tử hóa.

Bộ gán trọng số theo cảm quan bao gồm bộ cấp mô hình và bộ áp dụng mô hình. Bộ cấp mô hình được tạo cấu hình để cung cấp mô hình được gán trọng số theo cảm quan dựa trên tín hiệu đầu vào. Bộ áp dụng mô hình được tạo cấu hình để cung cấp phổ được gán trọng số theo cảm quan bằng cách áp dụng mô hình được gán trọng số theo cảm quan cho phổ dựa trên tín hiệu đầu vào.

Bộ lượng tử hóa được tạo cấu hình để lượng tử hóa phổ được gán trọng số theo cảm quan và để cung cấp dòng bit. Bộ lượng tử hóa bao gồm bộ áp dụng ma trận ngẫu nhiên và bộ tính toán hàm dấu. Bộ áp dụng ma trận ngẫu nhiên được tạo cấu hình để áp dụng ma trận ngẫu nhiên cho phổ được gán trọng số theo cảm quan để cung cấp phổ được biến đổi. Bộ tính toán hàm dấu được tạo cấu hình để tính toán hàm dấu (hoặc: hàm signum) của các thành phần của véctơ được biến đổi để cung cấp dòng bit.

Phép lượng tử hóa bao gồm ít nhất hai bước: trong bước thứ nhất, phổ được gán trọng số theo cảm quan được tổ hợp với ma trận ngẫu nhiên. Ma trận ngẫu nhiên này có lợi ích là đối với từng tín hiệu đầu vào, ma trận khác nhau được sử dụng. Điều này trở nên hữu ích khi nhiều cảm biến bao hàm cùng nguồn âm thanh và việc mã hóa quá mức phải được tránh khỏi. Bước thứ hai bao gồm hàm dấu được tính toán của các thành phần của véctơ được biến đổi.

Trong một phương án, bộ cấp mô hình được tạo cấu hình để cung cấp mô hình được gán trọng số theo cảm quan dựa trên phép nén phổ dựa trên tín hiệu đầu vào.

Trong một phương án, trong thiết bị mã hóa theo sáng chế, việc gán trọng số của tín hiệu đầu vào được dựa trên phép nén tín hiệu đầu vào. Trong một phương án,

việc gán trọng số dựa trên phép nén đường bao của phổ cường độ của tín hiệu đầu vào (đường bao là liên tục, thường có dạng nhẵn mô tả đặc tính của tín hiệu, ở đây là phổ cường độ). Dựa trên phép nén mô hình được gán trọng số theo cảm quan thu được, mà được sử dụng cuối cùng để gán trọng số theo cảm quan phổ của tín hiệu đầu vào.

Trong một phương án, thiết bị mã hóa cung cấp dòng bit với thông tin phụ bao hàm một số khía cạnh về quy trình mã hóa.

Theo một phương án, thông tin về phép lượng tử hóa được cung cấp như thông tin phụ bởi tín hiệu được mã hóa được xuất ra bởi thiết bị mã hóa như kết quả của quy trình mã hóa.

Trong phương án nữa, bộ gán trọng số theo cảm quan bao gồm bộ tính toán đường bao. Bộ tính toán đường bao được tạo cấu hình để cung cấp đường bao của phổ cường độ dựa trên tín hiệu đầu vào.

Trong một phương án, bộ cấp mô hình được tạo cấu hình để tính toán hàm nén mô tả phép nén của đường bao. Hơn nữa, bộ cấp mô hình được tạo cấu hình để tính toán mô hình được gán trọng số theo cảm quan dựa trên hàm nén. Trong phương án này, bộ cấp mô hình nén đường bao và tính toán hàm mô tả phép nén này. Dựa trên hàm này, thu được mô hình được gán trọng số theo cảm quan. Trong phương án này, phép nén đường bao của phổ dựa trên tín hiệu đầu vào được thực hiện, tức là phạm vi cường độ được giảm và, do đó, nhỏ hơn trước khi nén. Bằng cách nén đường bao, phổ như vậy cũng được nén, tức là phạm vi cường độ của phổ được giảm.

Trong phương án khác, mô hình được gán trọng số theo cảm quan hoặc hàm nén được tính toán trực tiếp từ tín hiệu đầu vào hoặc phổ cường độ/công suất dựa trên tín hiệu đầu vào.

Theo phương án, bộ cấp mô hình được tạo cấu hình để tính toán hàm nén mô tả phép nén phổ dựa trên tín hiệu đầu vào hoặc mô tả phép nén đường bao của phổ cường độ dựa trên tín hiệu đầu vào. Phép nén làm giảm phạm vi cường độ của phổ dựa trên tín hiệu đầu vào hoặc giảm phạm vi cường độ của đường bao. Hơn nữa, bộ cấp mô hình được tạo cấu hình để tính toán mô hình được gán trọng số theo cảm quan dựa trên hàm nén.

Trong phương án, hàm nén để nén, ví dụ, phổ hoặc đường bao cần đáp ứng hai

tiêu chuẩn:

Thứ nhất, hàm nén tăng hoàn toàn. Điều này ngụ ý rằng đối với giá trị vô hướng dương bất kỳ và giá trị nhỏ tùy chọn, giá trị hàm đối với giá trị vô hướng dương nhỏ hơn giá trị hàm đối với tổng của giá trị vô hướng dương và của giá trị nhỏ tùy chọn.

Thứ hai, đối với giá trị vô hướng dương thứ nhất và giá trị vô hướng dương thứ hai lớn hơn giá trị vô hướng dương thứ nhất, chênh lệch giữa hai giá trị hàm cho giá trị vô hướng dương thứ hai và giá trị hàm cho giá trị vô hướng dương thứ nhất là nhỏ hơn chênh lệch giữa giá trị vô hướng dương thứ hai và giá trị vô hướng dương thứ nhất.

Mục đích cũng đạt được bởi phương pháp xử lý tín hiệu đầu vào. Tín hiệu đầu vào này tốt hơn là tín hiệu âm thanh và/hoặc tín hiệu tiếng nói.

Phương pháp xử lý tín hiệu đầu vào là tín hiệu âm thanh và/hoặc tiếng nói bao gồm ít nhất các bước sau:

tính toán mô hình được gán trọng số theo cảm quan dựa trên tín hiệu đầu vào,

cung cấp phổ được gán trọng số theo cảm quan bằng cách áp dụng mô hình được gán trọng số theo cảm quan cho phổ dựa trên tín hiệu đầu vào, và

lượng tử hóa phổ được gán trọng số theo cảm quan để cung cấp dòng bit,

trong đó bước lượng tử hóa phổ được gán trọng số theo cảm quan bao gồm:

áp dụng ma trận ngẫu nhiên cho phổ được gán trọng số theo cảm quan để cung cấp phổ được biến đổi, và

tính toán hàm dấu của các thành phần của phổ biến đổi để cung cấp dòng bit.

Trong phương án, các bước sau đây được thực hiện:

tính toán đường bao của phổ cường độ dựa trên tín hiệu đầu vào,

tính toán mô hình được gán trọng số theo cảm quan dựa trên phép nén đường bao.

Các phương án của thiết bị mã hóa cũng có thể được thực hiện bởi các bước của phương pháp và các phương án tương ứng của phương pháp. Do đó, các giải thích đã nêu cho các phương án của thiết bị cũng đúng cho phương pháp.

Mục đích của sáng chế còn đạt được bởi phương pháp xử lý tín hiệu đầu vào,

bao gồm các bước:

cung cấp mô hình gán trọng số theo cảm quan dựa trên tín hiệu đầu vào,

gán trọng số phổ của tín hiệu đầu vào bằng cách áp dụng mô hình gán trọng số theo cảm quan cho phổ của tín hiệu đầu vào, và

lượng tử hóa phổ được gán trọng số bằng cách tính toán hàm dấu hoặc các phép chiếu ngẫu nhiên của phổ được gán trọng số.

Theo các phương án, phương pháp còn bao gồm bước:

thu được các phép chiếu ngẫu nhiên của phổ được gán trọng số bằng cách áp dụng ma trận ngẫu nhiên cho phổ được gán trọng số.

Trong một phương án, bước cung cấp mô hình gán trọng số theo cảm quan bao gồm nén đường bao phổ cường độ của tín hiệu đầu vào.

Theo phương án, phương pháp còn bao gồm bước:

thu được đường bao bằng cách sử dụng giàn lọc và ma trận chéo chứa các thừa số chuẩn cho từng băng.

Mục đích cũng đạt được bởi thiết bị giải mã để xử lý tín hiệu được mã hóa.

Thiết bị giải mã bao gồm ít nhất bộ giải lượng tử hóa và bộ giải trọng số theo cảm quan.

Bộ giải lượng tử hóa được tạo cấu hình để giải lượng tử hóa dòng bit được chứa bởi tín hiệu được mã hóa và để cung cấp phổ được gán trọng số theo cảm quan được tính toán. Hơn nữa, bộ giải lượng tử hóa được tạo cấu hình để giải lượng tử hóa dòng bit bằng cách áp dụng sự giả nghịch đảo của ma trận ngẫu nhiên cho dòng bit. Bộ giải lượng tử hóa đảo ngược các tác động của phép lượng tử hóa xảy ra trong suốt quy trình mã hóa. Bộ giải lượng tử hóa được theo dõi bởi bộ giải trọng số theo cảm quan sao cho phổ thu được từ phép giải lượng tử hóa được giải trọng số theo cảm quan. Phổ được gán trọng số theo cảm quan được tính toán sẽ được nhận bởi bộ giải trọng số để được giải trọng số theo cảm quan. Do đó, phổ thu được cuối cùng là dòng bit đã giải lượng tử hóa được giải trọng số theo cảm quan được chứa bởi tín hiệu đầu vào.

Bộ giải trọng số theo cảm quan đảo ngược các tác động của việc gán trọng số

theo cảm quan mà xảy ra trong suốt quy trình mã hóa dẫn đến tín hiệu được mã hóa. Trong phương án, điều này được thực hiện mà không có thông tin phụ của tín hiệu được mã hóa bao gồm mô hình được gán trọng số theo cảm quan. Mô hình được khôi phục từ tín hiệu âm thanh được mã hóa như vậy.

Bộ giải trọng số theo cảm quan bao gồm bộ làm gần đúng phổ và bộ làm gần đúng mô hình.

Trong một phương án, sự khôi phục mô hình được thực hiện theo cách khác mà các giá trị bắt đầu hoặc các giá trị khởi tạo được yêu cầu. Do đó, bộ cấp giả định ban đầu được chứa bởi bộ giải trọng số theo cảm quan được tạo cấu hình để cung cấp dữ liệu cho sự giả định ban đầu của mô hình được gán trọng số theo cảm quan mà tín hiệu đã mã hóa được kết hợp. Theo một phương án, dữ liệu cho sự giả định ban đầu bao gồm véctơ với các phần tử chéo của ma trận mô tả mô hình được gán trọng số theo cảm quan.

Bộ làm gần đúng phổ được tạo cấu hình để tính toán phép làm gần đúng của phổ dựa trên phổ được gán trọng số theo cảm quan được tính toán. Hơn nữa, bộ làm gần đúng mô hình được tạo cấu hình để tính toán phép làm gần đúng của mô hình được gán trọng số theo cảm quan mà tín hiệu đã mã hóa được kết hợp (tức là được sử dụng để mã hóa tín hiệu đầu vào và, do đó, tạo ra tín hiệu được mã hóa) dựa trên phép làm gần đúng của phổ.

Theo một phương án, bộ giải trọng số theo cảm quan bao gồm bộ cấp giả định ban đầu. Bộ cấp giả định ban đầu được tạo cấu hình để cung cấp dữ liệu cho sự giả định ban đầu của mô hình được gán trọng số theo cảm quan. Bộ làm gần đúng phổ được tạo cấu hình để tính toán phép làm gần đúng của phổ dựa trên tín hiệu đã mã hóa và sự giả định ban đầu hoặc phép làm gần đúng của mô hình được gán trọng số theo cảm quan. Hơn nữa, theo phương án khác, phép làm gần đúng của phổ được dựa trên sự giả định ban đầu hoặc dựa trên phép làm gần đúng được tính toán theo cách đặc biệt của mô hình được gán trọng số theo cảm quan. Lựa chọn này phụ thuộc vào việc liệu phép lặp chỉ mới bắt đầu với sự giả định ban đầu hay ít nhất một phép lặp với ít nhất một sự cải tiến của phép làm gần đúng dẫn đến phép làm gần đúng của mô hình được gán trọng số theo cảm quan đã xảy ra. Trong một phương án, phép lặp được thực hiện đến khi tiêu chuẩn hội tụ đã được đáp ứng.

Trong một phương án, bộ giải lượng tử hóa được tạo cấu hình để giải lượng tử hóa dòng bit được chứa tín hiệu được mã hóa dựa trên thông tin phụ về ma trận ngẫu nhiên được chứa tín hiệu được mã hóa. Phương án này đề cập đến quy trình mã hóa mà phép lượng tử hóa được thực hiện bằng cách sử dụng ma trận ngẫu nhiên. Thông tin về ma trận ngẫu nhiên đã sử dụng được chứa bởi thông tin phụ của tín hiệu được mã hóa. Trong một phương án, thông tin phụ bao gồm chỉ hạt giống của các cột của ma trận ngẫu nhiên.

Trong một phương án, nhiều tín hiệu được mã hóa được xử lý theo cách kết hợp. Từng tín hiệu được mã hóa bao gồm dòng bit với thông tin phụ liên quan đến ít nhất là thông tin về sự lượng tử hóa được thực hiện trong khi cung cấp tín hiệu được mã hóa tương ứng. Với mục đích này, trong một phương án, bộ giải lượng tử hóa được tạo cấu hình để nhận nhiều tín hiệu đầu vào và để cung cấp chỉ mộ phổ được gán trọng số theo cảm quan được tính toán dựa trên nhiều tín hiệu đầu vào. Các tín hiệu đầu vào tốt hơn là đề cập đến tín hiệu âm thanh/tiếng nói bắt nguồn từ cùng nguồn tín hiệu.

Trong một phương án khác, thiết bị giải mã được tạo cấu hình để áp dụng việc mô hình hóa nguồn. Mô hình nguồn mô tả phạm vi và khả năng của các đầu vào khác nhau.

Mục đích của sáng chế cũng đạt được bởi phương pháp xử lý tín hiệu được mã hóa. Tín hiệu được mã hóa này tốt hơn là tín hiệu âm thanh được mã hóa và/hoặc tín hiệu tiếng nói được mã hóa.

Phương pháp xử lý (hoặc giải mã) tín hiệu được mã hóa bao gồm ít nhất là các bước sau:

giải lượng tử hóa dòng bit được chứa bởi tín hiệu được mã hóa và để cung cấp phổ được gán trọng số theo cảm quan được tính toán,

trong đó việc giải lượng tử hóa dòng bit bao gồm việc áp dụng sự giả nghịch đảo của ma trận ngẫu nhiên cho dòng bit,

tính toán phép làm gần đúng của phổ dựa trên phổ được gán trọng số theo cảm quan được tính toán, và

tính toán phép làm gần đúng của mô hình được gán trọng số theo cảm quan mà tín hiệu đã mã hóa được kết hợp (tức là được sử dụng để tạo ra tín hiệu được mã hóa)

dựa trên phép làm gần đúng của phổ.

Trong một phương án, phương pháp bao gồm các bước:

tính toán phép làm gần đúng của phổ hoặc dựa trên phổ được gán trọng số theo cảm quan được tính toán và sự giả định ban đầu hoặc dựa trên phổ được gán trọng số theo cảm quan được tính toán và phép làm gần đúng của mô hình được gán trọng số theo cảm quan mà tín hiệu đã mã hóa được kết hợp.

Các phép làm gần đúng được tính toán của mô hình được gán trọng số theo cảm quan tốt hơn là được sử dụng cho việc tính toán tiếp theo của phép làm gần đúng của phổ.

Mục đích của sáng chế cũng đạt được bởi phương pháp xử lý tín hiệu được mã hóa, bao gồm:

cung cấp tín hiệu theo cảm quan được lượng tử hóa bằng cách áp dụng sự giả nghịch đảo của ma trận ngẫu nhiên cho tín hiệu được mã hóa,

tính toán phép ước lượng phổ dựa trên tín hiệu theo cảm quan được lượng tử hóa, và

tính toán phép làm gần đúng của mô hình gán trọng số theo cảm quan được sử dụng để cung cấp tín hiệu được mã hóa dựa trên phép làm gần đúng của phổ.

