



(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)



1-0026185

(51)⁷ H04S 7/00; H04S 3/00

(13) B

- (21) 1-2015-04341 (22) 07/04/2014
(86) PCT/EP2014/056917 07/04/2014 (87) WO 2014/166863 A1 16/10/2014
(30) 13163621.9 12/04/2013 EP; 13182103.5 28/08/2013 EP
(45) 25/11/2020 392 (43) 25/01/2016 334A
(73) Fraunhofer-Gesellschaft zur Foerderung der angewandten Forschung e.V. (DE)
Hansastrasse 27c, 80686 Muenchen, Germany (DE)
(72) UHLE, Christian (DE); PROKEIN, Peter (DE); HELLMUTH, Oliver (DE);
SCHARRER, Sebastian (DE); HABETS, Emanuel (NL).
(74) CÔNG TY LUẬT TRÁCH NHIỆM HỮU HẠN AMBYS HÀ NỘI (AMBYS
HANOI)

(54) THIẾT BỊ, HỆ THỐNG VÀ PHƯƠNG PHÁP TẠO RA TÍN HIỆU ÂM THANH ĐƯỢC BIẾN ĐỔI

(57) Sáng chế đề cập đến thiết bị, hệ thống và phương pháp tạo ra tín hiệu âm thanh được biến đổi. Thiết bị tạo ra tín hiệu âm thanh được biến đổi bao gồm hai hoặc nhiều hơn hai kênh âm thanh được biến đổi từ tín hiệu đầu vào âm thanh bao gồm hai hoặc nhiều hơn hai kênh đầu vào âm thanh được cung cấp. Thiết bị bao gồm bộ tạo thông tin (110) để tạo ra thông tin tín hiệu trên trộn giảm. Bộ tạo thông tin (110) được làm thích ứng để tạo ra thông tin tín hiệu bằng cách tổ hợp giá trị phổ của mỗi kênh đầu vào âm thanh trong số hai hoặc nhiều hơn hai kênh đầu vào âm thanh theo cách thứ nhất. Hơn nữa, bộ tạo thông tin (110) được làm thích ứng để tạo ra thông tin trộn giảm bằng cách tổ hợp giá trị phổ của mỗi kênh đầu vào âm thanh trong số hai hoặc nhiều hơn hai kênh đầu vào âm thanh theo cách thứ hai khác cách thứ nhất. Hơn nữa, bộ tạo thông tin (110) được làm thích ứng để kết hợp thông tin tín hiệu và thông tin trộn giảm để thu được thông tin tín hiệu trên trộn giảm. Hơn nữa, thiết bị bao gồm bộ suy giảm tín hiệu (120) để làm suy giảm hai hoặc nhiều hơn hai kênh đầu vào âm thanh phụ thuộc vào thông tin tín hiệu trên trộn giảm để thu được hai hoặc nhiều hơn hai kênh âm thanh được biến đổi.



Lĩnh vực kỹ thuật của sáng chế

Sáng chế đề cập đến việc xử lý tín hiệu âm thanh, và, cụ thể, đề cập đến việc định tỷ lệ tín hiệu trung tâm và cải thiện âm lập thể dựa vào tỷ lệ tín hiệu trên trộn giảm.

Tình trạng kỹ thuật của sáng chế

Thông thường các tín hiệu âm thanh là hỗn hợp của các âm thanh thẳng và các âm thanh (hoặc khuếch tán) ở môi trường xung quanh. Các tín hiệu thẳng được phát ra bởi các nguồn âm thanh, ví dụ nhạc cụ, người hát hoặc loa phát thanh, và đến bộ nhận theo đường ngắn nhất có thể, ví dụ tai người nghe hoặc micrô. Khi nghe âm thanh thẳng, nó được hiểu là đến từ hướng của nguồn âm thanh. Các tín hiệu thính giác thích hợp đối với việc định vị và đối với các đặc tính âm thanh trong không gian khác là chênh lệch mức độ âm nội tại (interaural level difference - ILD), chênh lệch thời gian âm nội tại (interaural time difference - ITD) và sự nhất quán âm nội tại. Các sóng âm thanh thẳng gây ra ILD và ITD đồng nhất được hiểu là đến từ cùng hướng. Trong trường hợp không có âm thanh ở môi trường xung quanh, các tín hiệu đi đến tai phải và tai trái hoặc bộ các cảm biến cách quãng khác bất kỳ là nhất quán.

Ngược lại, các âm thanh ở môi trường xung quanh được phát ra bởi nhiều nguồn âm thanh cách quãng hoặc các biên phản xạ âm thanh góp thành âm thanh giống nhau. Khi sóng âm thanh chạm tới tường chắn trong phòng, một phần sóng âm thanh được phản xạ, và sự chồng chất của tất cả các phản xạ trong phòng, tiếng vang, là ví dụ nổi bật về các âm thanh ở môi trường xung quanh. Các ví dụ khác là tiếng vỗ tay, tiếng ồn gây nhiễu và tiếng ồn của gió. Các âm thanh ở môi trường xung quanh được hiểu là ở trạng thái khuếch tán, không thể xác định vị trí, và gây ra vết hằn của đường bao (thuộc trạng thái "bị nhúng chìm trong âm thanh") cho người nghe. Khi giữ lại trường âm thanh ở môi trường xung quanh sử dụng bộ các cảm biến cách quãng, các tín hiệu ghi được ít nhất là một phần không nhất quán.

Liên quan đến tình trạng kỹ thuật về việc tách, phân tích hoặc định tỷ lệ hoặc là dựa vào thông tin quét, tức là các chênh lệch mức độ liên kênh (inter-channel level differences - ICLD) và các chênh lệch thời gian liên kênh (inter-channel time differences - ICTD), hoặc dựa vào các đặc trưng tín hiệu của âm thanh thẳng và âm thanh ở môi trường xung quanh. Các phương pháp tận dụng ICLD trong những bản thu âm lập thể hai-kênh là phương pháp trộn tăng được mô tả trong tài liệu tham khảo [7], thuật toán Phân biệt Azimuth và Resynthesis (ADRes) trong tài liệu tham khảo [8], trộn tăng từ các tín hiệu đầu vào hai-kênh thành ba kênh được đề xuất bởi Vickers trong tài liệu tham khảo [9], và trích xuất tín hiệu trung tâm được mô tả trong tài liệu tham khảo [10].

Kỹ thuật ước lượng không trộn suy biến (Degenerate Unmixing Estimation Technique - DUET) trong tài liệu tham khảo [11, 12] dựa vào việc tạo nhóm các ngăn thời gian-tần số thành các bộ có ICLD và ICTD tương tự nhau. Hạn chế của các phương pháp ban đầu là tần số cực đại mà có thể được xử lý bằng nửa tốc độ âm thanh qua khoảng cách micro cực đại (do những mơ hồ trong ước lượng ICTD) mà được giải quyết trong tài liệu tham khảo [13]. Hiệu suất của phương pháp giảm xuống khi các nguồn chồng lấp trong miền thời gian-tần số và khi tiếng vang tăng lên. Các phương pháp khác dựa vào ICLD và ICTD là thuật toán ADRes biến đổi trong tài liệu tham khảo [14], mà mở rộng thuật toán ADRes trong tài liệu tham khảo [8] để xử lý những bản thu micro cách quãng, phương pháp dựa vào tương quan thời gian-tần số (AD-TIFCORR) trong tài liệu tham khảo [15] cho các hỗn hợp thời gian-trễ, ước lượng hướng của ma trận trộn (Direction Estimation of Mixing Matrix - DEMIX) cho các hỗn hợp không có tiếng vang trong tài liệu tham khảo [16], mà bao gồm phép đo tin cậy mà chỉ một nguồn là hoạt động ở ngăn thời gian-tần số cụ thể, tách và định vị nguồn kỳ vọng tối đa hóa theo mô hình (Model-based Expectation-Maximization Source Separation and Localization - MESSL) trong tài liệu tham khảo [17], và các phương pháp bắt trước cơ chế nghe ở hai tai người như trong tài liệu tham khảo [18, 19].