Theo phương án, phương pháp còn bao gồm bước:

cung cấp phép làm gần đúng thứ không của mô hình gán trọng số theo cảm quan sử dụng sự giả định ban đầu, và

tính toán phép ước lượng thứ không của phổ dựa trên phép làm gần đúng thứ không của mô hình gán trọng số theo cảm quan.

Theo phương án nữa, phương pháp còn bao gồm bước:

thu được sự giả định ban đầu bằng cách sử dụng giàn lọc và ma trận chéo chứa các thừa số chuẩn cho từng băng.

Các phương án của thiết bị cũng có thể được thực hiện bởi các bước của phương pháp và các phương án tương ứng của phương pháp. Do đó, các giải thích đã nêu cho các phương án của thiết bị cũng đúng cho phương pháp.

Mục đích này cũng đạt được bởi hệ thống bao gồm ít nhất một thiết bị mã hóa và thiết bị giải mã. Trong một phương án, nhiều thiết bị mã hóa được sử dụng mà, trong một phương án, được kết hợp với các nút cảm biến, ví dụ, các micrô.

Mục đích của sáng chế cũng đạt được bởi chương trình máy tính để thực hiện phương pháp theo phương án bất kỳ trong số các phương án nêu trên khi chạy trên máy tính hoặc bộ xử lý.

Mô tả vắn tắt các hình vẽ

Sáng chế sẽ được giải thích trong phần sau đây đối với các hình vẽ kèm theo và các phương án được mô tả trong các hình vẽ kèm theo, trong đó:

Fig.1 là sơ đồ khối của hệ thống theo phương án thứ nhất bao gồm thiết bị mã hóa và thiết bị giải mã,

Fig.2 là sơ đồ khối của phương án về thiết bị mã hóa,

Fig.3 là sơ đồ khối của bộ gán trọng số theo cảm quan như một phần của thiết bị mã hóa,

Fig.4 là các tín hiệu thuộc về quy trình mã hóa,

Fig.5 là sơ đồ khối của phương án thứ hai của hệ thống bao gồm nhiều thiết bị mã hóa,

Fig.6 là sơ đồ khối chi tiết hơn của phương án thứ ba của hệ thống,

Fig.7 là sơ đồ khối của một phương án của thiết bị giải mã,

Fig.8 là sơ đồ khối của một phương án khác của thiết bị giải mã,

Fig.9 là phương án của bộ giải trọng số theo cảm quan như một phần của thiết bị giải mã,

Fig.10 là các tín hiệu thuộc về quy trình giải mã,

Fig.11 là các giá trị SNR trung bình với các tốc độ bit khác nhau cho dữ liệu dựa trên thực nghiệm và

Fig.12 là các điểm MUSHRA (Multi Stimulus test with Hidden Reference and Anchor – thử nghiệm đa kích thích với tham chiếu ẩn và neo) vi phân cho dữ liệu dựa trên thực nghiệm.

Mô tả chi tiết sáng chế

Fig.1 là hệ thống xử lý tín hiệu đầu vào 3 mà là tín hiệu tiếng nói và/hoặc âm thanh. Tín hiệu đầu vào 3 này được mã hóa bởi thiết bị mã hóa 1. Tín hiệu được mã hóa 4 tốt hơn là dòng bit, ví dụ, được truyền thông qua internet đến thiết bị giải mã 2 mà giải mã tín hiệu được mã hóa 4 và cung cấp tín hiệu âm thanh được trích xuất 5 đến người nghe - không được thể hiện.

Thiết bị mã hóa 1 sử dụng mô hình được gán trọng số theo cảm quan để xử lý tín hiệu đầu vào 3 nhưng mô hình này không được truyền bởi tín hiệu được mã hóa 4. Thiết bị giải mã 2 trích xuất mô hình từ tín hiệu được mã hóa 4 để hoàn tác các ảnh hưởng của mô hình.

Thiết bị mã hóa 1 được thể hiện trên Fig.2 bao gồm bộ biến đổi 15, bộ gán trọng số theo cảm quan 10, và bộ lượng tử hóa 14.

Bộ biến đổi 15 cung cấp phổ s dựa trên tín hiệu đầu vào 3 là tín hiệu thời gian. Điều này được thực hiện, ví dụ, bởi phép biến đổi Fourier thời gian ngắn (short-time Fourier transform - STFT).

Phổ s trải qua quá trình gán trọng số theo cảm quan bởi bộ gán trọng số theo cảm quan 10 trở thành phổ được gán trọng số theo cảm quan x. Phổ x này được đưa đến bộ lượng tử hóa 14 mà lượng tử hóa nó và cung cấp tín hiệu được lượng tử hóa $\hat{u}$ là dòng bit. Trong phương án này, tín hiệu được lượng tử hóa $\hat{u}$ đó được tổ hợp với thông tin phụ tương ứng bao hàm thông tin về phép lượng tử hóa, nhưng trong phương án này, không bao hàm thông tin về mô hình được gán trọng số theo cảm quan để cung cấp tín hiệu được mã hóa 4.

Fig.3 thể hiện phương án về bộ gán trọng số theo cảm quan 10.

Bộ gán trọng số theo cảm quan 10 nhận tín hiệu đầu vào 3 hoặc phổ tương ứng sau phép biến đổi tín hiệu đầu vào 3 thành miền tần số và cung cấp phổ được gán trọng số theo cảm quan x. Với mục đích này, bộ gán trọng số theo cảm quan 10 bao gồm bộ tính toán đường bao 11, bộ cấp mô hình 12, và bộ áp dụng mô hình 13.

Bộ tính toán đường bao 11 nhận tín hiệu đầu vào 3 hoặc phổ tương ứng và cung cấp đường bao y có phổ cường độ |x| dựa trên tín hiệu đầu vào 3. Trong phương án được thể hiện, bộ tính toán đường bao 11 cung cấp đường bao y bởi ma trận A và ma

trận chéo Λ. Ma trận A là giàn lọc và ma trận chéo Λ chứa các thừa số chuẩn hóa cho mỗi băng của băng lọc được sử dụng. Đường bao y sau đó thu được dựa trên phổ cường độ |x| của tín hiệu đầu vào 3 bởi phương trình $y = A\Lambda A^T|x|$, trong đó A^T là ma trận được chuyển vị của ma trận A.

Dựa trên đường bao y này, bộ cấp mô hình 12 nén đường bao y. Mục đích của việc nén là thu được hàm mà làm gần đúng hiệu suất theo cảm quan của tai. Trong phần sau đây, phương án được mô tả trong đó đường bao được nén bằng cách tăng y tới công suất p. Khi, ví dụ, p bằng 0,3 thì phạm vi của y^p sẽ nhỏ hơn phạm vi của y ban đầu. Do đó, trong ví dụ này, lượng nén phụ thuộc vào giá trị p. Ví dụ, đường bao được giảm tới phạm vi mong muốn. Trong một phương án, hàm giảm phạm vi hoặc hàm nén để nén đường bao y được cho bởi y^p với $0 < p < 1$. Điều này có nghĩa là phép nén được thực hiện bởi hàm mũ của đường bao với số mũ lớn hơn không và nhỏ hơn một. Trong một phương án, phép nén này được thực hiện theo mẫu.

Hàm nén $f(y)$ mô tả véco w, tức là $w = f(y)$ mà cho các phần tử chéo của mô hình được gán trọng số theo cảm quan W. Ở đây, đầu vào y được cho như véctơ và hàm f được áp dụng trên tất cả các mẫu của vectơ y, thu được véctơ w. Do đó, nếu mẫu thứ k' của y là y_k, thì mẫu thứ k' của w = f(y) là $w_k = f(y_k)$.

Do đó, dựa trên hàm nén, có thể thu được mô hình được gán trọng số theo cảm quan W, ở đây, dưới dạng ma trận.

Nói cách khác: đường bao của phổ cường độ được nén và từ hàm mô tả đường bao được nén, mô hình được gán trọng số theo cảm quan được tính toán mà được sử dụng để gán trọng số theo cảm quan cho phổ.

Bộ áp dụng mô hình 13 áp dụng mô hình được gán trọng số theo cảm quan W cho phổ s dựa trên tín hiệu đầu vào 3. Trong phương án được thể hiện, bộ áp dụng mô hình 13 áp dụng ma trận của mô hình được gán trọng số theo cảm quan W cho véctơ dựa trên phổ.

Sự mô hình hóa theo cảm quan sẽ được giải thích một lần nữa:

Các bộ mã hóa-giải mã âm thanh và tiếng nói được dựa trên việc mô hình hóa hiệu quả của cảm quan thính giác của người. Mục đích của sáng chế là để thu được việc gán trọng số của các lỗi lượng tử hóa mà sự tối ưu hóa tỉ số tín hiệu trên nhiễu

trong miền được gán trọng số đem lại chất lượng có khả năng tốt nhất theo cảm quan.

Các bộ mã hóa-giải mã âm thanh thường hoạt động trong miền phổ, tại đó phổ của khung đầu vào s có thể được gán trọng số theo cảm quan với ma trận chéo W sao cho phổ được gán trọng số $x = Ws$ có thể được lượng tử hóa $\hat{x} = \lfloor Ws \rceil$, mà ở đó các dấu ngoặc⌊ ⌉ biểu thị sự lượng tử hóa.

Tại bộ giải mã, phép toán nghịch đảo $\hat{s} = W^{-1}\hat{x}$ có thể được khôi phục.

Đặc biệt là, mô hình gán trọng số theo cảm quan bao gồm hai phần:

i) Phần cố định tương ứng với các giới hạn của cảm quan ở các băng tần số khác nhau. Các mô hình theo cảm quan như các tỉ lệ Bark và ERB mô hình hóa mật độ tần số sao cho các trục bị biến dạng có độ chính xác theo cảm quan đồng đều [17]. Tuy nhiên, vì mục đích là đo năng lượng lỗi trên tỉ lệ bị biến dạng nên cường độ của các thành phần phổ có thể được định tỉ lệ theo cách tương đương sao cho phép toán làm biến dạng phức tạp về mặt tính toán có thể được tránh khỏi [18]. Phép toán này cũng giống với phép toán chỉnh tăng được áp dụng trong các bộ mã hóa-giải mã tiếng nói [1 – 3]. Vì phần gán trọng số này được cố định nên không cần truyền theo cách rõ ràng. Nó có thể được áp dụng ở bộ mã hóa và được nghịch đảo trực tiếp ở bộ giải mã.

ii) Phần thích ứng tín hiệu của mô hình theo cảm quan tương ứng với các đặc tính che tần số của cảm quan. Cụ thể, các thành phần năng lượng cao của tín hiệu sẽ che các thành phần có năng lượng thấp hơn và do đó kết xuất chúng không thể nghe được, nếu hai thành phần đủ gần theo tần số [5]. Hình dạng của đường cong che phổ do đó giống với hình dạng của đường bao tín hiệu, nhưng với cường độ nhỏ hơn.

Trong một phương án, nếu |x| là phổ cường độ của tín hiệu đầu vào thì đường bao phổ y của nó có thể thu được bởi $y = A\Lambda A^T|x|$, trong đó ma trận A là giàn lọc như trên Fig.4.

Trong một phương án, trong chênh lệch của các giàn lọc loại MFCC phổ biến [19], các cửa sổ loại sửa sổ Hann bất đối xứng được sử dụng với phần chồng lấp kéo dài từ bộ lọc thứ k đến bộ lọc (k – 2) và (k + 2) (xem Fig.4).

Ma trận chéo Λ chứa các thừa số chuẩn hóa cho từng băng sao cho thu được độ

khuếch đại đơn vị.

Phụ thuộc vào phương án, tỉ lệ Mel, Bark hoặc ERB được sử dụng với số lượng băng phù hợp.

Ở tốc độ lấy mẫu 12,8kHz, giàn lọc Mel với 20 băng đã được sử dụng.

Cách khác, ma trận giàn lọc loại MFCC là để sử dụng việc trải rộng bằng cách lọc, do đó A trở thành ma trận tích chập. Vì các phép toán lọc đã được hiểu là các phương pháp xử lý tín hiệu số nên nghịch đảo của chúng cũng đã được tìm thấy.

Các thừa số gán trọng số theo cảm quan mô hình hóa hiệu quả che tần số, mà lần lượt tương ứng với sự trải rộng và định tỉ lệ năng lượng trên các tần số [20, 5]. Ma trận mô hình đường bao A đã đạt được các hiệu quả trải rộng, do đó việc định tỉ lệ năng lượng vẫn phải được mô hình hóa.

Việc định tỉ lệ năng lượng tương ứng với phép nén tín hiệu, mà làm giảm phạm vi cường độ của đường bao (xem Fig.4 (b)). Do đó, nếu phổ s được nhân với ma trận gán trọng số theo cảm quan W thì thu được phổ $x = Ws$ mà có phạm vi được giảm (xem Fig.4 (c)).

Do đó việc gán trọng số theo cảm quan làm giảm phạm vi hoặc làm phẳng phổ, nhưng không tạo ra phổ với đường bao hoàn toàn phẳng. Phạm vi của đường bao được giảm, bằng cách đó một phần của phạm vi của nó được giữ lại và phạm vi còn lại có thể được sử dụng để lấy lại tín hiệu gốc sau đường bao được trải rộng.

Hàm nén hoặc giảm phạm vi $w = f(y)$ cho đường bao y (mà véctơ w mang lại các phần tử chéo của W), có thể được áp dụng, ví dụ, như số mũ theo mẫu $f(y) = y^p$ với $0 < p < 1$.

Cho phổ s và mẫu thứ k' của nó là s_k, việc gán trọng số được áp dụng bởi phép nhân sao cho $x_k = w_k * s_k$, trong đó x_k là mẫu thứ k' của phổ được gán trọng số x và w_k là mẫu thứ k' của véctơ gán trọng số w. Phép toán tương tự có thể được trình bày như phép toán ma trận bằng cách tạo ra ma trận W, mà có các giá trị gán trọng số trên đường chéo $W_{kk} = w_k$, và ma trận bằng không ở tất cả các vị trí khác. Nó tuân theo x = W*s.

Trong khi có khả năng sử dụng hàm mà nén phạm vi của y, hàm mũ có lợi ích

là nó dẫn tới biểu thức phân tích đơn giản trong sự khôi phục đường bao ở bộ giải mã.

Các hàm nén phù hợp f() đáp ứng các yêu cầu sau đây:

1. các hàm nén tăng hoàn toàn, tức là f(t) < f(t+eps), trong đó t là giá trị vô hướng dương bất kỳ và eps là giá trị nhỏ tùy chọn,
2. đối với các giá trị vô hướng dương bất kỳ (giá trị vô hướng dương thứ nhất và thứ hai, t1, t2) t1<t2, nó giữ f(t2) - f(t1) <t2– t1. Nói cách khác, hàm bất kỳ mà làm giảm khoảng cách giữa hai giá trị vô hướng dương t2 –t1 này là hàm phù hợp.

Trong phương án khác, ngoài hàm mũ $f(y)=y^p$, với số mũ p nhỏ, hàm nén là hàm loga, tức là, f(y) = log(y).

Trong một phương án, thuật toán mã hóa như phương pháp mã hóa hoặc được thực hiện bởi thiết bị mã hóa là:

1. Tính toán đường bao của phổ cường độ.
2. Nén đường bao để thu được mô hình gán trọng số theo cảm quan.
3. Áp dụng việc gán trọng số lên phổ $x = Ws$.
4. Lượng tử hóa và truyền phổ được gán trọng số $sign(Px)$.

Thuật toán này được áp dụng độc lập ở mọi nút cảm biến.

Việc gán trọng số theo cảm quan được tuân theo bởi phép lượng tử hóa.

Do đó, sáng chế bao gồm hai phần:

1. lượng tử hóa được phân phối của tín hiệu đầu vào sử dụng các phép chiếu ngẫu nhiên và 1 bit lượng tử hóa, và
2. truyền ẩn mô hình theo cảm quan.

Bằng việc lượng tử hóa các phép chiếu ngẫu nhiên, từng bit được truyền mã hóa mảnh thông tin đơn nhất và việc mã hóa quá mức được tránh khỏi.

Mô hình theo cảm quan được tạo ra độc lập ở từng nút cảm biến (ví dụ bao gồm micrô) và tín hiệu được gán trọng số theo cảm quan được lượng tử hóa sẽ được tuyền. Việc gán trọng số theo cảm quan làm cho tín hiệu phẳng hơn, nhưng giữ lại được hình dạng cơ bản. Do đó, có thể kết luận theo cách nghịch đảo ở phía bộ giải mã cái gì mà đường bao gốc phải có, ngay cả từ tín hiệu được gán trọng số theo cảm quan.

Điều này tuân theo sự mô tả về phép lượng tử hóa được phân phối:

Dù việc mã hóa nguồn được phân phối là đối tượng được nghiên cứu kỹ (ví dụ [7, 8]) và đã được áp dụng trong những ứng dụng khác như video [9], chỉ một số ít đã làm việc trên mã hóa âm thanh được phân phối (ví dụ [10 – 13]), và, tuy nhiên, không có ứng dụng nào trong số đó giải quyết vấn đề mã hóa quá mức đối với các mô hình cảm quan và đường bao. Ngay cả phương pháp mã hóa có thể thay đổi tỉ lệ trong [14] bao gồm mã hóa đường bao với các thừa số tỉ lệ. Ngoài ra, nhiều phương pháp mã hóa đã mô tả đã được áp dụng chỉ cho sự che giấu tổn hao gói [15, 16].

Nó tuân theo sơ đồ lượng tử hóa có thể thực hiện dễ dàng. So sánh phương pháp lượng tử hóa 1 bit mà đã được sử dụng trong các hệ thống cảm biến nén [6].

Mục đích của bộ lượng tử hóa và quy trình lượng tử hóa là để cho phép sự lượng tử hóa ở các cảm biến độc lập, sao cho được đảm bảo rằng từng bit được truyền cải thiện chất lượng, mà không có sự truyền thông giữa các nút cảm biến. Ở cực biên, cảm biến có thể chỉ gửi một bit và bit đơn có thể được sử dụng để cải thiện chất lượng.

Sơ đồ lượng tử hóa được đề xuất của một phương án được dựa trên các phép chiếu ngẫu nhiên của sự biểu diễn giá trị thực của phổ tín hiệu và truyền dấu của từng kích thước.

Để x là véctơ N nhân 1 có giá trị thực chứa phổ của tín hiệu đầu vào, và P là ma trận ngẫu nhiên K nhân N mà các cột của ma trận được chuẩn hóa tới độ dài đơn vị. Sau đó, x sẽ được biến đổi bởi $u = Px$. Điều này được tuân theo bởi phép lượng tử hóa của dấu của từng thành phần u, tức là, phép lượng tử hóa là $\hat{u} = sign(u)$, mà có thể được truyền theo cách không tổn hao với K bit.

Do đó, số lượng bit của dòng bit xác định một chiều của ma trận ngẫu nhiên.

Các mẫu P tốt hơn là các giá trị giả ngẫu nhiên mà có nghĩa là chúng giống các giá trị ngẫu nhiên, nhưng được tạo ra thực sự bởi một số thuật toán hoặc công thức toán học. Các bộ tạo giả ngẫu nhiên là các công cụ toán học được chuẩn hóa mà mọi máy tính và mọi thư viện phần mềm toán học có. Thích hợp là: ma trận P phải được biết ở cả bộ mã hóa và bộ nhận/bộ giải mã, và sự phân phối các số lượng ngẫu nhiên là giống nhau cho tất cả các mẫu trong P.

Sự khôi phục như phép làm gần đúng của x có thể đã được tính toán bởi

$$\hat{x} = P^{\dagger}\hat{u} = P^{\dagger}sign(Px) \quad (1)$$

trong đó $P^{\dagger}$ là nghịch đảo giả của ma trận ngẫu nhiên P.

Miễn là hạt giống cho các cột giả ngẫu nhiên của P được biết tại bộ giải mã, do đó bộ giải mã có thể chỉ giải mã tín hiệu từ $\hat{u}$. Do đó, trong một phương án, hạt giống của cột giả ngẫu nhiên được cho như thông tin phụ của tín hiệu được mã hóa. Bộ tạo giả ngẫu nhiên thường tạo ra các chuỗi giá trị ngẫu nhiên mà cho giá trị trước x(k) theo chuỗi, nó tạo ra các mẫu ngẫu nhiên tiếp theo, x(k+1) = f(x(k)). Tức là, nếu điểm bắt đầu x(1) – được gọi là giá trị "hạt giống" của chuỗi giả ngẫu nhiên – được biết đến, thì có thể tạo ra chuỗi toàn bộ. Do đó, ở phía mã hóa và giải mã, hàm tương tự để tạo ra các mẫu ngẫu nhiên được sử dụng.

Trong trường hợp có nhiều nút cảm biến, tín hiệu đầu vào x được giả định để là các phiên bản tương tự hoặc nhiễu của tín hiệu tương tự, nhưng từng cảm biến có ma trận ngẫu nhiên P_k của nó. Ở bộ giải mã, các ma trận ngẫu nhiên có thể được đối chiếu tới ma trận lớn đơn lẻ P = [P_1, P_2, …] nhờ bởi Eq. 1 giữ nguyên không đổi.

Được biết rằng nếu K <<N, thì P gần như là trực tâm, $P^T P \approx I$ và phép lượng tử hóa gần như tối ưu.

Ở đây, K không nhất thiết nhỏ hơn N, do đó trực tâm trở nên ít chính xác hơn. Sử dụng phép chuyển vị thay vì phép giả nghịch đảo làm giảm độ phức tạp thuật toán và hiệu quả mã hóa, nhưng không không áp đặt giới hạn cho các thí nghiệm của chúng ta về mô hình hóa theo cảm quan, vì mọi bit được truyền vẫn cải thiện độ chính xác của tín hiệu đầu ra.

Có thể được mong đợi rằng mô hình nguồn sau đó sẽ được áp dụng lên phía bộ giải mã và mô hình này sẽ làm tăng độ chính xác của sự khôi phục. Tuy nhiên không cần thiết thực hiện mô hình nguồn, vì hiệu quả của nó có thể được mô phỏng bằng cách tăng độ chính xác bằng cách gửi nhiều bit hơn.

Sơ đồ dòng của phương án về hệ thống (ngoại trừ mô hình theo cảm quan) được minh họa trên Fig.5. Được thể hiện là n micrô thu các tín hiệu âm thanh từ một nguồn âm thanh. Trong phương án được thể hiện, các thiết bị mã hóa 1 sau đây là một phần

của micrô hoặc nút cảm biến tương ứng.

n tín hiệu đầu vào 3 được biến đổi bởi các bộ biến đổi 15 của n thiết bị mã hóa 1 thành n phổ s mà được chuyển đổi thành n phổ được gán trọng số theo cảm quan $x_k = W_k s_k$ mà được thực hiện ở đây bởi các bộ gán trọng số theo cảm quan không được thể hiện.

n phổ được gán trọng số theo cảm quan x_k được đưa tới n bộ lượng tử hóa 14.

Từng bộ lượng tử hóa 14 bao gồm bộ áp dụng ma trận ngẫu nhiên 16 và bộ tính toán hàm dấu 17.

Bộ áp dụng ma trận ngẫu nhiên 16 áp dụng ma trận ngẫu nhiên P_i là ma trận với kích thước K nhân N cho véctơ N nhân 1 dựa trên phổ được gán trọng số theo cảm quan tương ứng x_i với i = 1, 2, …, n. N là vô hướng nguyên tương ứng với số lượng các mẫu trong phổ x. K là vô hướng nguyên tương ứng với số lượng hàng trong phổ được ngẫu nhiên hóa. Bằng phép lượng tử hóa với toán tử dấu, từng hàng được lượng tử hóa với một bit, do đó số lượng bit được truyền là K. Khi nhiều bộ mã hóa, ví dụ các micrô, được sử dụng, thì từng bộ mã hóa có ma trận P_k của nó, mà có kích thước K_k x N. Tức là, từng bộ mã hóa gửi K_k bit tới bộ giải mã, trong đó số lượng các bit có thể thay đổi từ bộ mã hóa đến bộ mã hóa.

Phổ được biến đổi được cho bởi: $u_i = P_i x_i$.

Bộ tính toán hàm dấu 17 tính toán dấu hoặc hàm dấu của phổ biến đổi tương ứng: $\hat{u}_i = sign(u_i)$. Điều này được thực hiện cho K thành phần của phổ được biến đổi dẫn đến K bit được truyền thông qua kênh truyền.

Fig.6 thể hiện một thiết bị mã hóa 1 bao gồm bộ gán trọng số theo cảm quan 10 và bộ lượng tử hóa 14.

Tín hiệu âm thanh 3 được truyền bởi bộ biến đổi 15 thành phổ s_k. Bộ gán trọng số theo cảm quan 10 áp dụng ma trận gán trọng số theo cảm quan W cho phổ s_k để cung cấp phổ được gán trọng số theo cảm quan x_k mà được lượng tử hóa bởi bộ lượng tử hóa 14.

Bộ lượng tử hóa 14 bao gồm bộ áp dụng ma trận ngẫu nhiên 16 mà nhận phổ được gán trọng số theo cảm quan x_k và áp dụng ma trận ngẫu nhiên P_k cho nó thông

qua P_kx_k. Các thành phần của phổ biến đổi thu được u_k được lượng tử hóa bởi bộ tính toán hàm dấu 17 thông qua việc tính toán hàm dấu của từng thành phần. Điều này dẫn đến các dòng bit với số lượng các bit bằng số lượng các thành phần của phổ biến đổi. Do đó, tín hiệu được mã hóa 4 được cho bởi dòng bit $\hat{u}_k$.

Nó tuân theo mô tả về cách giải mã tín hiệu được mã hóa.

Fig.7 thể hiện phương án của thiết bị giải mã 2 để xử lý tín hiệu được mã hóa 4.

Tín hiệu được mã hóa 4 dưới dạng dòng bit $\hat{u}$ được giải lượng tử hóa bởi bộ giải lượng tử hóa 20.

Tín hiệu được mã hóa 4 được giải lượng tử hóa bằng cách sử dụng phép giả nghịch đảo $P^{\dagger}$ của ma trận ngẫu nhiên P. Thông tin về ma trận ngẫu nhiên thu được từ thông tin phụ của tín hiệu được mã hóa 4. Phổ được gán trọng số theo cảm quan đã tính toán được cho bởi: $\hat{x} = P^{\dagger}\hat{u}$.

Phổ được gán trọng số theo cảm quan được tính toán $\hat{x}$ được giải trọng số bởi bộ giải trọng số theo cảm quan 21. Phổ được tính toán $\hat{s}$ được cung cấp bởi bộ giải trọng số theo cảm quan 21 được biến đổi bởi bộ giải biến đổi 22 (ví dụ thông qua biến đổi Fourier thời gian ngắn nghịch đảo, $STFT^{-1}$) thành tín hiệu thời gian mà là tín hiệu âm thanh được trích xuất 5.

Fig.8 thể hiện phương án trong đó bộ giải lượng tử hóa 20 nhận nhiều tín hiệu được mã hóa 4 từ các nút cảm biến khác nhau, tức là từ các micrô khác nhau. Các ma trận ngẫu nhiên riêng lẻ P_k được sắp xếp thành một ma trận lớn $P = [P_1, P_2, \ldots]$.

Phổ được gán trọng số theo cảm quan được tính toán đơn lẻ sau đó được cho bởi: $\hat{x} = \begin{bmatrix} P_1 \\ P_2 \\ \ldots \\ P_n \end{bmatrix}^{\dagger} \begin{bmatrix} u_1 \\ u_2 \\ \ldots \\ u_n \end{bmatrix}$.

Trong phương án thay thế, các dòng bit được nghịch đảo với ma trận ngẫu nhiên tương ứng của chúng $\hat{x}_k = P_k^{\dagger}u_k$ và các phổ được hợp nhất sau đó.

Phổ được gán trọng số theo cảm quan được tính toán đã thu được $\hat{x}$ sẽ được xử lý như được mô tả cho phương án được thể hiện trên Fig.7.

Trên Fig.9 là phương án của bộ giải trọng số theo cảm quan 21 như một phần

của thiết bị giải mã được thể hiện.

Bộ giải trọng số theo cảm quan 21 nhận, từ bộ giải lượng tử hóa 20, phổ được gán trọng số theo cảm quan được tính toán $\hat{x}$ mà là dòng bit được giải lượng tử hóa của tín hiệu được mã hóa 4.

Bộ giải trọng số theo cảm quan 21 phục hồi mô hình được gán trọng số theo cảm quan W được sử dụng trong suốt quá trình mã hóa của tín hiệu đầu vào 3 bằng cách sử dụng thuật toán lặp. Ở đây, W_k là phép làm gần đúng hoặc sự giả định thứ k của mô hình được gán trọng số theo cảm quan W.

Thuật toán bắt đầu với bộ cấp giả định ban đầu 23. Bộ cấp giả định ban đầu 23 cung cấp phép ước lượng thứ không W_0 của mô hình được gán trọng số theo cảm quan W.

Trong phương án được thể hiện, véctơ w_k được sử dụng mà bao gồm các phần tử chéo của ma trận tương ứng W_k. Véctơ này được sử dụng với hàm nén đường bao của phổ cường độ trong suốt quá trình gán trọng số theo cảm quan của tín hiệu đầu vào.

Do đó, sự giả định thứ 0 của véctơ w_0 được cho và phép làm gần đúng thứ 0 W_0 của mô hình được gán trọng số theo cảm quan W được thiết lập theo cách thích hợp.

Trong một phương án, sự giả định ban đầu của véctơ gần đúng w_0 được thiết lập bằng cách sử dụng phương trình sử dụng cho sự tính toán đường bao y của phổ cường độ.

Do đó, sự giả định ban đầu trở thành: $w_0 = A\Lambda A^T \hat{x}$ với ma trận A và ma trận chéo Λ.

Ma trận A là giàn lọc và ma trận chéo Λ chứa các thừa số chuẩn hóa cho từng băng của băng lọc được sử dụng. Lựa chọn của ma trận A là lựa chọn thiết kế, mà thường sẽ được cố định trước. Trong một phương án, trong đó ma trận A không thay đổi theo thời gian, nó được nối cứng vào cả bộ mã hóa và bộ giải mã. Điều này ngụ ý rằng bộ mã hóa và bộ giải mã sử dụng cùng ma trận A đã biết. Trong một phương án, ma trận A được sử dụng cho phép mã hóa được truyền như phần của tín hiệu được mã hóa. Trong một phương án, lựa chọn về ma trận A được truyền trong bước khởi tạo

hoặc bước tạo cấu hình, trước khi sự truyền thông bắt đầu thực sự. Phần đã nói ở trên cũng đề cập đến ma trận chéo Λ.

Ma trận giả định ban đầu W_0 được chuyển tới bộ làm gần đúng phổ 24. Điều này cũng bắt đầu phép lặp thực sự với chỉ số chạy k được thiết lập bằng 0.

Bộ làm gần đúng phổ 24 tính toán dựa trên dữ liệu w_0 cho sự giả định ban đầu W_0 của mô hình được gán trọng số theo cảm quan W phép làm gần đúng của phổ.

Điều này được thực hiện thông qua việc áp dụng phép nghịch đảo của ma trận giả định ban đầu W_0 cho véctơ chứa phổ được gán trọng số theo cảm quan đã tính toán $\hat{x}$.

Phổ được làm gần đúng trong bước này với k = 0 bởi $\hat{s}_0 = W_0^{-1}\hat{x}$.

Phổ được làm gần đúng $\hat{s}_0$ được cung cấp với bộ làm gần đúng mô hình 25 mà tính toán phép làm gần đúng mới W_1 cho mô hình được gán trọng số theo cảm quan W.

Điều này dựa trên báo nhận rằng vectơ của phần tử chéo w_k là hàm của phổ $\hat{s}_k$.

Trong một phương án, hàm nén f(), mà được sử dụng trong suốt quá trình mã hóa để tạo ra ma trận W, được biết đến ở phía giải mã. Giống như đã thảo luận trước liên quan đến ma trận A, hàm nén f() hoặc là được thiết lập ở cả hai phía, ví dụ, được cố định trong phần mềm tương ứng để mã hóa và giải mã hoặc được truyền như phần của tín hiệu được mã hóa.

Do đó, phép làm gần đúng của véctơ được tính toán thông qua: $w_1 = f(\hat{s}_0)$ và thu được phép làm gần đúng tiếp theo, ở đây là phép làm gần đúng thứ nhất W_1.

Phép làm gần đúng W_1 này được quay lại bộ làm gần đúng phổ 24 để tính toán phổ được làm gần đúng tiếp theo $\hat{s}_k = W_k^{-1}\hat{x}$. Chỉ số k được tăng lên theo.

Phép ước lượng này đóng vai trò sau đó để thu được phép ước lượng tiếp theo W_{k+1} của mô hình được gán trọng số theo phổ W dựa trên véctơ với phần tử chéo được cho bởi công thức:$w_{k+1} = f(\hat{s}_k)$.