Dù các phương pháp tách nguồn mù (Blind Source Separation - BSS) sử dụng các tín hiệu trong không gian của các thành phần tín hiệu thẳng được đề cập ở trên, ngoài ra việc trích xuất và làm suy giảm các tín hiệu ở môi trường xung quanh liên

quan đến phương pháp đã có. Các phương pháp dựa vào sự kết hợp liên kênh (inter - channel coherence - ICC) trong các tín hiệu hai-kênh được mô tả trong tài liệu tham khảo [22, 7, 23]. Áp dụng việc lọc thích ứng được đề xuất trong tài liệu tham khảo [24], với lý do căn bản là các tín hiệu thẳng có thể được dự báo qua các kênh trong khi các âm thanh khuếch tán thu được từ sai số dự báo.

Phương pháp trộn tăng của các tín hiệu âm lập thể hai-kênh dựa vào việc lọc Wiener đa kênh ước lượng cả hai, ICLD của các âm thanh thẳng và các mật độ phổ công suất (power spectral densities - PSD) của thành phần tín hiệu ở môi trường xung quanh và thành phần tín hiệu thẳng trong tài liệu tham khảo [25].

Gần như việc lấy ra các tín hiệu ở môi trường xung quanh từ các ghi chép kênh đơn bao gồm sử dụng việc tìm thừa số ma trận không âm thuộc biểu diễn thời gian-tần số của tín hiệu đầu vào, trong đó tín hiệu ở môi trường xung quanh được thu từ phần dư của phép tính xấp xỉ trong tài liệu tham khảo [26], việc lấy ra đặc điểm mức độ-thấp và việc tìm hiểu được kiểm soát trong tài liệu tham khảo [27], và việc ước lượng đáp ứng xung của hệ thống vang và lọc nghịch đảo trong miền tần số trong tài liệu tham khảo [28].

Tài liệu tham khảo

- [1] International Telecommunication Union, Radiocommunication Assembly, “Multichannel stereophonic sound system with and without accompanying picture,” Recommendation ITU-R BS.775-2, 2006, Geneva, Switzerland.
- [2] J. Berg and F. Rumsey, “Identification of quality attributes of spatial sound by repertory grid technique,” *J. Audio Eng. Soc.*, vol. 54, pp. 365–379, 2006.
- [3] J. Blauert, *Spatial Hearing*, MIT Press, 1996.
- [4] F. Rumsey, “Controlled subjective assessment of two-to-five channel surround sound processing algorithms,” *J. Audio Eng. Soc.*, vol. 47, pp. 563–582, 1999.
- [5] H. Fuchs, S. Tuff, and C. Bustad, “Dialogue enhancement - technology and experiments,” *EBU Technical Review*, vol. Q2, pp. 1–11, 2012.

- [6] J.-H. Bach, J. Anemüller, and B. Kollmeier, “Robust speech detection in real acoustic backgrounds with perceptually motivated features,” *Speech Communication*, vol. 53, pp. 690–706, 2011.
- [7] C. Avendano and J.-M. Jot, “A frequency-domain approach to multi-channel upmix,” *J. Audio Eng. Soc.*, vol. 52, 2004.
- [8] D. Barry, B. Lawlor, and E. Coyle, “Sound source separation: Azimuth discrimination and resynthesis,” in *Proc. Int. Conf. Digital Audio Effects (DAFx)*, 2004.
- [9] E. Vickers, “Two-to-three channel upmix for center channel derivation and speech enhancement,” in *Proc. Audio Eng. Soc. 127th Conv.*, 2009.
- [10] D. Jang, J. Hong, H. Jung, and K. Kang, “Center channel separation based on spatial analysis,” in *Proc. Int. Conf. Digital Audio Effects (DAFx)*, 2008.
- [11] A. Jourjine, S. Rickard, and O. Yilmaz, “Blind separation of disjoint orthogonal signals: Demixing N sources from 2 mixtures,” in *Proc. Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 2000.
- [12] O. Yilmaz and S. Rickard, “Blind separation of speech mixtures via time-frequency masking,” *IEEE Trans. on Signal Proc.*, vol. 52, pp. 1830–1847, 2004.
- [13] S. Rickard, “The DUET blind source separation algorithm,” in *Blind Speech Separation*, S. Makino, T.-W. Lee, and H. Sawada, Eds. Springer, 2007.
- [14] N. Cahill, R. Cooney, K. Humphreys, and R. Lawlor, “Speech source enhancement using a modified ADress algorithm for applications in mobile communications,” in *Proc. Audio Eng. Soc. 127th Conv.*, 2006.
- [15] M. Puigt and Y. Deville, “A time-frequency correlation-based blind source separation method for time-delay mixtures,” in *Proc. Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 2006.

- [16] Simon Arberet, Remi Gribonval, and Frederic Bimbot, “A robust method to count and locate audio sources in a stereophonic linear anechoic mixture,” in *Proc. Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 2007.
- [17] M.I. Mandel, R.J. Weiss, and D.P.W. Ellis, “Model-based expectation-maximization source separation and localization,” *IEEE Trans. on Audio, Speech and Language Proc.*, vol. 18, pp. 382–394, 2010.
- [18] H. Viste and G. Evangelista, “On the use of spatial cues to improve binaural source separation,” in *Proc. Int. Conf. Digital Audio Effects (DAFx)*, 2003.
- [19] A. Favrot, M. Erne, and C. Faller, “Improved cocktail-party processing,” in *Proc. Int. Conf. Digital Audio Effects (DAFx)*, 2006.
- [20] US patent 7,630,500 B1, P.E. Beckmann, 2009
- [21] US patent 7,894,611 B1, P.E. Beckmann, 2011
- [22] J.B. Allen, D.A. Berkeley, and J. Blauert, “Multimicrophone signal-processing technique to remove room reverberation from speech signals,” *J. Acoust. Soc. Am.*, vol. 62, 1977.
- [23] J. Merimaa, M. Goodwin, and J.-M. Jot, “Correlation-based ambience extraction from stereo recordings,” in *Proc. Audio Eng. Soc. 127th Conv.*, 2007.
- [24] J. Usher and J. Benesty, “Enhancement of spatial sound quality: A new reverberation-extraction audio upmixer,” *IEEE Trans. on Audio, Speech, and Language Processing*, vol. 15, pp. 2141–2150, 2007.
- [25] C. Faller, “Multiple-loudspeaker playback of stereo signals,” *J. Audio Eng. Soc.*, vol. 54, 2006.
- [26] C. Uhle, A. Walther, O. Hellmuth, and J. Herre, “Ambience separation from mono recordings using Non-negative Matrix Factorization,” in *Proc. Audio Eng. Soc. 30th Int. Conf.*, 2007.
- [27] C. Uhle and C. Paul, “A supervised learning approach to ambience extraction from mono recordings for blind upmixing,” in *Proc. Int. Conf. Digital Audio Effects (DAFx)*, 2008.

- [28] G. Soulodre, "System for extracting and changing the reverberant content of an audio input signal," US Patent 8,036,767, Oct. 2011.
- [29] International Telecommunication Union, Radiocommunication Assembly, "Algorithms to measure audio programme loudness and true-peak audio level," Recommendation ITUR BS.1770-2, March 2011, Geneva, Switzerland.

Bản chất kỹ thuật của sáng chế

Mục đích của sáng chế là để đề xuất các khái niệm được cải thiện để xử lý tín hiệu âm thanh. Mục đích của sáng chế được giải quyết bởi thiết bị theo yêu cầu bảo hộ 1, bởi hệ thống theo yêu cầu bảo hộ 8, bởi phương pháp theo yêu cầu bảo hộ 9 và bởi vật ghi có thể đọc được bằng máy tính bao gồm chương trình máy tính theo yêu cầu bảo hộ 10.

Thiết bị tạo ra tín hiệu âm thanh được biến đổi bao gồm hai hoặc nhiều hơn hai kênh âm thanh được biến đổi từ tín hiệu đầu vào âm thanh bao gồm hai hoặc nhiều hơn hai kênh đầu vào âm thanh được cung cấp. Thiết bị bao gồm bộ tạo thông tin để tạo ra thông tin tín hiệu trên trộn giảm. Bộ tạo thông tin được làm thích ứng để tạo ra thông tin tín hiệu bằng cách tổ hợp giá trị phổ của mỗi kênh đầu vào âm thanh trong số hai hoặc nhiều hơn hai kênh đầu vào âm thanh theo cách thứ nhất. Hơn nữa, bộ tạo thông tin được làm thích ứng để tạo ra thông tin trộn giảm bằng cách tổ hợp giá trị phổ của mỗi kênh đầu vào âm thanh trong số hai hoặc nhiều hơn hai kênh đầu vào âm thanh theo cách thứ hai mà khác so với cách thứ nhất. Hơn nữa, bộ tạo thông tin được làm thích ứng để tổ hợp thông tin tín hiệu và thông tin trộn giảm để thu được thông tin tín hiệu trên trộn giảm. Hơn nữa, thiết bị bao gồm bộ suy giảm tín hiệu để làm suy giảm hai hoặc nhiều hơn hai kênh đầu vào âm thanh phụ thuộc vào thông tin tín hiệu trên trộn giảm để thu được hai hoặc nhiều hơn hai kênh âm thanh được biến đổi.

Trong phương án cụ thể, thiết bị có thể, ví dụ, được làm thích ứng để tạo ra tín hiệu âm thanh được biến đổi bao gồm ba hoặc nhiều hơn ba kênh âm thanh được biến đổi từ tín hiệu đầu vào âm thanh bao gồm ba hoặc nhiều hơn ba kênh đầu vào âm thanh.

Trong phương án, số kênh âm thanh được biến đổi bằng hoặc nhỏ hơn số kênh đầu vào âm thanh, hoặc trong đó số kênh âm thanh được biến đổi nhỏ hơn số kênh đầu vào âm thanh. Ví dụ, theo phương án cụ thể, thiết bị có thể được làm thích ứng để tạo ra tín hiệu âm thanh được biến đổi bao gồm hai hoặc nhiều hơn hai kênh âm thanh được biến đổi từ tín hiệu đầu vào âm thanh bao gồm hai hoặc nhiều hơn hai kênh đầu vào âm thanh, trong đó số kênh âm thanh được biến đổi bằng số kênh đầu vào âm thanh.

Các phương án đề xuất các khái niệm mới để định tỷ lệ mức độ của trung tâm ảo trong các tín hiệu âm thanh được đề xuất. Các tín hiệu đầu vào được xử lý trong miền thời gian-tần số sao cho các thành phần âm thanh thẳng có năng lượng xấp xỉ bằng nhau trong tất cả các kênh được khuếch đại hoặc bị suy giảm. Các trọng số phổ được định giá trị thực được thu từ tỷ lệ của tổng các mật độ phổ công suất của tất cả tín hiệu kênh đầu vào với mật độ phổ công suất của tín hiệu tổng. Những áp dụng của các khái niệm đã có là trộn tăng các bản ghi âm lập thể hai-kênh để việc tái tạo nó sử dụng các thiết lập âm thanh bao quanh, cải thiện âm lập thể, cải thiện âm thanh hội thoại, và như xử lý trước để phân tích âm thanh về mặt ngữ nghĩa.

Các phương án đề xuất các khái niệm mới để khuếch đại hoặc suy giảm tín hiệu trung tâm trong tín hiệu âm thanh. Ngược lại với các khái niệm có trước, cả việc dịch chuyển ngang và khuếch tán của các thành phần tín hiệu được xem xét. Hơn nữa, việc sử dụng các tham số có nghĩa về mặt ngữ nghĩa được mô tả để hỗ trợ người sử dụng khi việc thực hiện các khái niệm được sử dụng.

Một số phương án tập trung vào việc định tỷ lệ tín hiệu trung tâm, tức là sự khuếch đại hoặc suy giảm của các tín hiệu trung tâm trong các bản ghi âm thanh. Tín hiệu trung tâm là, ví dụ, được xác định ở đây là tổng của tất cả các thành phần tín hiệu thẳng có cường độ xấp xỉ bằng nhau trong tất cả các kênh và các chênh lệch thời gian không đáng kể giữa các kênh.

Những áp dụng khác nhau của việc xử lý tín hiệu âm thanh và lợi ích tái tạo từ việc định tỷ lệ tín hiệu trung tâm, ví dụ, trộn tăng, cải thiện âm thanh hội thoại, và phân tích âm thanh về mặt ngữ nghĩa.

Trộn tăng đề cập đến quy trình tạo thành tín hiệu đầu ra đã đưa tín hiệu đầu vào với ít kênh hơn. Ứng dụng chủ yếu của nó là tái tạo các tín hiệu hai-kênh sử dụng các thiết lập âm thanh bao quanh như ví dụ được chỉ rõ trong tài liệu tham khảo [1]. Nghiên cứu về chất lượng chủ quan của âm thanh trong không gian trong tài liệu tham khảo [2] cho biết việc định vị trong tài liệu tham khảo [3], sự khoanh vùng và chiều rộng là những thuộc tính mô tả nổi bật của âm thanh. Những kết quả về việc đánh giá chủ quan của 2 đến 5 thuật toán trộn tăng trong tài liệu tham khảo [4] được thể hiện để việc sử dụng loa phát thanh trung tâm bổ sung có thể thu hẹp ảnh vô tuyến âm lập thể. Công việc đã có được thúc đẩy bằng cách giả định rằng việc định vị, khoanh vùng và chiều rộng có thể được duy trì hoặc thậm chí được cải thiện khi loa phát thanh trung tâm bổ sung làm giảm chủ yếu các thành phần tín hiệu thẳng mà có thể được quét đến trung tâm, và khi các thành phần tín hiệu này bị suy giảm trong các tín hiệu loa phát thanh lệch tâm.

Cải thiện âm thanh hội thoại đề cập đến việc cải thiện độ rõ tiếng, ví dụ trong phát thanh và âm thanh phim ảnh, và thường được yêu cầu khi các âm thanh nền quá to liên quan đến hội thoại trong tài liệu tham khảo [5]. Cụ thể việc này áp dụng cho những người nặng tai, những người điếc, trong những môi trường ồn ào hoặc khi chênh lệch mức chấn hai tai bị giảm xuống do việc dịch chuyển loa phát thanh hẹp. Phương pháp theo các khái niệm có thể được áp dụng cho việc xử lý các tín hiệu đầu vào trong đó hội thoại được quét đến trung tâm để làm suy giảm các âm thanh nền và do đó cho phép độ rõ tiếng tốt hơn.

Phân tích âm thanh về mặt ngữ nghĩa (hoặc phân tích nội dung âm thanh) bao gồm các quy trình suy luận các mô tả có nghĩa từ các tín hiệu âm thanh, ví dụ dò phách hoặc phiên âm giai điệu chính. Hiệu suất của các phương pháp tính toán thường bị giảm khi các âm thanh quan tâm được nhúng trong các âm thanh nền, xem ví dụ trong tài liệu tham khảo [6]. Vì đây là thực tiễn chung trong việc tạo ra âm thanh mà những nguồn âm thanh quan tâm (ví dụ các nhạc cụ và người hát chính) được quét đến trung tâm, việc trích xuất trung tâm có thể được áp dụng như bước tiền xử lý để làm suy giảm các âm thanh nền và tiếng vang.

Theo phương án, bộ tạo thông tin có thể được tạo cấu hình để tổ hợp thông tin tín hiệu và thông tin trộn giảm để thông tin tín hiệu trên trộn giảm cho biết tỷ lệ của thông tin tín hiệu với thông tin trộn giảm.

Trong phương án, bộ tạo thông tin có thể được tạo cấu hình để xử lý giá trị phổ của mỗi kênh đầu vào âm thanh trong số hai hoặc nhiều hơn hai kênh đầu vào âm thanh để thu được hai hoặc nhiều hơn hai giá trị được xử lý, và trong đó bộ tạo thông tin có thể được tạo cấu hình để tổ hợp hai hoặc nhiều hơn hai giá trị được xử lý để thu được thông tin tín hiệu. Hơn nữa, bộ tạo thông tin có thể được tạo cấu hình để tổ hợp giá trị phổ của mỗi kênh đầu vào âm thanh trong số hai hoặc nhiều hơn hai kênh đầu vào âm thanh để thu được giá trị kết hợp, và trong đó bộ tạo thông tin có thể được tạo cấu hình để xử lý giá trị kết hợp để thu được thông tin trộn giảm.