Việc này được lặp lại đến khi được hội tụ. Phép hội tụ thường được kiểm tra bằng cách so sánh vectơ đầu ra hiện thời w(k) với véctơ trước w(k-1). Nếu chênh lệch ||w(k) – w(k-1)|| nằm dưới một số ngưỡng được thiết lập trước, thì phép lặp đã được

hội tụ và các phép lặp có thể được dừng.

Phổ được làm gần đúng thu được cuối cùng $\hat{s}$ là đầu ra của bộ giải trọng số theo cảm quan 21 và là phép làm gần đúng của phổ của tín hiệu đầu vào gốc.

Sự khôi phục của mô hình theo cảm quan được giải thích một lần nữa:

Ở phía bộ giải mã, phép ước lượng $\hat{x}$ của tín hiệu theo cảm quan x (xem Eq. 1) có thể được phục hồi $\hat{x} = P^{\dagger} sign(Px)$ trong đó $\hat{u} = sign(Px)$ là phổ được gán trọng số theo cảm quan được lượng tử hóa x. Việc này được thực hiện bởi bộ giải lượng tử hóa 20.

Do đó nhiệm vụ chính là khôi phục phép ước lượng $\hat{s}$ của tín hiệu s từ tín hiệu theo cảm quan được lượng tử hóa $\hat{x}$.

Phổ được gán trọng số theo cảm quan x được dựa trên phổ của tín hiệu đầu vào thông qua mô hình được gán trọng số theo cảm quan W bởi phương trình: $x = Ws$.

Mục đích là phép ước lượng bằng với phổ, tức là $\hat{x} \approx x$.

Do đó, $Ws \approx W\hat{s}$.

Hơn nữa, véctơ w cho các phần tử chéo của ma trận W là hàm của phổ s của tín hiệu đầu vào: w = f(s).

Do đó, với phép ước lượng w, $\hat{s}$ có thể được ước lượng, do đó w có thể được ước lượng. Việc này có thể được lặp lại đến khi hội tụ.

Do đó đây là thuật toán loại cực đại hóa kỳ vọng mà có thể được mô tả như:

1. Có được sự giả định ban đầu w_0 bởi, ví dụ, $w_0 = A\Lambda A^T \hat{x}$ và thiết lập W_0 theo cách thích hợp.

2. Lặp lại từ k = 0 đến khi được hội tụ:

   (a) Tính toán $\hat{s}_k = W_k^{-1} \hat{x}$.

   (b) Tính toán $w_{k+1} = f(\hat{s}_k)$ và thiết lập W_{k+1} theo cách thích hợp.

   (c) Tăng k.

   Các giá trị cuối cùng $\hat{s}_k$ và W_k là các phép ước lượng cuối cùng của $\hat{s}$ và $\widehat{W}$.

   Thông thường, ít hơn 20 phép lặp được yêu cầu để hội tụ.

Các phương án khác nhau của thiết bị giải mã 2 cũng được thể hiện bởi Fig.5 và Fig.6. Thiết bị giải mã 2 của Fig.5 được thể hiện mà không có phép giải trọng số. Trên Fig.6, phép làm gần đúng của mô hình được gán trọng số theo cảm quan W được cho và được sử dụng cho phép giải trọng số của phổ được giải lượng tử hóa $\hat{x}$. Điều này làm rõ rằng việc giải mã bao gồm sự khôi phục mô hình W dựa trên tín hiệu được mã hóa.

Để đánh giá việc thực hiện của từng phần của hệ thống được đề xuất, các thí nghiệm sau được thực hiện.

Ba phiên bản của âm thanh đầu vào được so sánh:

tín hiệu được lượng tử hóa và được khôi phục, 1) không có và 2) có phép mô hình hóa theo cảm quan sao cho cảm quan được biết ở bộ giải mã,

cũng như 3) tín hiệu được lượng tử hóa theo cảm quan mà sự khôi phục được thực hiện với mô hình cảm quan được ước lượng mù theo sáng chế.

Như vật liệu thử nghiệm, các mẫu tiếng nói ngẫu nhiên từ tập hợp dữ liệu NTT-AT [21] đã được sử dụng (xem Fig.10 hàng phía trên). Các tín hiệu đầu vào được lấy mẫu lại đến 12,8kHz, STFT được thực hiện với biến đổi cosin rời rạc để thu được mô hình phổ và đường bao có giá trị thực được sử dụng với 20 băng, được phân phối theo tỉ lệ Mel [20, 5].

Như phép làm gần đúng của mô hình theo cảm quan, hàm giảm phạm vi của $f(y) = y^p$ với p = 0,5 đã được sử dụng. Mô hình theo cảm quan này chỉ được chọn như cách để giải thích việc thực hiện của sự khôi phục mù và cần được xem xét như sản phẩm cuối cùng được điều chỉnh. Việc thực hiện của mô hình đường bao cũng như mô hình theo cảm quan đã được minh họa trên Fig.4.

Trước hết, SNR theo cảm quan cho phép lượng tử hóa được đề xuất mà không có (SNR_O) và với sự khôi phục mù (SNR_B) của mô hình cảm quan, theo cách tương ứng, sẽ được ước lượng bởi:

$$SNR_O = \frac{\|x\|}{\|x-\hat{x}\|}, \text{ và } SNR_B = \frac{\|x\|}{\|x-W\hat{W}^{-1}\hat{x}\|} \quad (2)$$

Fig.10 là SNR theo cảm quan cho tập tin tiếng nói được lượng tử hóa với các phương pháp khác nhau (K = 3000).

Rõ ràng là khi mô hình theo cảm quan được biết đến (phương thức oracle), SNR gần với 8,4dB. Sự khôi phục mù của mô hình cảm quan làm giảm rõ ràng chất lượng (mù) đặc biệt là cho các âm vị có tiếng. Tuy nhiên, SNR của hệ thống mà không phải mô hình cảm quan (No perc) kém hơn hai lần khôi phục mù.

Để định lượng thêm lợi ích của sự khôi phục mù thay vì mô hình hóa không cảm quan, SNR trung bình với các tốc độ bit khác nhau K được đo (xem Fig.11).

Phương pháp phục hồi mù và mô hình hóa không cảm quan có trung bình 1,1 dB và kém hơn 5,8 dB so với phương pháp oracle. Rõ ràng, SNR cải thiện với tốc độ bit, dù trường hợp mô hình không cảm quan cải thiện ít hơn với mô hình cảm quan. Hơn nữa, trong việc tăng SNR, phép phục hồi mù tiếp cận chất lượng của phương pháp oracle theo cách tiệm cận.

Cuối cùng, để đánh giá chất lượng đối tượng, thử nghiệm nghe MUSHRA được thực hiện với tám người nghe và sáu mục chọn ngẫu nhiên từ tập hợp dữ liệu NTT-AT. Tín hiệu được lượng tử hóa với 3kbit/khung. Tức là số lượng tương đối thấp được cho mà không có mô hình hóa nguồn được sử dụng cho các thí nghiệm, bằng cách đó SNR đầu ra cũng tương đối thấp. Ngữ cảnh này được chọn để chứng minh các điều kiện có vấn đề và hiệu suất được mong đợi để cải thiện đáng kể ở tốc độ bit cao cũng như khi áp dụng mô hình nguồn.

Từ các điểm MUSHRA khác nhau trên Fig.12, có thể thấy rằng đối với tất cả các mục, mô hình hóa cảm quan cải thiện chất lượng với cả ước lượng oracle và ước lượng mù, bởi lần lượt là 29,9 và 22,3 điểm trung bình. Ý nghĩa thống kê của các chênh lệch đã được xác nhận với Student's t-test ở $p > 99\%$.

Phép lượng tử hóa 1 bit và sơ đồ mã hóa được đề xuất hoặc là được tổ hợp trong một phương án hoặc trong các phương án khác nhau có một số kết quả và đặc tính đáng chú ý.

Thứ nhất, để phân tích các đặc tính lượng tử hóa, lưu ý là từng cột của P là phép chiếu lên không gian con một chiều của không gian N chiều của véctơ x.

Bằng cách mã hóa dấu của một phép chiếu, không gian N chiều được chia thành hai phần. Bằng cách mã hóa lặp lại các dấu của Px, không gian N chiều được chia thành các ô lượng tử hóa nhỏ hơn bao giờ hết. Vì P là ma trận ngẫu nhiên, các cột của

nó là gần như trực giao với nhau, do đó các ô lượng tử hóa vẫn gần tối ưu.

Trong hệ thống nút đơn, phương pháp lượng tử hóa có thể được thiết kế mà hiệu quả hơn. Tuy nhiên, trong hệ thống được phân phối, nó trở nên phức tạp hơn - phương pháp đơn giản là cần thiết để ngăn ngừa các nút khỏi mã hóa cùng một thông tin, nghĩa là, mã hóa quá mức phải được tránh trong khi vẫn giữ được độ phức tạp thuật toán thấp. Phép lượng tử hóa nghịch đảo là rất đơn giản và cung cấp hiệu suất gần tối ưu.

Thứ hai, các phương pháp mã hóa nguồn không được sử dụng.

Tuy nhiên, được biết rằng việc mô hình hóa này có thể được sử dụng để cải thiện đáng kể hiệu quả mã hóa. Mô hình hóa nguồn có thể được áp dụng ở phía bộ giải mã bằng cách mô hình hóa sự phân phối xác suất của các tín hiệu âm thanh và tiếng nói (ví dụ [22]). Mô hình hóa nguồn có thể, vì tín hiệu được lượng tử hóa có thể được xử lý như sự theo dõi nhiễu của tín hiệu "đúng", do đó, bằng cách áp dụng sự phân phối của nguồn, tối ưu hóa có khả năng cực đại (hoặc tương tự) có thể được áp dụng để làm gần đúng tín hiệu "đúng". Vì sự tối ưu hóa này được áp dụng trong mạng hoặc ở bộ giải mã, tải tính toán được giữ khỏi nút cảm biến và các nút cảm biến có thể giữ nguyên công suất thấp.

Thứ ba, từ quan điểm riêng tư, phương pháp chiếu ngẫu nhiên có thể được thiết kế để trở thành phép mật mã hóa hiệu quả cao.

Bộ nghe lén ifan không biết được hạt giống cho ma trận ngẫu nhiên, thì dữ liệu sẽ có vẻ hoàn toàn ngẫu nhiên và vô nghĩa. Giả định rằng hạt giống ngẫu nhiên được truyền thông theo cách an toàn, thì chỉ bộ mã hóa và bộ nhận được dự định có thể giải mật mã tin nhắn. Phương pháp này trái ngược với các phương pháp như được cho bởi [12, 13], trong đó sự truyền thông giữa các nút được cố ý sử dụng. Trong khi sự tương quan giữa các nút có thể được sử dụng để cải thiện SNR theo cảm quan, sự riêng tư sẽ khó bảo đảm hơn. Ngay cả khi giả định rằng các nút đang hoạt động trên mạng an toàn, có thể chỉ lấy một nút bị làm lộ để có được sự truy cập vào tất cả các giao tiếp. Ngược lại, trong phương pháp được đề xuất, nếu bộ nghe lén có được sự truy cập vào một nút cảm biến, nó làm lộ dữ liệu của chỉ nút đó, vì các nút có thể và nên sử dụng các hạt giống khác nhau. Để giới hạn công suất truyền của các nút cảm biến, tuy nhiên

có thể được cho phép là các nút phát lại các gói, vì các gói giữ nguyên khả năng đọc chỉ bởi sự nhận được dự định và do đó sự riêng tư không bị để lộ.

Theo một phương án, phương pháp theo sáng chế dựa trên ý tưởng lượng tử hóa 1 bit, mà ở trên phía bộ mã hóa, tín hiệu đầu vào được gán trọng số theo cảm quan được chiếu tới các khoảng trống nhỏ ngẫu nhiên, và trong đó các dấu của từng kích thước sau đó được truyền. Bộ giải mã có thể nghịch đảo phép lượng tử hóa với phép giả nghịch đảo, hoặc tương tự, để thu được tín hiệu được gán trọng số theo cảm quan được lượng tử hóa.

Phần chính của phương pháp được đề xuất sau đó là sự khôi phục của phép ước lượng của tín hiệu gốc, khi chúng ta chỉ có sự truy cập vào tín hiệu được gán trọng số theo cảm quan. Phương pháp được dựa trên thuật toán tối đa hóa ước lượng (estimation-maximization - EM), trong đó nó được luân phiên lặp lại giữa việc ước lượng mô hình cảm quan và tín hiệu gốc.

Do đó thuật toán mã hóa âm thanh và tiếng nói đã phân phối được đề xuất là phương pháp có thể thực hiện cho các ứng dụng cho internet vạn vật. Nó cung cấp hiệu suất có thể thay đổi cho số lượng nút cảm biến bất kỳ và mức tiêu thụ công suất. Hơn nữa, thuật toán được đảm bảo bởi thiết kế, vì sự riêng tư của kênh truyền thông có thể được đảm bảo bởi sự truyền thông được mật mã hóa của hạt giống ngẫu nhiên.

Sáng chế bao gồm ít nhất là các dấu hiệu, khía cạnh, cũng như các ứng dụng và mở rộng rõ ràng sau đây. Do đó, việc liệt kê đề cập đến phương án khác và cho phép các tổ hợp khác nhau của các dấu hiệu được đề cập.

1. Hệ thống mã hóa âm thanh và tiếng nói được phân phối với một hoặc một số bộ mã hóa, mà có thể được thực hiện trên các bộ xử lý có thể thay đổi, linh hoạt, công suất thấp và chi phí thấp với khả năng truyền và khả năng tính toán giới hạn và các tài nguyên khác.

1.1. Các bộ mã hóa được phân phối có thể được tạo cấu hình để mã hóa thông tin sao cho thông tin từ mỗi nút bộ mã hóa có thể được lượng tử hóa độc lập sao cho lượng thông tin từ việc tập hợp các nút có thể được tối đa hóa, ví dụ,

1.1.1. bằng cách che thông tin từng nút bộ mã hóa mà trực giao hoặc gần như trực giao với các nút khác,

1.1.2. bằng cách sử dụng các thuật toán lượng tử hóa được ngẫu nhiên hóa, mà có thể hoặc không thể trực giao hoặc gần như trực giao với các nút khác,

1.1.3. bằng cách sử dụng các phương pháp mã hóa entropy để nén thông tin được lượng tử hóa.

1.2. Các bộ mã hóa được phân phối có thể được tạo cấu hình để mã hóa thông tin trong miền cảm quan, mà làm gần đúng độ cảm quan theo cảm quan của hệ thống nghe của người

1.2.1. trong đó việc biến đổi sang miền cảm quan có thể được thiết kế để có thể nghịch đảo, sao cho các nút bộ mã hóa có thể phục hồi (mù) mô hình cảm quan mà không truyền rõ ràng mô hình theo cảm quan,

1.2.2. trong đó một số hoặc tất cả các tham số của các mô hình cảm quan có thể được truyền rõ ràng từ một số hoặc tất cả các nút bộ mã hóa như thông tin phụ.

1.2.3. trong đó các mô hình cảm quan có thể được mô tả bằng cách sử dụng các mô hình đường bao như mô hình dự báo tuyến tính, các giàn lọc, các biến đổi ma trận hoặc các băng thừa số tỉ lệ (hằng số theo mảnh hoặc được nội suy) và chúng có thể được thực hiện như các phép toán ma trận hoặc các phép toán lọc.

1.2.4. trong đó các mô hình cảm quan có thể có các phần không đổi và thích ứng, trong đó các phần không đổi luôn giống nhau, và các phần thích ứng phụ thuộc vào tín hiệu đầu vào và hoặc cấu hình của hệ thống và/hoặc lượng tài nguyên (phần cứng, sự tính toán, truyền, dung lượng pin, v.v.) khả dụng.

1.3. Các bộ mã hóa được phân phối có thể được tạo cấu hình để mã hóa thông tin sao cho lỗi lượng tử hóa được giảm thiểu trong miền cảm quan, ví dụ, giống như chu trình phân tích bằng tổng hợp trong các bộ mã hóa-giải mã loại CELP.

1.4. Các bộ mã hóa được phân phối có thể được tạo cấu hình để mã hóa thông tin sử dụng phép mật mã để ngăn ngừa sự nghe lén bởi, ví dụ,

1.4.1. mật mã hóa sự mô tả tín hiệu cuối cùng hoặc trung gian bằng cách sử dụng các phương pháp mà là thuật toán hoặc dựa trên các bảng tra cứu,

1.4.2. bằng cách sử dụng phép lượng tử hóa với phép ngẫu nhiên hóa, như các phép biến đổi hoặc chiếu ngẫu nhiên, trong đó phép ngẫu nhiên hóa đạt được bởi thuật toán hoặc bảng tra cứu với các giá trị hạt giống được biết (điểm bắt đầu).