Theo phương án, bộ tạo thông tin có thể được tạo cấu hình để xử lý giá trị phổ của mỗi kênh đầu vào âm thanh trong số hai hoặc nhiều hơn hai kênh đầu vào âm thanh bằng cách nhân giá trị phổ đã nêu với liên hợp phức của giá trị phổ đã nêu để thu được mật độ phổ công suất tự động của giá trị phổ đã nêu cho mỗi kênh đầu vào âm thanh trong số hai hoặc nhiều hơn hai kênh đầu vào âm thanh.

Trong phương án, bộ tạo thông tin có thể được tạo cấu hình để xử lý giá trị kết hợp bằng cách xác định mật độ phổ công suất của giá trị kết hợp.

Theo phương án, bộ tạo thông tin có thể được tạo cấu hình để tạo ra thông tin tín hiệu $s(m, k, \beta)$ theo công thức:

$$s(m, k, \beta) = \sum_{i=1}^N \Phi_{i,i}(m, k)^{\beta},$$

trong đó N là số kênh đầu vào âm thanh của tín hiệu đầu vào âm thanh, trong đó $\Phi_{i,i}(m, k)$ là mật độ phổ công suất tự động của giá trị phổ của kênh tín hiệu âm thanh thứ i , trong đó β là số thực với $\beta > 0$, trong đó m là chỉ số thời gian, và trong đó k là chỉ số tần số. Ví dụ, theo phương án cụ thể $\beta \geq 1$.

Theo phương án, bộ tạo thông tin có thể được tạo cấu hình để xác định tỷ lệ tín hiệu trên trộn giảm như thông tin tín hiệu trên trộn giảm theo công thức $R(m, k, \beta)$

$$R(m, k, \beta) = \left(\frac{\sum_{i=1}^N \Phi_{i,i}(m, k)^\beta}{\Phi_d(m, k)^\beta} \right)^{\frac{1}{2\beta-1}},$$

trong đó $\Phi_d(m, k)$ là mật độ phổ công suất của giá trị kết hợp, và trong đó $\Phi_d(m, k)^\beta$ là thông tin trộn giảm.

Theo phương án, bộ tạo thông tin có thể được tạo cấu hình để tạo ra thông tin tín hiệu $\Phi_1(m, k)$ theo công thức

$$\Phi_1(m, k) = \mathcal{E}\{\mathbf{W}\mathbf{X}(m, k)(\mathbf{W}\mathbf{X}(m, k))^H\},$$

trong đó bộ tạo thông tin được tạo cấu hình để tạo ra thông tin trộn giảm $\Phi_2(m, k)$ theo công thức

$$\Phi_2(m, k) = \mathcal{E}\{\mathbf{V}\mathbf{X}(m, k)(\mathbf{V}\mathbf{X}(m, k))^H\}, \text{ và}$$

trong đó bộ tạo thông tin được tạo cấu hình để tạo ra tỷ lệ tín hiệu trên trộn giảm như thông tin tín hiệu trên trộn giảm $R(m, k, \beta)$ theo công thức

$$R_g(m, k, \beta) = \left(\frac{\text{tr}\{\Phi_1(m, k)^\beta\}}{\text{tr}\{\Phi_2(m, k)^\beta\}} \right)^{\frac{1}{2\beta-1}}$$

trong đó $\mathbf{X}(m, k)$ là tín hiệu đầu vào âm thanh, trong đó

$$\mathbf{X}(m, k) = [X_1(m, k) \cdots X_N(m, k)]^T$$

trong đó N là số kênh đầu vào âm thanh của tín hiệu đầu vào âm thanh, trong đó m là chỉ số thời gian, và trong đó k là chỉ số tần số, trong đó $X_1(m, k)$ là kênh đầu vào âm thanh thứ nhất, trong đó $X_N(m, k)$ là kênh đầu vào âm thanh thứ N , trong đó $\mathbf{V}$ là ma trận hoặc vector, trong đó $\mathbf{W}$ là ma trận hoặc vector, trong đó H là chuyển vị liên hợp của ma trận hoặc vector, trong đó $\mathcal{E}\{\cdot\}$ là phép toán kỳ vọng, trong đó β là số thực với $\beta > 0$, và trong đó $\text{tr}\{\cdot\}$ là vết của ma trận. Ví dụ, theo phương án cụ thể $\beta \geq 1$.

Trong phương án, $\mathbf{V}$ có thể là vector hàng có chiều dài N mà các phần tử của nó bằng một và $\mathbf{W}$ có thể là ma trận đơn vị có kích thước $N \times N$.

Theo phương án, $\mathbf{V} = [1, 1]$, trong đó $\mathbf{W} = [1, -1]$ và trong đó $N = 2$.

Trong phương án, bộ suy giảm tín hiệu có thể được làm thích ứng để làm suy giảm hai hoặc nhiều hơn hai kênh đầu vào âm thanh phụ thuộc vào hàm khuếch đại $G(m, k)$ theo công thức

$$\mathbf{Y}(m, k) = G(m, k)\mathbf{X}(m, k),$$

trong đó hàm khuếch đại $G(m, k)$ phụ thuộc vào thông tin tín hiệu trên trộn giảm, và trong đó hàm khuếch đại $G(m, k)$ là hàm tăng đơn điệu của thông tin tín hiệu trên trộn giảm hoặc hàm giảm đơn điệu của thông tin tín hiệu trên trộn giảm,

trong đó $\mathbf{X}(m, k)$ là tín hiệu đầu vào âm thanh, trong đó $\mathbf{Y}(m, k)$ là tín hiệu âm thanh được biến đổi, trong đó m là chỉ số thời gian, và trong đó k là chỉ số tần số.

Theo phương án, hàm khuếch đại $G(m, k)$ có thể là hàm thứ nhất $G_{c1}(m, k, \beta, \gamma)$, hàm thứ hai $G_{c2}(m, k, \beta, \gamma)$, hàm thứ ba $G_{s1}(m, k, \beta, \gamma)$ hoặc hàm thứ tư $G_{s2}(m, k, \beta, \gamma)$,

trong đó

$$G_{c1}(m, k, \beta, \gamma) = (1 + R_{\min} - R(m, k, \beta))^{\gamma},$$

trong đó

$$G_{c2}(m, k, \beta, \gamma) = \left(\frac{R_{\min}}{R(m, k, \beta)}\right)^{\gamma},$$

trong đó

$$G_{s1}(m, k, \beta, \gamma) = R(m, k, \beta)^{\gamma},$$

trong đó

$$G_{s2}(m, k, \beta, \gamma) = \left(1 + R_{\min} - \frac{R_{\min}}{R(m, k, \beta)}\right)^{\gamma},$$

trong đó β là số thực với $\beta > 0$,

trong đó γ là số thực với $\gamma > 0$, và

trong đó $R_{\min}$ là cực tiểu của R .

Hơn nữa, hệ thống được đề xuất. Hệ thống bao gồm bộ bù pha để tạo ra tín hiệu âm thanh được bù pha bao gồm hai hoặc nhiều hơn hai kênh âm thanh được bù pha từ tín hiệu âm thanh không được xử lý bao gồm hai hoặc nhiều hơn hai kênh âm thanh không được xử lý. Hơn nữa, hệ thống bao gồm thiết bị theo một trong số các phương án được mô tả ở trên để nhận tín hiệu âm thanh được bù pha như tín hiệu đầu vào âm thanh và để tạo ra tín hiệu âm thanh được biến đổi bao gồm hai hoặc nhiều hơn hai kênh âm thanh được biến đổi từ tín hiệu đầu vào âm thanh bao gồm hai hoặc nhiều hơn hai kênh âm thanh được bù pha như hai hoặc nhiều hơn hai kênh đầu vào âm thanh. Một trong số hai hoặc nhiều hơn hai kênh âm thanh không được xử lý là kênh tham chiếu. Bộ bù pha được làm thích ứng để ước lượng cho mỗi kênh âm thanh không được xử lý trong số hai hoặc nhiều hơn hai kênh âm thanh không được xử lý mà không là kênh tham chiếu hàm truyền pha giữa kênh âm thanh không được xử lý với kênh chuẩn đã nêu và kênh tham chiếu. Hơn nữa, bộ bù pha được làm thích ứng để tạo ra tín hiệu âm thanh được bù pha bằng cách biến đổi mỗi kênh âm thanh không được xử lý trong số các kênh âm thanh không được xử lý mà không là kênh tham chiếu phụ thuộc vào hàm truyền pha của kênh âm thanh không được xử lý đã nêu.

Hơn nữa, phương pháp tạo ra tín hiệu âm thanh được biến đổi bao gồm hai hoặc nhiều hơn hai kênh âm thanh được biến đổi từ tín hiệu đầu vào âm thanh bao gồm hai hoặc nhiều hơn hai kênh đầu vào âm thanh được cung cấp. Phương pháp bao gồm các bước:

- Tạo ra thông tin tín hiệu bằng cách tổ hợp giá trị phổ của mỗi kênh đầu vào âm thanh trong số hai hoặc nhiều hơn hai kênh đầu vào âm thanh theo cách thứ nhất.
- Tạo ra thông tin trộn giảm bằng cách tổ hợp giá trị phổ của mỗi kênh đầu vào âm thanh trong số hai hoặc nhiều hơn hai kênh đầu vào âm thanh theo cách thứ hai khác cách thứ nhất.
- Tạo ra thông tin tín hiệu trên trộn giảm bằng cách tổ hợp thông tin tín hiệu và thông tin trộn giảm. Và:

- Làm suy giảm hai hoặc nhiều hơn hai kênh đầu vào âm thanh phụ thuộc vào thông tin tín hiệu trên trộn giảm để thu được hai hoặc nhiều hơn hai kênh âm thanh được biến đổi.

Hơn nữa, vật ghi có thể đọc được bằng máy tính chứa chương trình máy tính để thực hiện phương pháp được mô tả ở trên khi được chạy trên máy tính hoặc bộ suy giảm tín hiệu được cung cấp.

Mô tả vắn tắt các hình vẽ

Tiếp theo, các phương án của sáng chế được mô tả chi tiết hơn khi tham chiếu đến các hình vẽ, trong đó:

Fig.1 minh họa thiết bị theo một phương án,

Fig.2 minh họa tỷ lệ tín hiệu trên trộn giảm như hàm của các chênh lệch mức độ liên kênh và như hàm của sự kết hợp liên kênh theo một phương án,

Fig.3 minh họa các trọng số phổ như hàm của sự kết hợp liên kênh và của các chênh lệch mức độ liên kênh theo phương án,

Fig.4 minh họa các trọng số phổ như hàm của sự kết hợp liên kênh và của các chênh lệch mức độ liên kênh theo phương án khác,

Fig.5 minh họa các trọng số phổ như hàm của sự kết hợp liên kênh và của các chênh lệch mức độ liên kênh theo phương án nữa,

Các hình vẽ từ Fig.6a đến Fig.6e minh họa các ảnh phổ của các tín hiệu nguồn thẳng và tín hiệu kênh bên trái và tín hiệu kênh bên phải của tín hiệu hỗn hợp,

Fig.7 minh họa tín hiệu đầu vào và tín hiệu đầu ra để trích xuất tín hiệu trung tâm theo phương án,

Fig.8 minh họa các ảnh phổ của tín hiệu đầu ra theo phương án,

Fig.9 minh họa tín hiệu đầu vào và tín hiệu đầu ra để làm suy giảm tín hiệu trung tâm theo phương án khác,

Fig.10 minh họa các ảnh phổ của tín hiệu đầu ra theo phương án,

Các hình vẽ từ Fig.11a đến Fig.11d minh họa hai tín hiệu tiếng nói mà được trộn để thu được các tín hiệu đầu vào có và không có các chênh lệch thời gian liên kênh,

Các hình vẽ từ Fig.12a đến Fig.12c minh họa các trọng số phổ được tính toán từ hàm khuếch đại theo phương án, và

Fig.13 minh họa hệ thống theo phương án.

Mô tả chi tiết sáng chế

Fig.1 minh họa thiết bị tạo ra tín hiệu âm thanh được biến đổi bao gồm hai hoặc nhiều hơn hai kênh âm thanh được biến đổi từ tín hiệu đầu vào âm thanh bao gồm hai hoặc nhiều hơn hai kênh đầu vào âm thanh theo phương án.

Thiết bị bao gồm bộ tạo thông tin 110 để tạo ra thông tin tín hiệu trên trộn giảm.

Bộ tạo thông tin 110 được làm thích ứng để tạo ra thông tin tín hiệu bằng cách tổ hợp giá trị phổ của mỗi kênh đầu vào âm thanh trong số hai hoặc nhiều hơn hai kênh đầu vào âm thanh theo cách thứ nhất. Hơn nữa, bộ tạo thông tin 110 được làm thích ứng để tạo ra thông tin trộn giảm bằng cách tổ hợp giá trị phổ của mỗi kênh đầu vào âm thanh trong số hai hoặc nhiều hơn hai kênh đầu vào âm thanh theo cách thứ hai mà khác so với cách thứ nhất.

Hơn nữa, bộ tạo thông tin 110 được làm thích ứng để tổ hợp thông tin tín hiệu và thông tin trộn giảm để thu được thông tin tín hiệu trên trộn giảm. Ví dụ, thông tin tín hiệu trên trộn giảm có thể là tỷ lệ tín hiệu trên trộn giảm, ví dụ giá trị tín hiệu trên trộn giảm.

Hơn nữa, thiết bị bao gồm bộ suy giảm tín hiệu 120 để làm suy giảm hai hoặc nhiều hơn hai kênh đầu vào âm thanh phụ thuộc vào thông tin tín hiệu trên trộn giảm để thu được hai hoặc nhiều hơn hai kênh âm thanh được biến đổi.

Theo phương án, bộ tạo thông tin có thể được tạo cấu hình để tổ hợp thông tin tín hiệu và thông tin trộn giảm để thông tin tín hiệu trên trộn giảm là tỷ lệ của thông tin tín hiệu với thông tin trộn giảm. Ví dụ, thông tin tín hiệu có thể là giá trị thứ nhất và thông tin trộn giảm có thể là giá trị thứ hai và thông tin tín hiệu trên trộn giảm là tỷ lệ của giá trị tín hiệu với giá trị trộn giảm. Ví dụ, thông tin tín hiệu trên trộn giảm có thể

là giá trị thứ nhất được suy ra từ giá trị thứ hai. Hoặc, ví dụ, nếu giá trị thứ nhất và giá trị thứ hai là các giá trị lôgarit, thông tin tín hiệu trên trộn giảm có thể là chênh lệch giữa giá trị thứ nhất với giá trị thứ hai.

Tiếp theo, mô hình tín hiệu và các khái niệm tín hiệu nằm phía dưới được mô tả và áp dụng đối với trường hợp mà tín hiệu đầu vào mô tả đặc trưng âm lập thể chênh lệch biên độ.