1.5. Các bộ mã hóa được phân phối mà cấu hình của nó có thể là linh hoạt và được biến đổi trực tuyến, ví dụ,

1.5.1. dựa trên vị trí và sự di chuyển của các nguồn âm thanh (như loa phóng thanh có yếu tố con người), các nút cảm biến, cũng như tính sẵn có của các tài nguyên, các nút cảm biến có thể quyết định độc lập hoặc kết hợp mà các nút là hoạt động hoặc không hoạt động,

1.5.2. dựa trên vị trí và sự di chuyển của các nguồn âm thanh (như loa phóng thanh có yếu tố con người), các nút cảm biến cũng như tính khả dụng của các tài nguyên, các nút cảm biến có thể quyết định điều chỉnh độc lập hoặc phụ thuộc sự phân phối tài nguyên, ví dụ, sao cho nút cảm biến với micro gần loa phóng thanh sử dụng nhiều tài nguyên hơn các micro ở xa hơn.

1.6. Trường hợp đặc biệt của hệ thống được đề xuất là khi chỉ có bộ mã hóa đơn, do đó hệ thống giảm để tuân theo cấu trúc cổ điển của các bộ mã hóa-giải mã âm thanh và tiếng nói. Tuy nhiên, quan trọng là phép lượng tử hóa ngẫu nhiên và phục hồi mù của các khía cạnh mô hình cảm quan của sáng chế này cũng có thể được sử dụng trong các thiết kế bộ mã hóa và bộ giải mã cổ điển, ví dụ, đối với lợi ích về hiệu quả mã hóa, để thu được bộ mã hóa ít phức tạp, và để mật mã hóa truyền thông.

2. Hệ thống mã hóa âm thanh và tiếng nói được phân phối mà đầu vào từ bộ mã hóa được hợp nhất ở một hoặc một số nút mạng (sử dụng các phương pháp xử lý trong mạng) hoặc ở một hoặc một số nút bộ giải mã/bộ nhận.

2.1. Bộ giải mã hoặc các bộ phận xử lý có thể được tạo cấu hình để nghịch đảo phép lượng tử hóa bởi, ví dụ,

2.1.1. biến đổi nghịch đảo chính xác, giả nghịch đảo hoặc nghịch đảo gần đúng như đổi chỗ ma trận ngẫu nhiên.

2.1.2. ước lượng tín hiệu đầu vào bằng phương pháp tối ưu hóa, như cảm biến nén,

2.1.3. các dòng bit u_ktừ các nút $k \in [1, 2, \dots, n]$ có thể được hợp nhất bởi phép nghịch đảo kết hợp như $\hat{x} = \begin{bmatrix} P_1 \\ P_2 \\ \dots \\ P_n \end{bmatrix}^{\dagger} \begin{bmatrix} u_1 \\ u_2 \\ \dots \\ u_n \end{bmatrix}$ hoặc các dòng bit trước hết có thể được nghịch đảo $\hat{x}_k = P_k^{\dagger} u_k$ và chỉ được hợp nhất sau đó. Lợi ích của cái sau là ở chỗ sau đó có thể áp dụng việc mô hình hóa, như các mô hình cảm quan, nguồn không gian và cảm biến theo cách riêng rẽ hoặc kết hợp lên từng nút.

2.2. Bộ giải mã hoặc bộ phận xử lý có thể được tạo cấu hình để nghịch đảo mô hình cảm quan bằng cách sử dụng thông tin phụ được truyền và/hoặc bằng cách ước lượng mù nó từ tín hiệu được khôi phục (cảm quan) bởi, ví dụ,

2.2.1. các phương pháp trực tiếp, trong đó mô hình cảm quan thu được trực tiếp từ tín hiệu được lượng tử hóa,

2.2.2. các phương pháp lặp, trong đó phép ước lượng của mô hình cảm quan là phép lặp được cải thiện liên tục bởi phép lặp.

2.3. Bộ giải mã hoặc bộ phận xử lý có thể được tạo cấu hình để sử dụng nguồn và các mô hình nút cảm biến để cải thiện chất lượng tín hiệu được giải mã.

2.3.1. Bằng cách mô hình hóa các đặc tính thống kê của các tín hiệu nguồn và các nút cảm biến, có thể tối ưu hóa khả năng theo dõi (tín hiệu được lượng tử hóa) sự phân phối đã cho của tín hiệu và của nguồn.

2.3.1.1. Các dấu hiệu được mô hình hóa có thể gồm một hoặc một số đường bao phổ (các mô hình dự báo tuyến tính, lượng tử hóa phân phối, các thừa số tỉ lệ v.v.), các mô hình sóng hài (các bộ lọc kiểu răng lược, các bộ dự báo dài hạn v.v.), các đường bao theo thời gian (các mô hình dự báo tuyến tính, lượng tử hóa phân phối, các thừa số tỉ lệ, v.v.), các đường bao không gian và các mô hình di chuyển không gian).

2.3.1.2. Các dấu hiệu có thể được mô hình hóa với các bảng tra cứu như các bảng mã hoặc lượng tử hóa véctơ, hoặc với các mô hình tham số như các mô hình vật lý (tạo ra tiếng nói v.v.), cảm quan (các mô hình mạng che, các băng tới hạn v.v.) hoặc thống kê (các phân phối xác suất một hoặc nhiều chiều).

2.3.1.3. Các dấu hiệu có thể mô tả các dấu hiệu không gian và có âm của các nguồn, nhưng cũng là các dấu hiệu của các nút cảm biến, như chất lượng tín hiệu (ví dụ SNR cao hoặc thấp) và các dấu hiệu không gian (ví dụ, tĩnh, di chuyển, có thể mang, v.v.).

2.3.2. Bằng cách xác định độ méo gây nhiễu theo cảm quan của tín hiệu đầu ra, có thể biến đổi tín hiệu đầu ra bởi, ví dụ, các phương pháp lọc sau.

2.4. Bộ giải mã hoặc các bộ phận xử lý có thể được tạo cấu hình để sử dụng phương pháp nâng cao tín hiệu như phương pháp làm suy yếu nhiễu âm, tạo chùm và phân tách nguồn để nâng cao chất lượng tín hiệu.

2.4.1. Các phương pháp nhiều micro như tạo chùm có thể được sử dụng để có được

lợi ích của cấu hình không gian của các cảm biến để trích xuất các nguồn âm thanh được phân tách theo không gian và các thông tin khác về âm học trong phòng. Quan trọng là, các phương pháp nhiều micro thường có thể gồm ước lượng độ trễ và/hoặc ước lượng tương quan chéo.

2.4.2. Phương pháp làm suy giảm nhiễu âm có thể được sử dụng để triệt các nguồn không mong muốn như tạp âm nền.

2.4.3. Các phương pháp phân tách nguồn có thể được sử dụng để phân biệt giữa các nguồn âm thanh khác nhau.

3. Hệ thống mã hóa âm thanh và tiếng nói được phân phối mà thông tin về thông tin tín hiệu có thể được truyền, phát lại và xử lý bởi sự lựa chọn của các thiết bị/nút trong hệ thống.

3.1. Các nút cảm biến (các thiết bị với các cảm biến) nhận tín hiệu thô và có thể được tạo cấu hình để hoặc là mã hóa (lượng tử hóa và mã hóa) tín hiệu hoặc là truyền nó theo định dạng thô.

3.2. Nút bất kỳ (thường là nút cảm biến) mà có truy cập vào tín hiệu thô có thể mã hóa tín hiệu và truyền nó.

3.3. Nút bất kỳ có thể được tạo cấu hình để phát lại thông tin từ các nút khác.

3.4. Nút bất kỳ (thường là nút bộ nhận) mà có truy cập vào tín hiệu mã hóa (và thông tin giải mã phép mật mã hóa tùy chọn), có thể được tạo cấu hình để giải mã tín hiệu.

3.5. Nút trung gian như thiết bị máy chủ ở phía bộ giải mã, có thể được tạo cấu hình để hợp nhất thông tin từ các dòng khả dụng thành một hoặc một số dòng. (Các) dòng được tổ hợp có thể biểu diễn, ví dụ, sự biểu diễn âm học gốc (như hiệu suất âm nhạc) hoặc các nguồn âm thanh riêng rẽ (như các loa riêng trong hội nghị qua điện thoại). Dòng tổ hợp có thể còn được tái tạo bởi các loa phóng thanh, được lưu trữ, được truyền như hoặc được mã hóa bởi công cụ giống nhau hoặc một số công cụ mã hóa âm thanh và tiếng nói.

3.6. Cấu hình mạng có thể được tạo cấu hình để là tĩnh hoặc động, sao cho nó tối ưu hóa, ví dụ, một hoặc một vài tiêu chuẩn sau: chất lượng âm, phân phối tài nguyên, bảo mật/riêng tư.

4. Các ứng dụng được hình dung gồm ít nhất là

4.1. các ứng dụng điện thoại, trong đó sự lựa chọn thiết bị hỗ trợ được sử dụng để

chọn ra (các) tín hiệu âm thanh và tiếng nói mong muốn.

4.1.1. Các ứng dụng trong nhà và văn phòng, trong đó sự lựa chọn các thiết bị hỗ trợ được sử dụng để chọn ra (các) tín hiệu âm thanh và tiếng nói mong muốn và được truyền đến một hoặc một số vị trí từ xa.

4.1.2. Các ứng dụng truyền đạt trong hội nghị qua điện thoại, trong đó sự lựa chọn các thiết bị hỗ trợ được sử dụng để chọn ra (các) tín hiệu âm thanh và tiếng nói mong muốn và được truyền đến một hoặc một số vị trí từ xa.

4.1.3. Các hệ thống điện thoại trên xe hơi, trong đó các micro được cố định trên xe hơi và/hoặc các thiết bị hỗ trợ trên xe hơi được sử dụng để chọn ra (các) tín hiệu âm thanh và tiếng nói mong muốn và được truyền đến một hoặc một số vị trí từ xa.

4.2. Các ứng dụng trò chơi và thực tế ảo tăng cường, trong đó cảnh âm thanh của người chơi được truyền đến người chơi khác hoặc máy chủ.

4.3. Các ứng dụng tạo ra buổi hòa nhạc, biểu diễn, sân khấu, opera, thuyết trình, thể thao và các ứng dụng tạo ra sự kiện khác, trong đó các âm của người thực hiện, người chơi, khán giả hoặc toàn bộ cảnh âm thanh được ghi lại hoặc được truyền với hệ thống mã hóa âm thanh và tiếng nói được phân phối.

4.3.1. Ứng dụng có thể được thiết kế để có độ trễ thấp hoặc rất thấp để cho phép tương tác và/hoặc phát lại đồng thời và khuếch đại.

4.3.2. Ứng dụng có thể được thiết kế để cho phép tương tác giữa những người thực hiện, trong khán giả, hoặc giữa tất cả những người tham gia.

4.4. Các ứng dụng an toàn và đảm bảo, trong đó cảnh âm thanh được theo dõi để phát hiện các sự kiện nguy hiểm (ví dụ sự hoảng loạn ở sân vận động, các tai nạn ở nhà (ví dụ, người già bị ngã) v.v..

4.5. Các ứng dụng đa phương tiện, trong đó các tín hiệu âm thanh và tiếng nói được tổ hợp với video và/hoặc phương tiện khác.

Dù một số khía cạnh đã được mô tả trong ngữ cảnh của thiết bị, sẽ là rõ ràng rằng các khía cạnh này cũng biểu diễn sự mô tả của phương pháp tương ứng, trong đó khối hoặc thiết bị tương ứng với bước phương pháp hoặc dấu dấu hiệu của bước trong phương pháp. Ngoài ra, các khía cạnh được mô tả trong ngữ cảnh của bước trong phương pháp cũng biểu diễn sự mô tả của khối hoặc mục hoặc dấu hiệu tương ứng của thiết bị tương ứng. Một số hoặc tất cả các bước phương pháp có thể được chạy bởi

(hoặc sử dụng) thiết bị phần cứng, ví dụ như bộ vi xử lý, máy tính có thể lập trình hoặc mạch điện tử. Trong một số phương án, một hoặc nhiều trong số các bước quan trọng nhất trong phương pháp có thể được chạy bởi thiết bị này.

Tín hiệu được mã hóa hoặc được truyền theo sáng chế có thể được lưu trữ trên vật ghi lưu trữ số hoặc có thể được truyền trên vật ghi truyền như vật ghi truyền không dây hoặc vật ghi truyền có dây như internet.

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

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

Nói chung, các phương án của sáng chế có thể được thực hiện như sản phẩm chương trình máy tính với mã chương trình, mã chương trình có tác dụng thực hiện một trong số 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 ghi đọc được bằng máy.

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

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

Do đó, phương án nữa của phương pháp theo sáng chế là vật mang dữ liệu (hoặc vật ghi lưu trữ không tạm thời như vật ghi lưu trữ số, hoặc vật ghi có thể đọc được bằng máy tính) bao gồm chương trình máy tính đã được lưu trữ trên máy tính để

thực hiện một trong số các phương pháp được mô tả ở đây. Vật mang dữ liệu, vật ghi lưu trữ số hoặc vật ghi đã được ghi là thường là hữu hình và/hoặc không tạm thời.

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

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

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

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

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

Các phương án được mô tả ở trên chỉ minh họa cho các nguyên tắc của sáng chế. Được hiểu là các biến đổi và thay đổi về cách sắp đặt hoặc các chi tiết được mô tả ở đây sẽ là hiển nhiên với người có hiểu biết trung bình trong lĩnh vực kỹ thuật tương ứng. Do đó, ý định là sáng chế chỉ được giới hạn bởi phạm vi của các điểm yêu cầu bảo hộ sáng chế sau đây và không bị giới hạn bởi các chi tiết cụ thể được thể hiện bằng cách mô tả và giải thích của các phương án được nêu trong bản mô tả.

YÊU CẦU BẢO HỘ

1. Thiết bị mã hóa để xử lý tín hiệu đầu vào,

trong đó thiết bị bao gồm:

bộ gán trọng số theo cảm quan; và

bộ lượng tử hóa,

trong đó bộ gán trọng số theo cảm quan bao gồm bộ cấp mô hình và bộ áp dụng mô hình, trong đó bộ cấp mô hình được tạo cấu hình để cung cấp mô hình được gán trọng số theo cảm quan dựa trên tín hiệu đầu vào, và trong đó bộ áp dụng mô hình được tạo cấu hình để cung cấp phổ được gán trọng số theo cảm quan bằng cách áp dụng mô hình được gán trọng số theo cảm quan cho phổ dựa trên tín hiệu đầu vào, và

trong đó bộ lượng tử hóa được tạo cấu hình để lượng tử hóa phổ được gán trọng số theo cảm quan và để cung cấp dòng bit, trong đó bộ lượng tử hóa bao gồm bộ áp dụng ma trận ngẫu nhiên và bộ tính toán hàm dấu, trong đó bộ áp dụng ma trận ngẫu nhiên được tạo cấu hình để áp dụng ma trận ngẫu nhiên cho phổ được gán trọng số theo cảm quan để cung cấp phổ được biến đổi, và trong đó bộ tính toán hàm dấu được tạo cấu hình để tính toán hàm dấu của các thành phần của phổ được biến đổi để cung cấp dòng bit.

2. Thiết bị mã hóa theo điểm 1, trong đó bộ cấp mô hình được tạo cấu hình để cung cấp mô hình được gán trọng số theo cảm quan dựa trên phép nén phổ dựa trên tín hiệu đầu vào.

3. Thiết bị mã hóa theo điểm 1,

trong đó bộ gán trọng số theo cảm quan bao gồm bộ tính toán đường bao,

trong đó bộ tính toán đường bao được tạo cấu hình để cung cấp đường bao của phổ cường độ dựa trên tín hiệu đầu vào, và

trong đó bộ cấp mô hình được tạo cấu hình để cung cấp mô hình được gán trọng số theo cảm quan dựa trên phép nén của đường bao.

4. Thiết bị mã hóa theo điểm 1,

trong đó bộ cấp mô hình được tạo cấu hình để tính toán hàm nén mô tả phép nén của phổ dựa trên tín hiệu đầu vào hoặc của đường bao của phổ cường độ dựa trên tín hiệu đầu vào, và

trong đó bộ cấp mô hình được tạo cấu hình để tính toán mô hình được gán trọng số theo cảm quan dựa trên hàm nén.