Lý do căn bản là để tính toán và áp dụng các trọng số phổ được định giá trị thực như hàm khuếch tán và vị trí bên của các nguồn thẳng. Việc xử lý như được thể hiện ở đây được áp dụng trong miền STFT, tuy nhiên nó không hạn chế với giàn lọc riêng biệt. Tín hiệu đầu vào kênh N được biểu thị bởi

$$\mathbf{x}[n] = [x_1[n] \cdots x_N[n]]^T. \quad (1)$$

trong đó n biểu thị chỉ số thời gian rời rạc. Tín hiệu đầu vào được giả định là hỗn hợp bổ sung của các tín hiệu thẳng $s_i[n]$ và các âm thanh ở môi trường xung quanh $a_i[n]$,

$$x_l[n] = \sum_{i=1}^K d_{i,l}[n] * s_i[n] + a_l[n], \quad l = 1, \dots, N \quad (2)$$

trong đó P là số nguồn âm thanh, $d_{i,l}[n]$ biểu thị các đáp ứng xung của những đường thẳng của nguồn thứ i vào trong kênh thứ l của chiều dài các mẫu $L_{i,l}$, và các thành phần tín hiệu ở môi trường xung quanh không tương quan lẫn nhau hoặc tương quan yếu. Trong mô tả sau đây mà được giả định là mô hình tín hiệu tương ứng với âm lập thể chênh lệch biên độ, tức là $L_{i,l} = 1, \forall i, l$. Biểu diễn miền thời gian-tần số của $\mathbf{x}[n]$ được xác định bởi

$$\mathbf{X}(m, k) = [X_1(m, k) \cdots X_N(m, k)]^T, \quad (3)$$

với chỉ số thời gian m và chỉ số tần số k . Các tín hiệu đầu ra được biểu thị bởi

$$\mathbf{Y}(m, k) = [Y_1(m, k) \cdots Y_N(m, k)]^T, \quad (4)$$

và thu được bằng cách gán trọng số phổ

$$\mathbf{Y}(m, k) = G(m, k)\mathbf{X}(m, k) , \quad (5)$$

với các trọng số được định giá trị thực $G(m, k)$. Các tín hiệu đầu ra miền thời gian được tính toán bằng cách áp dụng việc xử lý nghịch đảo của giàn lọc. Để tính toán các trọng số phổ, tín hiệu tổng, sau đó được biểu thị như tín hiệu trộn giảm, được tính toán là

$$X_d(m, k) = \sum_{i=1}^N X_i(m, k) , \quad (6)$$

Ma trận của PSD của tín hiệu đầu vào, bao gồm các ước lượng của PSD (-tự động) trên đường chéo chính, trong khi các phần tử ngoại chéo là các ước lượng của PSD-chéo, được xác định bởi

$$\Phi_{i,l}(m, k) = \mathcal{E}\{X_i(m, k)X_l^*(m, k)\} , \quad i, l = 1 \dots N , \quad (7)$$

trong đó X^* biểu thị liên hợp phức của X , và $\mathcal{E}\{\cdot\}$ là phép toán kỳ vọng liên quan đến phân đoạn thời gian. Trong các mô phỏng đã có các giá trị kỳ vọng được ước lượng sử dụng việc lấy trung bình đệ quy đơn cực.

$$\Phi_{i,l}(m, k) = \alpha X_i(m, k)X_l^*(m, k) + (1 - \alpha)\Phi_{i,l}(m - 1, k) , \quad (8)$$

trong đó hệ số bộ lọc α xác định thời gian tích hợp. Hơn nữa, số lượng $R(m, k; \beta)$ được chỉ rõ là

$$R(m, k, \beta) = \left(\frac{\sum_{i=1}^N \Phi_{i,i}(m, k)^\beta}{\Phi_d(m, k)^\beta} \right)^{\frac{1}{2\beta-1}} . \quad (9)$$

trong đó $\Phi_d(m, k)$ là PSD của tín hiệu trộn giảm và β là tham số mà sẽ được đề cập sau đây. Số lượng $R(m, k; 1)$ là tỷ lệ tín hiệu trên trộn giảm (SDR), tức là tỷ lệ của PSD tổng với PSD của tín hiệu trộn giảm. Lũy thừa bằng $\frac{1}{2\beta-1}$ đảm bảo rằng khoảng $R(m, k; \beta)$ là không phụ thuộc vào β .

Bộ tạo thông tin 110 có thể được tạo cấu hình để xác định tỷ lệ tín hiệu trên trộn giảm theo Phương trình (9).

Theo phương trình (9) thông tin tín hiệu $s(m, k, \beta)$ mà có thể được xác định bởi bộ tạo thông tin 110 được chỉ rõ là

$$s(m, k, \beta) = \sum_{i=1}^N \Phi_{i,i}(m, k)^\beta,$$

Như có thể nhìn thấy ở trên, $\Phi_{i,i}(m, k)$ được chỉ rõ là $\Phi_{i,i}(m, k) = \mathcal{E}\{X_i(m, k)X_i^*(m, k)\}$. Do đó, để xác định thông tin tín hiệu $s(m, k, \beta)$, giá trị phổ $X_i(m, k)$ của mỗi kênh đầu vào âm thanh trong số hai hoặc nhiều hơn hai kênh đầu vào âm thanh được xử lý để thu được giá trị được xử lý $\Phi_{i,i}(m, k)^\beta$ cho mỗi kênh đầu vào âm thanh trong số hai hoặc nhiều hơn hai kênh đầu vào âm thanh, và các giá trị được xử lý đã thu được $\Phi_{i,i}(m, k)^\beta$ sau đó kết hợp, ví dụ như trong Phương trình (9) bằng cách cộng các giá trị được xử lý đã thu được $\Phi_{i,i}(m, k)^\beta$.

Do đó, bộ tạo thông tin 110 có thể được tạo cấu hình để xử lý giá trị phổ $X_i(m, k)$ của mỗi kênh đầu vào âm thanh trong số hai hoặc nhiều hơn hai kênh đầu vào âm thanh để thu được hai hoặc nhiều hơn hai giá trị được xử lý $\Phi_{i,i}(m, k)^\beta$, và bộ tạo thông tin 110 có thể được tạo cấu hình để tổ hợp hai hoặc nhiều hơn hai giá trị được xử lý để thu được thông tin tín hiệu $s(m, k, \beta)$. Khái quát hơn, bộ tạo thông tin 110 được làm thích ứng để tạo ra thông tin tín hiệu $s(m, k, \beta)$ bằng cách tổ hợp giá trị phổ $X_i(m, k)$ của mỗi kênh đầu vào âm thanh trong số hai hoặc nhiều hơn hai kênh đầu vào âm thanh theo cách thứ nhất.

Hơn nữa, theo Phương trình (9) thông tin trộn giảm $d(m, k, \beta)$ mà có thể được xác định bởi bộ tạo thông tin 110 được chỉ rõ là

$$s(m, k, \beta) = \Phi_d(m, k)^\beta,$$

Để tạo thành $\Phi_d(m, k)$, đầu tiên $X_d(m, k)$ được tạo thành theo Phương trình (6) ở trên:

$$X_d(m, k) = \sum_{i=1}^N X_i(m, k)$$

Như có thể thấy, thứ nhất, giá trị phổ $X_i(m,k)$ của mỗi kênh đầu vào âm thanh trong số hai hoặc nhiều hơn hai kênh đầu vào âm thanh kết hợp để thu được giá trị kết hợp $X_d(m,k)$, ví dụ, như trong Phương trình (6), bằng cách cộng giá trị phổ $X_i(m,k)$ của mỗi kênh đầu vào âm thanh trong số hai hoặc nhiều hơn hai kênh đầu vào âm thanh.

Sau đó, để thu được $\Phi_d(m,k)$, mật độ phổ công suất của $X_d(m,k)$ được tạo thành, ví dụ theo

$$\Phi_d(m,k) = \mathcal{E} \{ X_d(m,k) X_d^*(m,k) \},$$

và sau đó, $\Phi_d(m,k)^\beta$ có thể được xác định. Nói một cách khái quát hơn, giá trị kết hợp thu được $X_d(m,k)$ được xử lý để thu được thông tin trộn giảm $d(m, k, \beta) = \Phi_d(m,k)^\beta$.

Do đó, bộ tạo thông tin 110 có thể được tạo cấu hình để tổ hợp giá trị phổ $X_i(m,k)$ của mỗi kênh đầu vào âm thanh trong số hai hoặc nhiều hơn hai kênh đầu vào âm thanh để thu được giá trị kết hợp, và bộ tạo thông tin 110 có thể được tạo cấu hình để xử lý giá trị kết hợp để thu được thông tin trộn giảm $d(m, k, \beta)$. Khái quát hơn, bộ tạo thông tin 110 được làm thích ứng để tạo ra thông tin trộn giảm $d(m, k, \beta)$ bằng cách tổ hợp giá trị phổ $X_i(m,k)$ của mỗi kênh đầu vào âm thanh trong số hai hoặc nhiều hơn hai kênh đầu vào âm thanh theo cách thứ hai. Cách mà thông tin trộn giảm được tạo ra ("cách thứ hai") không giống với cách, mà thông tin tín hiệu được tạo ra ("cách thứ nhất") và do đó, cách thứ hai không giống cách thứ nhất.

Bộ tạo thông tin 110 được làm thích ứng để tạo ra thông tin tín hiệu bằng cách tổ hợp giá trị phổ của mỗi kênh đầu vào âm thanh trong số hai hoặc nhiều hơn hai kênh đầu vào âm thanh theo cách thứ nhất. Hơn nữa, bộ tạo thông tin 110 được làm thích ứng để tạo ra thông tin trộn giảm bằng cách tổ hợp giá trị phổ của mỗi kênh đầu vào âm thanh trong số hai hoặc nhiều hơn hai kênh đầu vào âm thanh theo cách thứ hai mà khác so với cách thứ nhất.

Fig.2, đồ thị ở trên minh họa tỷ lệ tín hiệu trên trộn giảm $R(m, k; 1)$ đối với $N=2$ như hàm của CLD $\Theta(m, k)$, được thể hiện cho $\Psi(m, k) \in \{0, 0,2, 0,4, 0,6, 0,8, 1\}$. Fig.2, đồ thị ở dưới minh họa tỷ lệ tín hiệu trên trộn giảm $R(m, k; 1)$ đối với $N=2$ như hàm của ICC $\Psi(m, k)$ và ICLD $\Theta(m, k)$ trong đồ thị 2 chiều được mã hóa màu sắc.

Cụ thể, Fig.2 minh họa SDR đối với $N=2$ như hàm của ICC $\Psi(m, k)$ và ICLD $\Theta(m, k)$, với

$$\Psi(m, k) = \frac{|\Phi_{1,2}(m, k)|}{\sqrt{\Phi_{1,1}(m, k)\Phi_{2,2}(m, k)}}, \quad (10)$$

và

$$\Theta(m, k) = \frac{\Phi_{1,1}(m, k)}{\Phi_{2,2}(m, k)}. \quad (11)$$

Fig.2 thể hiện rằng SDR có các đặc tính như sau:

1. Nó đơn điệu đối với cả hai, $\Psi(m, k)$ và $|\log \Theta(m, k)|$.
2. Đối với các tín hiệu đầu vào khuếch tán, tức là $\Psi(m, k)=0$, giả sử giá trị cực đại của SDR, $R(m, k; 1) = 1$.
3. Đối với các âm thanh thẳng được quét đến trung tâm, tức là $\Theta(m, k) = 1$, giả sử giá trị cực tiểu của SDR là $R_{\min}$, trong đó $R_{\min} = 0,5$ đối với $N=2$.

Vì những đặc tính này, các trọng số phổ gần đúng để định tỷ lệ tín hiệu trung tâm có thể được tính toán từ SDR bằng cách sử dụng các hàm giảm đơn điệu để trích xuất các tín hiệu trung tâm và các hàm tăng đơn điệu để làm suy giảm các tín hiệu trung tâm.

Để trích xuất tín hiệu trung tâm, các hàm gần đúng của $R(m, k; \beta)$ là, ví dụ,

$$G_{c1}(m, k, \beta, \gamma) = (1 + R_{\min} - R(m, k, \beta))^{\gamma}, \quad (12)$$

, và

$$G_{c2}(m, k, \beta, \gamma) = \left(\frac{R_{\min}}{R(m, k, \beta)}\right)^{\gamma}. \quad (13)$$

trong đó tham số để kiểm soát việc làm suy giảm tối đa được đưa vào.

Để làm suy giảm tín hiệu trung tâm, các hàm gần đúng của $R(m, k; \beta)$ là, ví dụ,

$$G_{s1}(m, k, \beta, \gamma) = R(m, k, \beta)^\gamma, \quad (14)$$

và

$$G_{s2}(m, k, \beta, \gamma) = (1 + R_{\min} - \frac{R_{\min}}{R(m, k, \beta)})^\gamma, \quad (15)$$

Fig.3 và Fig.4 lần lượt minh họa các hàm khuếch đại (13) và (15), đối với $\beta = 1$, $\gamma = 3$. Các trọng số phổ là không đổi đối với $\Psi(m, k) = 0$. Việc làm suy giảm tối đa là $\gamma \cdot 6\text{dB}$, mà cũng áp dụng với các hàm khuếch đại (12) và (14).

Cụ thể, Fig.3 minh họa các trọng số phổ $G_{c2}(m, k; 1, 3)$ trong dB (deciben) như hàm của ICC $\Psi(m, k)$ và ICLD $\Theta(m, k)$.

Hơn nữa, Fig.4 minh họa các trọng số phổ $G_{s2}(m, k; 1, 3)$ trong dB như hàm của ICC $\Psi(m, k)$ và ICLD $\Theta(m, k)$.

Hơn nữa, Fig.5 minh họa các trọng số phổ $G_{c2}(m, k; 2, 3)$ trong dB như hàm của ICC $\Psi(m, k)$ và ICLD $\Theta(m, k)$.

Tác dụng của tham số β được thể hiện trên Fig.5 đối với hàm khuếch đại trong Phương trình (13) với $\beta = 2$, $\gamma = 3$. Với những giá trị lớn hơn đối với β , ảnh hưởng của Ψ đối với các trọng số phổ giảm đi trong khi ảnh hưởng của Θ tăng lên. Việc này dẫn đến các thành phần tín hiệu khuếch tán rò rỉ nhiều hơn vào trong tín hiệu đầu ra, và dẫn đến làm suy giảm nhiều hơn các thành phần tín hiệu thẳng được quét lệch tâm, khi so sánh với hàm khuếch đại trên Fig.3.

Xử lý sau các trọng số phổ: Trước khi gán trọng số phổ, các trọng số $G(m, k; \beta, \gamma)$ có thể còn được xử lý bằng các hoạt động tron tru. Việc lọc thông thấp pha không theo trục tần số làm giảm các thành phần lạ tích chập vòng mà có thể xảy ra ví dụ khi đệm-không trong phép tính STFT quá ngắn hoặc cửa sổ tổng hợp hình chữ nhật được áp dụng. Lọc thông-thấp theo trục thời gian có thể làm giảm các thành phần lạ xử lý, đặc biệt khi hằng số thời gian để ước lượng PSD là khá nhỏ.

Tiếp theo, các trọng số phổ suy rộng được cung cấp.

Các trọng số phổ tổng quát hơn thu được khi viết lại Phương trình (9) là

$$R_g(m, k, \beta) = \left(\frac{\text{tr}\{\Phi_1(m, k)^\beta\}}{\text{tr}\{\Phi_2(m, k)^\beta\}} \right)^{\frac{1}{2\beta-1}}. \quad (16)$$

với

$$\Phi_1(m, k) = \mathcal{E}\{\mathbf{W}\mathbf{X}(m, k)(\mathbf{W}\mathbf{X}(m, k))^H\} \quad (17)$$

$$\Phi_2(m, k) = \mathcal{E}\{\mathbf{V}\mathbf{X}(m, k)(\mathbf{V}\mathbf{X}(m, k))^H\} \quad (18)$$

trong đó số mũ H biểu thị chuyển vị liên hợp của ma trận hoặc vector, và $\mathbf{W}$ và $\mathbf{V}$ là các ma trận tròn hoặc các vector (hàng) tròn.

Ở đây, $\Phi_1(m, k)$ có thể được xem như thông tin tín hiệu và $\Phi_2(m, k)$ có thể được xem như thông tin trộn giảm.

Ví dụ, $\Phi_2 = \Phi_d$ khi $\mathbf{V}$ là vector có chiều dài N mà các phần tử của nó bằng một. Phương trình (16) bằng phương trình (9) khi $\mathbf{V}$ là vector hàng có chiều dài N mà các phần tử của nó bằng một và $\mathbf{W}$ là ma trận đơn vị có kích thước $N \times N$.