5. Thiết bị mã hóa theo điểm 4,

trong đó hàm nén hoàn toàn tăng,

trong đó đối với giá trị vô hướng dương bất kỳ và giá trị nhỏ tùy chọn, giá trị hàm đối với giá trị vô hướng dương nhỏ hơn giá trị hàm đối với tổng của giá trị vô hướng dương và của giá trị nhỏ tùy chọn, và

trong đó đối với giá trị vô hướng dương thứ nhất và giá trị vô hướng dương thứ hai lớn hơn giá trị vô hướng dương thứ nhất, chênh lệch giữa hai giá trị hàm cho giá trị vô hướng dương thứ hai và giá trị hàm cho giá trị vô hướng dương thứ nhất là nhỏ hơn chênh lệch giữa giá trị vô hướng dương thứ hai và giá trị vô hướng dương thứ nhất.

6. Thiết bị giải mã để xử lý tín hiệu được mã hóa, trong đó thiết bị bao gồm:

bộ giải lượng tử hóa; và

bộ giải trọng số theo cảm quan,

trong đó bộ giải lượng tử hóa được tạo cấu hình để giải lượng tử hóa dòng bit được chứa bởi tín hiệu được mã hóa và để cung cấp phổ được gán trọng số theo cảm quan được tính toán,

trong đó bộ giải lượng tử hóa được tạo cấu hình để giải lượng tử hóa dòng bit bằng cách áp dụng sự giả nghịch đảo của ma trận ngẫu nhiên cho dòng bit,

trong đó bộ giải trọng số theo cảm quan bao gồm bộ làm gần đúng phổ và bộ làm gần đúng mô hình, trong đó bộ làm gần đúng phổ được tạo cấu hình để tính toán phép làm gần đúng của phổ dựa trên phổ được gán trọng số theo cảm quan được tính toán, và trong đó bộ làm gần đúng mô hình được tạo cấu hình để tính toán phép làm gần đúng của mô hình được gán trọng số theo cảm quan mà tín hiệu được mã hóa được kết hợp dựa trên phép làm gần đúng của phổ.

7. Thiết bị giải mã theo điểm 6,

trong đó bộ giải trọng số theo cảm quan bao gồm cấp giả định ban đầu,

trong đó bộ cấp giả định ban đầu được tạo cấu hình để cung cấp dữ liệu cho sự giả định ban đầu của mô hình được gán trọng số theo cảm quan, và

trong đó bộ làm gần đúng phổ được tạo cấu hình để tính toán phép làm gần

đúng của phổ dựa trên tín hiệu được mã hóa và sự giả định ban đầu hoặc phép làm gần đúng của mô hình được gán trọng số theo cảm quan.

8. Thiết bị giải mã theo điểm 6,

trong đó bộ giải lượng tử hóa được tạo cấu hình để giải lượng tử hóa dòng bit dựa trên thông tin phụ về ma trận ngẫu nhiên được chứa bởi tín hiệu được mã hóa.

9. Thiết bị giải mã theo điểm 6,

trong đó bộ giải lượng tử hóa được tạo cấu hình để nhận các dòng bit và để cung cấp phổ được gán trọng số theo cảm quan được tính toán dựa trên các dòng bit.

10. Thiết bị giải mã theo điểm 6,

trong đó thiết bị giải mã được tạo cấu hình để áp dụng phép mô hình hóa nguồn.

11. Phương pháp mã hóa tín hiệu đầu vào, trong đó phương pháp bao gồm các bước:

cung cấp mô hình gán trọng số theo cảm quan dựa trên tín hiệu đầu vào,

gán trọng số phổ của tín hiệu đầu vào bằng cách áp dụng mô hình gán trọng số theo cảm quan cho phổ của tín hiệu đầu vào; và

lượng tử hóa phổ được gán trọng số bằng cách tính toán hàm dấu của các phép chiếu ngẫu nhiên của phổ được gán trọng số,

trong đó phương pháp còn bao gồm bước thu được các phép chiếu ngẫu nhiên của phổ được gán trọng số bằng cách áp dụng ma trận ngẫu nhiên cho phổ được gán trọng số, hoặc

trong đó bước cung cấp mô hình gán trọng số theo cảm quan bao gồm nén đường bao của phổ cường độ của tín hiệu đầu vào, trong đó bước thu được đường bao được thực hiện bằng cách sử dụng giàn lọc và ma trận chéo chứa các thừa số chuẩn hóa cho từng băng,

trong đó một hoặc nhiều bước trong số các bước cung cấp, gán trọng số, và lượng tử hóa được thực hiện, ít nhất theo phần, bởi một hoặc nhiều phần tử phần cứng của thiết bị xử lý tín hiệu.

12. Phương pháp giải mã tín hiệu được mã hóa, trong đó phương pháp bao gồm các bước:

cung cấp tín hiệu theo cảm quan được lượng tử hóa bằng cách áp dụng phép giả

nghịch đảo của ma trận ngẫu nhiên cho tín hiệu được mã hóa,

tính toán phép ước lượng phổ dựa trên tín hiệu theo cảm quan được lượng tử hóa, và

tính toán phép làm gần đúng của mô hình gán trọng số theo cảm quan được sử dụng để cung cấp tín hiệu được mã hóa dựa trên phép làm gần đúng của phổ,

trong đó phương pháp còn bao gồm các bước:

cung cấp phép làm gần đúng thứ không của mô hình gán trọng số theo cảm quan sử dụng sự giả định ban đầu, và

tính toán phép ước lượng thứ không của phổ dựa trên phép làm gần đúng thứ không của mô hình gán trọng số theo cảm quan,

trong đó một hoặc nhiều bước trong số cung cấp tín hiệu theo cảm quan được lượng tử hóa, tính toán phép ước lượng, tính toán phép làm gần đúng, cung cấp phép làm gần đúng thứ không của mô hình gán trọng số theo cảm quan, và tính toán phép ước lượng thứ không của phổ được thực hiện, ít nhất theo phần, bởi một hoặc nhiều phần tử phần cứng của thiết bị xử lý tín hiệu.

13. Phương pháp theo điểm 12, trong đó phương pháp còn bao gồm bước:

thu được sự giả định ban đầu bằng cách sử dụng giàn lọc và ma trận chéo chứa các thừa số chuẩn hóa cho từng băng.

14. Vật ghi lưu trữ số không chuyển tiếp đã lưu trữ trên đó chương trình máy tính để thực hiện, khi chương trình máy tính đã nêu được chạy bởi máy tính, phương pháp mã hóa tín hiệu đầu vào, trong đó phương pháp này bao gồm các bước:

cung cấp mô hình gán trọng số theo cảm quan dựa trên tín hiệu đầu vào,

gán trọng số phổ của tín hiệu đầu vào bằng cách áp dụng mô hình gán trọng số theo cảm quan cho phổ của tín hiệu đầu vào; và

lượng tử hóa phổ được gán trọng số bằng cách tính toán hàm dấu của các phép chiếu ngẫu nhiên của phổ được gán trọng số,

trong đó phương pháp còn bao gồm bước thu được các phép chiếu ngẫu nhiên của phổ được gán trọng số bằng cách áp dụng ma trận ngẫu nhiên cho phổ được gán trọng số, hoặc

trong đó bước cung cấp mô hình gán trọng số theo cảm quan bao gồm nén

đường bao của phổ cường độ của tín hiệu đầu vào, trong đó bước thu được đường bao được thực hiện bằng cách sử dụng giàn lọc và ma trận chéo chứa các thừa số chuẩn hóa cho từng băng.

15. Vật ghi lưu trữ số không chuyển tiếp đã lưu trữ trên đó chương trình máy tính để thực hiện, khi chương trình máy tính đã nêu được chạy bởi máy tính, phương pháp giải mã tín hiệu được mã hóa, trong đó phương pháp này bao gồm các bước:

cung cấp tín hiệu theo cảm quan được lượng tử hóa bằng cách áp dụng phép giả nghịch đảo của ma trận ngẫu nhiên cho tín hiệu được mã hóa,

tính toán phép ước lượng phổ dựa trên tín hiệu theo cảm quan được lượng tử hóa, và

tính toán phép làm gần đúng của mô hình gán trọng số theo cảm quan được sử dụng để cung cấp tín hiệu được mã hóa dựa trên phép làm gần đúng của phổ,

trong đó phương pháp còn bao gồm các bước:

cung cấp phép làm gần đúng thứ không của mô hình gán trọng số theo cảm quan sử dụng sự giả định ban đầu, và

tính toán phép ước lượng thứ không của phổ dựa trên phép làm gần đúng thứ không của mô hình gán trọng số theo cảm quan.

16. Phương pháp mã hóa để xử lý tín hiệu đầu vào, trong đó phương pháp bao gồm các bước:

gán trọng số theo cảm quan; và

lượng tử hóa,

trong đó bước gán trọng số theo cảm quan bao gồm bước cung cấp mô hình và bước áp dụng mô hình, trong đó bước cung cấp bao gồm cung cấp mô hình được gán trọng số theo cảm quan dựa trên tín hiệu đầu vào, và trong đó bước áp dụng mô hình bao gồm cung cấp phổ được gán trọng số theo cảm quan bằng cách áp dụng mô hình được gán trọng số theo cảm quan cho phổ dựa trên tín hiệu đầu vào, và

trong đó bước lượng tử hóa bao gồm lượng tử hóa phổ được gán trọng số theo cảm quan và cung cấp dòng bit, trong đó bước lượng tử hóa phổ được gán trọng số theo cảm quan bao gồm áp dụng ma trận ngẫu nhiên và tính toán hàm dấu, trong đó bước áp dụng ma trận ngẫu nhiên bao gồm áp dụng ma trận ngẫu nhiên cho phổ được

gán trọng số theo cảm quan để cung cấp phổ được biến đổi, và trong đó bước tính toán hàm dấu bao gồm tính toán hàm dấu của các thành phần của phổ được biến đổi để cung cấp dòng bit.

17. Vật ghi lưu trữ số không chuyển tiếp đã lưu trữ trên đó chương trình máy tính để thực hiện, khi chương trình máy tính đã nêu được chạy bởi máy tính, phương pháp mã hóa để xử lý tín hiệu đầu vào, trong đó phương pháp mã hóa này bao gồm các bước:

gán trọng số theo cảm quan; và

lượng tử hóa,

trong đó bước gán trọng số theo cảm quan bao gồm bước cung cấp mô hình và bước áp dụng mô hình, trong đó bước cung cấp bao gồm cung cấp mô hình được gán trọng số theo cảm quan dựa trên tín hiệu đầu vào, và trong đó bước áp dụng mô hình bao gồm cung cấp phổ được gán trọng số theo cảm quan bằng cách áp dụng mô hình được gán trọng số theo cảm quan cho phổ dựa trên tín hiệu đầu vào, và

trong đó bước lượng tử hóa bao gồm lượng tử hóa phổ được gán trọng số theo cảm quan và cung cấp dòng bit, trong đó bước lượng tử hóa phổ được gán trọng số theo cảm quan bao gồm áp dụng ma trận ngẫu nhiên và tính toán hàm dấu, trong đó bước áp dụng ma trận ngẫu nhiên bao gồm áp dụng ma trận ngẫu nhiên cho phổ được gán trọng số theo cảm quan để cung cấp phổ được biến đổi, và trong đó bước tính toán hàm dấu bao gồm tính toán hàm dấu của các thành phần của phổ được biến đổi để cung cấp dòng bit.

18. Phương pháp giải mã để xử lý tín hiệu âm thanh được mã hóa, trong đó phương pháp này bao gồm các bước:

giải lượng tử hóa; và

giải trọng số theo cảm quan,

trong đó bước giải lượng tử hóa bao gồm giải lượng tử hóa dòng bit được chứa bởi tín hiệu được mã hóa và cung cấp phổ được gán trọng số theo cảm quan được tính toán, trong đó bước giải lượng tử hóa dòng bit bao gồm giải lượng tử hóa dòng bit bằng cách áp dụng sự giả nghịch đảo của ma trận ngẫu nhiên cho dòng bit, và

trong đó bước giải trọng số theo cảm quan bao gồm thực hiện phép làm gần

đúng phổ và thực hiện phép làm gần đúng mô hình, trong đó bước thực hiện phép làm gần đúng phổ bao gồm tính toán phép làm gần đúng của phổ dựa trên phổ được gán trọng số theo cảm quan được tính toán, và trong đó bước thực hiện phép làm gần đúng mô hình bao gồm tính toán phép làm gần đúng của mô hình được gán trọng số theo cảm quan mà tín hiệu được mã hóa được kết hợp dựa trên phép làm gần đúng của phổ.

19. Vật ghi lưu trữ số không chuyển tiếp đã lưu trữ trên đó chương trình máy tính để thực hiện, khi chương trình máy tính đã nêu được chạy bởi máy tính, phương pháp giải mã để xử lý tín hiệu được mã hóa, trong đó phương pháp này bao gồm các bước:

giải lượng tử hóa; và

giải trọng số theo cảm quan,

trong đó bước giải lượng tử hóa bao gồm giải lượng tử hóa dòng bit được chứa bởi tín hiệu được mã hóa và cung cấp phổ được gán trọng số theo cảm quan được tính toán, trong đó bước giải lượng tử hóa dòng bit bao gồm giải lượng tử hóa dòng bit bằng cách áp dụng sự giả nghịch đảo của ma trận ngẫu nhiên cho dòng bit, và

trong đó bước giải trọng số theo cảm quan bao gồm thực hiện phép làm gần đúng phổ và thực hiện phép làm gần đúng mô hình, trong đó bước thực hiện phép làm gần đúng phổ bao gồm tính toán phép làm gần đúng của phổ dựa trên phổ được gán trọng số theo cảm quan được tính toán, và trong đó bước thực hiện phép làm gần đúng mô hình bao gồm tính toán phép làm gần đúng của mô hình được gán trọng số theo cảm quan mà tín hiệu được mã hóa được kết hợp dựa trên phép làm gần đúng của phổ.

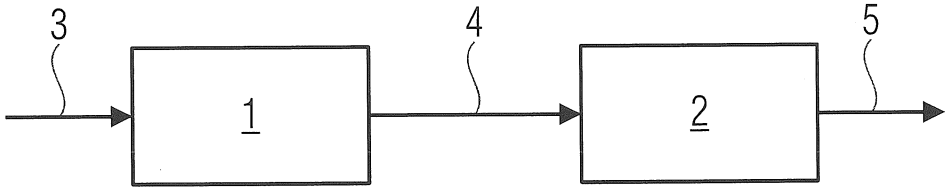


Fig. 1

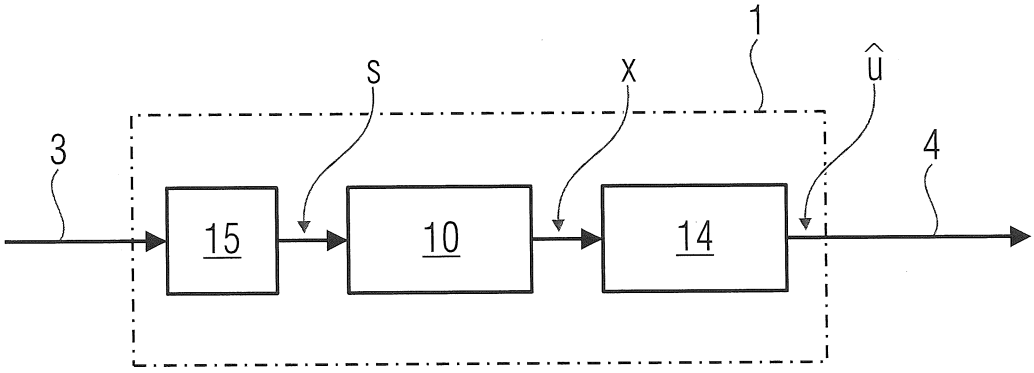


Fig. 2

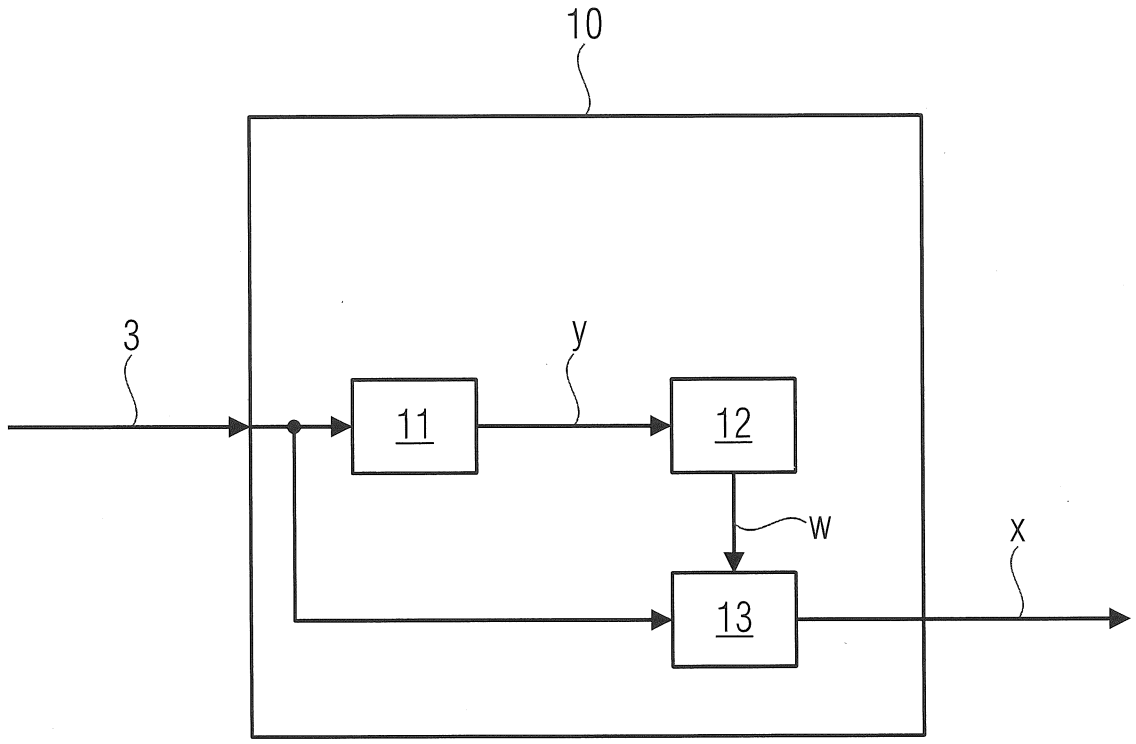


Fig. 3

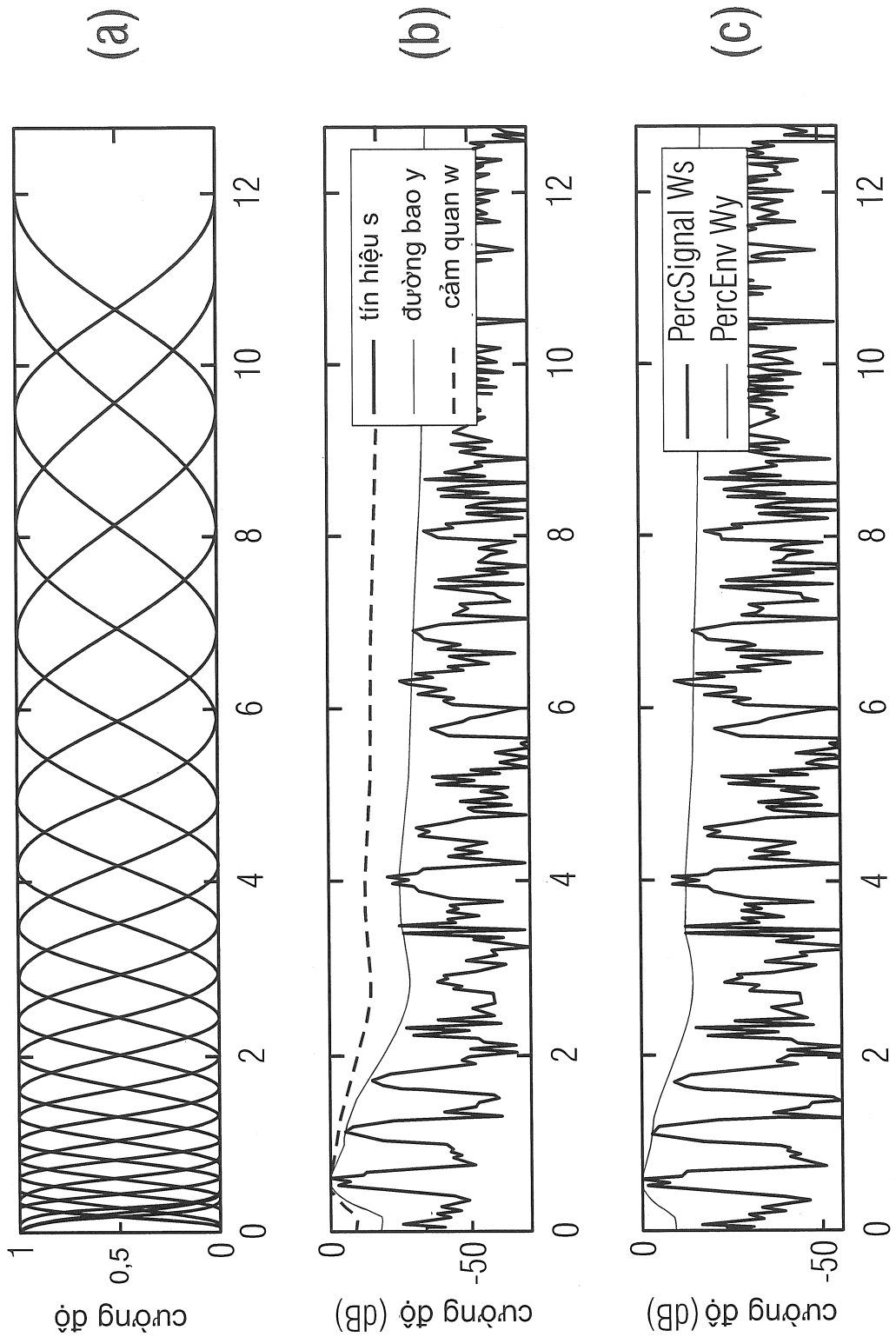

(a)
(b)
(c)
cường độ
cường độ (dB)
cường độ (dB)
tín hiệu s
đường bao y
cảm quan w
PercSignal Ws
PercEnv Wy
tần số (kHz)
1
0,5
0
-50
0
2
4
6
8
10
12

Fig. 4

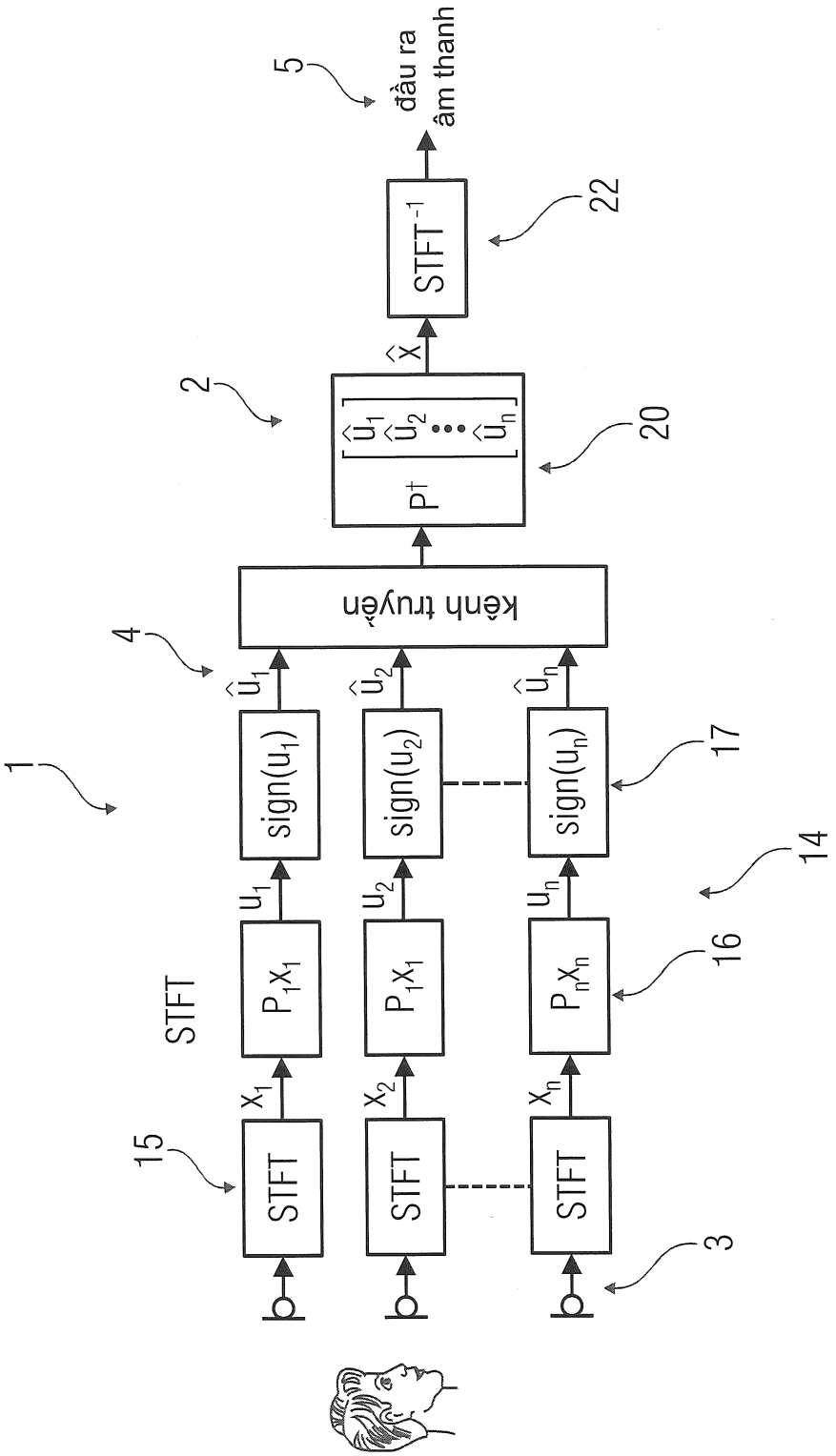

1
STFT
15
STFT
STFT
STFT
3
x_1
x_2
x_n
P_1x_1
P_1x_1
P_nx_n
16
14
u_1
u_2
u_n
sign(u_1)
sign(u_2)
sign(u_n)
17
4
$\hat{u}_1$
$\hat{u}_2$
$\hat{u}_n$
kênh truyền
2
$P^\dagger$
$\hat{u}_1$ $\hat{u}_2$ ••• $\hat{u}_n$
20
$\hat{x}$
$STFT^{-1}$
22
5
đầu ra
âm thanh

Fig. 5

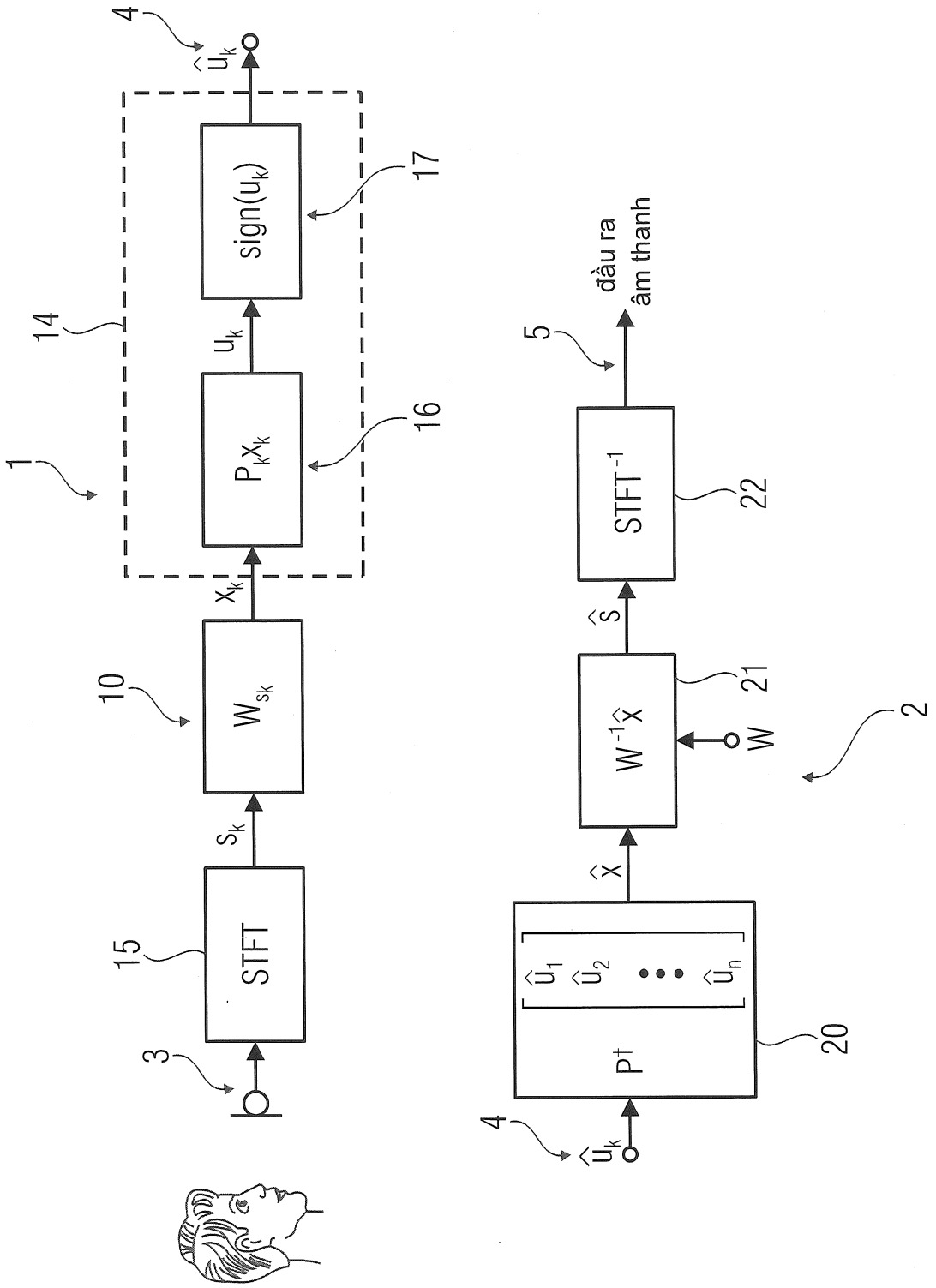

4
$\hat{u}_k$
sign(u_k)
17
14
u_k
$P_k x_k$
16
1
x_k
10
W_{sk}
s_k
15
STFT
3
đầu ra
âm thanh
5
STFT^{-1}
22
$\hat{s}$
$W^{-1}\hat{x}$
21
W
2
$\hat{x}$
$\hat{u}_1$ $\hat{u}_2$ ••• $\hat{u}_n$
$P^\dagger$
20
4
$\hat{u}_k$

Fig. 6

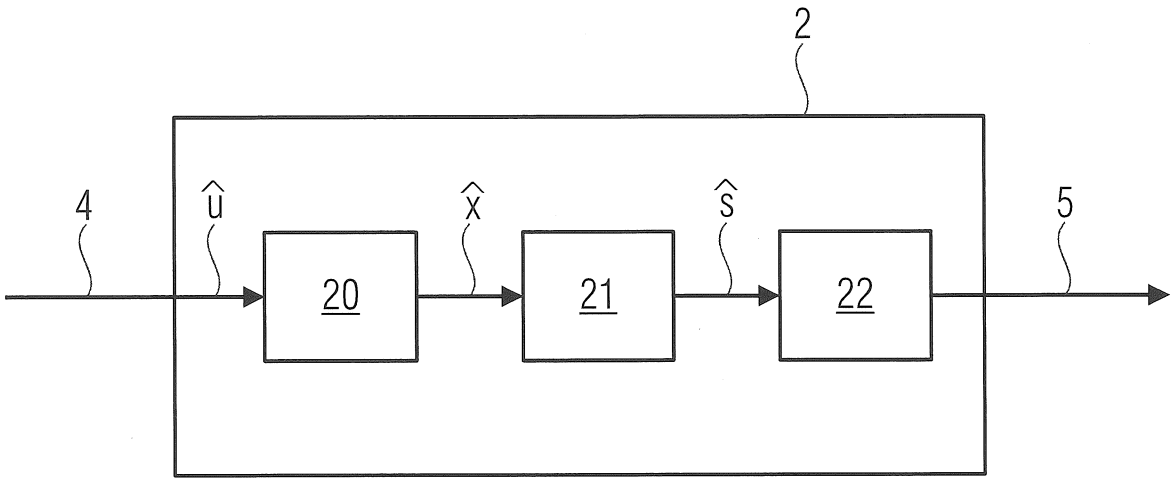


Fig. 7

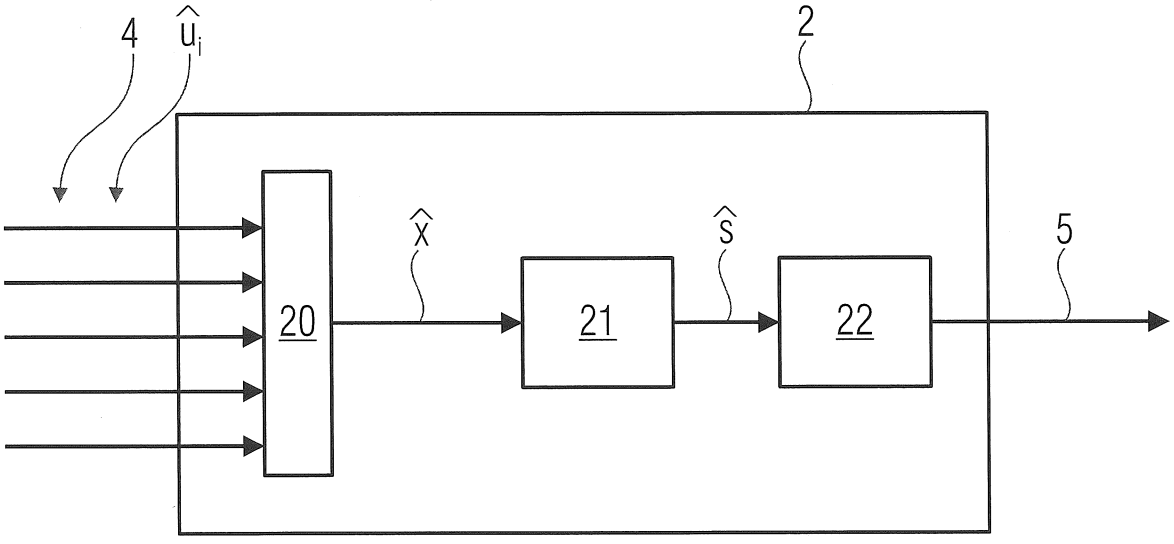


Fig. 8

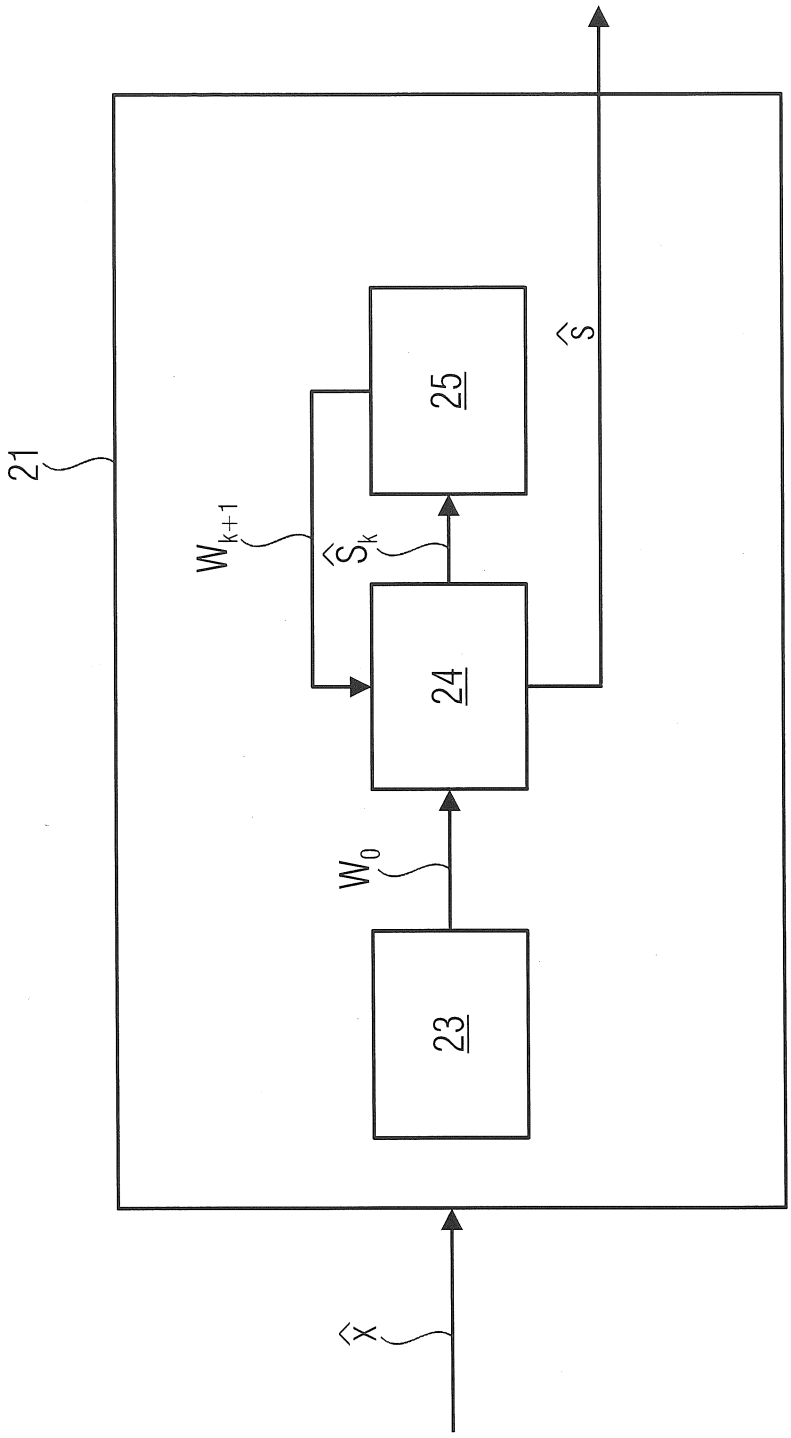


Fig. 9

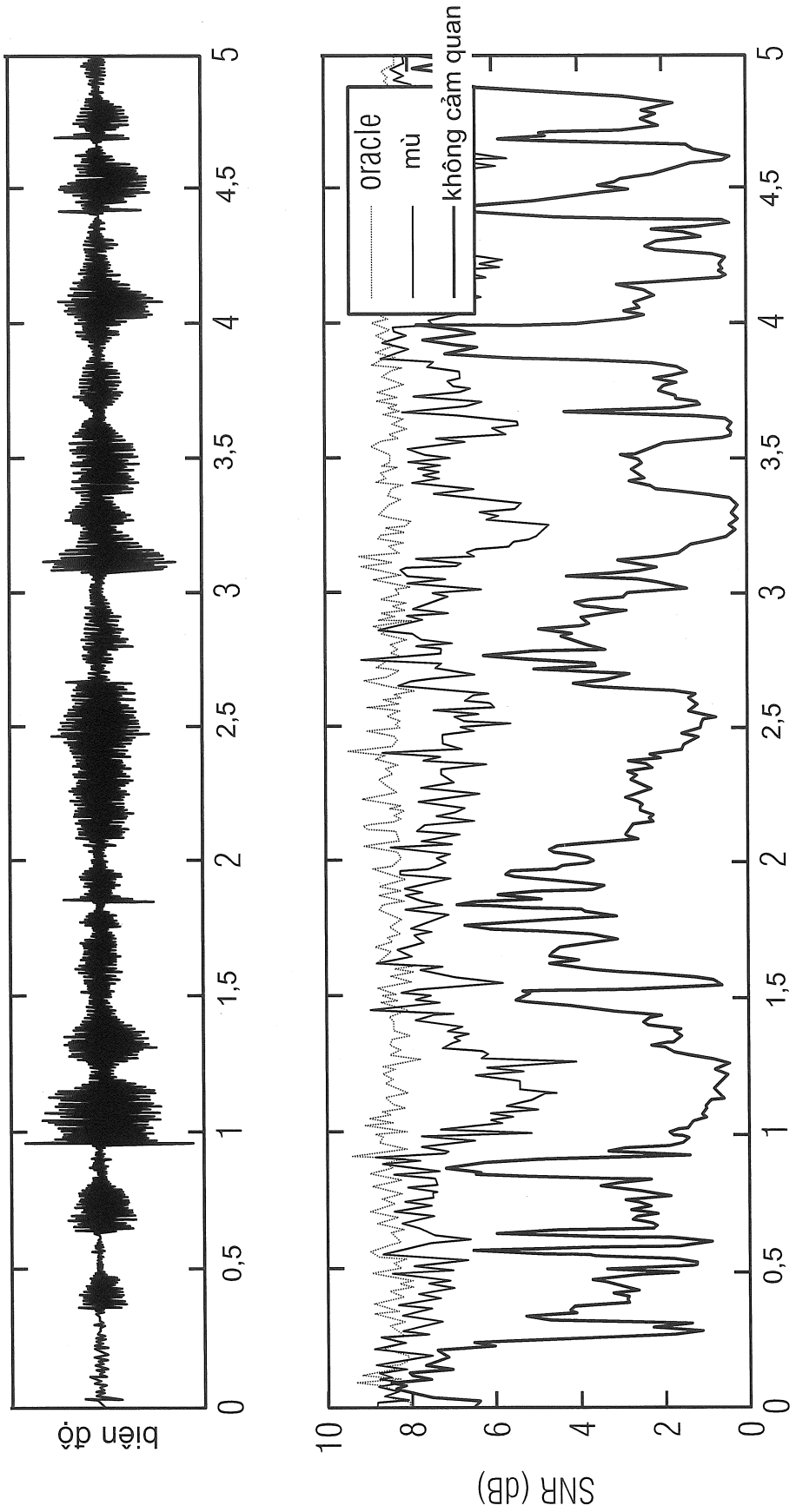

biên độ
0
0,5
1
1,5
2
2,5
3
3,5
4
4,5
5
SNR (dB)
10
8
6
4
2
0
oracle
mù
không cảm quan
thời gian (s)

Fig. 10

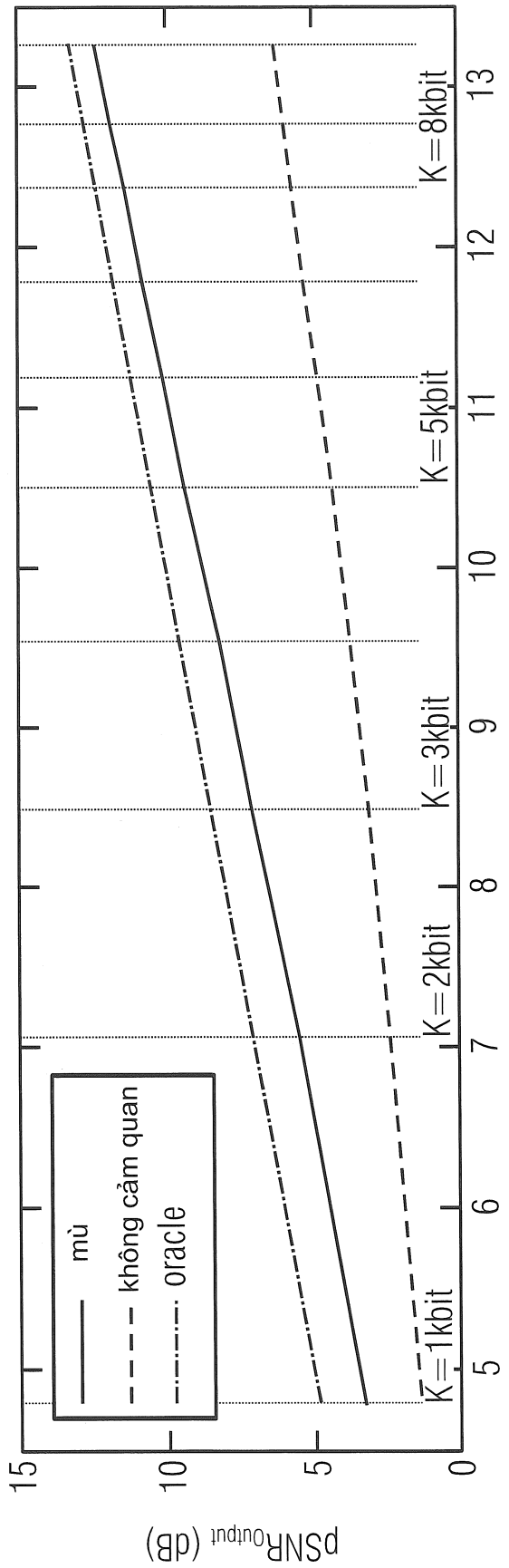

mù
không cảm quan
oracle
K=1kbit
K=2kbit
K=3kbit
K=5kbit
K=8kbit
pSNR_Output (dB)
pSNR_Oracle (dB)
0
5
10
15
5
6
7
8
9
10
11
12
13

Fig. 11

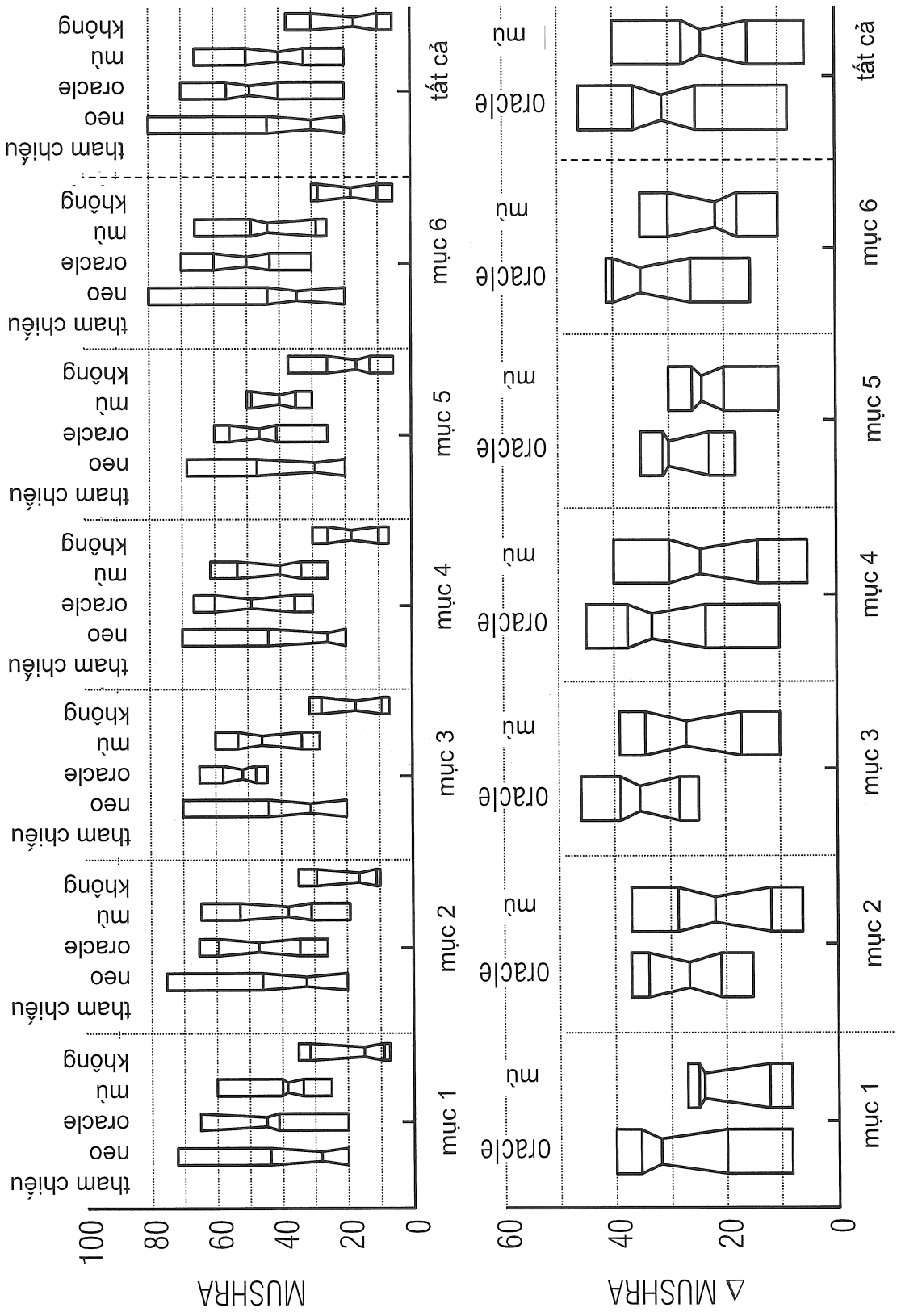

MUSHRA
Δ MUSHRA
0
20
40
60
80
100
mục 1
mục 2
mục 3
mục 4
mục 5
mục 6
tất cả
tham chiếu
neo
oracle
mũ
không

Fig. 12