SDR suy rộng $R_g(m, k, \beta, \mathbf{W}, \mathbf{V})$ bao trùm ví dụ tỷ lệ của PSD của tín hiệu phụ với PSD của tín hiệu trộn giảm, đối với $\mathbf{W} = [1, -1]$, $\mathbf{V} = [1, 1]$, và $N = 2$.

$$R(m, k, \beta) = \left(\frac{\Phi_s(m, k)^\beta}{\Phi_d(m, k)^\beta} \right)^{\frac{1}{2\beta-1}} \quad (19)$$

trong đó $\Phi_s(m, k)$ là PSD của tín hiệu phụ.

Theo phương án, bộ tạo thông tin 110 được làm thích ứng để tạo ra thông tin tín hiệu $\Phi_1(m, k)$ bằng cách tổ hợp giá trị phổ $X_i(m, k)$ của mỗi kênh đầu vào âm thanh trong số hai hoặc nhiều hơn hai kênh đầu vào âm thanh theo cách thứ nhất. Hơn nữa, bộ tạo thông tin 110 được làm thích ứng để tạo ra thông tin trộn giảm $\Phi_2(m, k)$ bằng cách tổ hợp giá trị phổ $X_i(m, k)$ của mỗi kênh đầu vào âm thanh trong số hai hoặc nhiều hơn hai kênh đầu vào âm thanh theo cách thứ hai mà khác so với cách thứ nhất.

Tiếp theo, trường hợp khái quát hơn của các mô hình trộn mô tả đặc trưng âm lập thể thời gian-đến được mô tả.

Việc suy ra các trọng số phổ được mô tả ở trên liên quan đến giả định là $L_{i,l} = 1$, $\forall i, l$, tức là các nguồn âm thanh thẳng được căn chỉnh thời gian giữa các kênh đầu vào. Khi việc trộn các tín hiệu nguồn thẳng không bị hạn chế với âm lập thể chênh lệch biên độ ($L_{i,l} > 1$), ví dụ khi ghi chép với các micrô cách quãng, sự trộn giảm của tín hiệu đầu vào $X_d(m, k)$ là mục đích để triệt tiêu pha. Việc triệt tiêu pha trong $X_d(m, k)$ dẫn đến làm tăng các giá trị SDR và do đó làm tăng các thành phần lọc kiểu răng lược điển hình khi áp dụng việc gán trọng số phổ như được mô tả ở trên.

Các rãnh của bộ lọc kiểu răng lược tương ứng với các tần số

$$f_n = \frac{of_s}{2d}$$

đối với các hàm khuếch đại (12) và (13) và

$$f_n = \frac{ef_s}{2d}$$

đối với các hàm khuếch đại (14) và (15), trong đó f_s là tần số lấy mẫu, o là các số nguyên lẻ, e là các số nguyên chẵn, và d là độ trễ của các mẫu.

Cách tiếp cận thứ nhất để giải quyết vấn đề này là bù các chênh lệch pha thu được từ ICTD trước khi tính toán $f X_d(m, k)$. Việc bù chênh lệch pha (PDC) đạt được bằng cách ước lượng hàm truyền pha liên kênh thời gian-biến thiên $\hat{P}_i(m, k)$ $\hat{P}_i(m, k) \in [-\pi \pi]$ giữa kênh thứ i với kênh tham chiếu được biểu thị bởi chỉ số r ,

$$\hat{P}_i(m, k) = \arg X_r(m, k) - \arg X_i(m, k) , \quad i \in [1, \dots, N] \setminus r \quad (20)$$

trong đó bộ hoạt động $A \setminus B$ biểu thị chênh lệch tập hợp-lý thuyết của tập B với tập A , và áp dụng bộ lọc bù thông toàn bộ thời gian-biến thiên $H_{C,i}(m, k)$ với tín hiệu kênh thứ i .

$$\tilde{X}_i(m, k) = H_{C,i}(m, k) X_i(m, k) . \quad (21)$$

trong đó hàm truyền pha của $H_{C,i}(m, k)$ là

$$\arg H_{C,i}(m, k) = -\mathcal{E}\{\hat{P}_i(m, k)\} . \quad (22)$$

Giá trị kỳ vọng được ước lượng sử dụng việc lấy trung bình đệ quy đơn cực. Lưu ý rằng các bước nhảy pha của 2π xảy ra ở những tần số gần với các tần số rãnh cần được bù trước khi lấy trung bình đệ quy.

Tín hiệu trộn giảm được tính toán theo

$$X_d(m, k) = \sum_{i=1}^N \tilde{X}_i(m, k) . \quad (23)$$

sao cho PDC chỉ được áp dụng để tính toán X_d và không ảnh hưởng đến pha của tín hiệu đầu ra.

Fig.13 minh họa hệ thống theo phương án.

Hệ thống bao gồm bộ bù pha 210 để tạo ra tín hiệu âm thanh được bù pha bao gồm hai hoặc nhiều hơn hai kênh âm thanh được bù pha từ tín hiệu âm thanh không được xử lý bao gồm hai hoặc nhiều hơn hai kênh âm thanh không được xử lý.

Hơn nữa, hệ thống bao gồm thiết bị 220 theo một trong số các phương án được mô tả ở trên để nhận tín hiệu âm thanh được bù pha như tín hiệu đầu vào âm thanh và để tạo ra tín hiệu âm thanh được biến đổi bao gồm hai hoặc nhiều hơn hai kênh âm thanh được biến đổi từ tín hiệu đầu vào âm thanh bao gồm hai hoặc nhiều hơn hai kênh âm thanh được bù pha như hai hoặc nhiều hơn hai kênh đầu vào âm thanh.

Một trong số hai hoặc nhiều hơn hai kênh âm thanh không được xử lý là kênh tham chiếu. Bộ bù pha 210 được làm thích ứng để ước lượng cho mỗi kênh âm thanh không được xử lý của hai hoặc nhiều hơn hai kênh âm thanh không được xử lý mà không là kênh tham chiếu hàm truyền pha giữa kênh âm thanh không được xử lý với kênh tham chiếu đã nêu. Hơn nữa, bộ bù pha 210 được làm thích ứng để tạo ra tín hiệu âm thanh được bù pha bằng cách biến đổi mỗi kênh âm thanh không được xử lý trong số các kênh âm thanh không được xử lý mà không là kênh tham chiếu phụ thuộc vào hàm truyền pha của kênh âm thanh không được xử lý đã nêu.

Tiếp theo, những lý giải theo trực giác của các tham số điều chỉnh được cung cấp, ví dụ ý nghĩa về mặt ngữ nghĩa của các tham số điều chỉnh.

Đối với hoạt động của các tác dụng âm thanh dạng số là lợi thế để cung cấp các điều chỉnh với các tham số có nghĩa về mặt ngữ nghĩa. Các hàm khuếch đại từ (12)

đến (15) được điều chỉnh bởi các tham số α , β và γ . Kỹ sư âm và kỹ sư âm thanh được sử dụng với các hằng số thời gian, và chỉ rõ α là hằng số thời gian là theo trục giác và theo thực tiễn chung. Tác dụng của thời gian tích hợp có thể được hiểu tốt nhất bằng thử nghiệm. Để hỗ trợ cho hoạt động của các khái niệm, các bộ mô tả cho các tham số còn lại được đề xuất, cụ thể tham số tác động biểu diễn bởi γ và tham số khuếch tán biểu diễn bởi β .

tham số tác động có thể được so sánh tốt nhất theo thứ tự của bộ lọc. Tương tự với hàm phổ dốc trong lọc, sự suy giảm tối đa bằng $\gamma \cdot 6\text{dB}$, đối với $N = 2$.

Sự khuếch tán nhận được đề xuất ở đây để nhấn mạnh thực tế là sau khi việc làm suy giảm được quét và các âm thanh khuếch tán, các giá trị lớn hơn của β dẫn đến việc rò rỉ các âm khuếch tán nhiều hơn. Ánh xạ phi tuyến của tham số người dùng β_u , ví dụ $\beta = \sqrt{\beta_u + 1}$, với $0 \leq \beta_u \leq 10$, là có thuận lợi theo cách mà nó có thể tạo trạng thái nhất quán hơn cho việc xử lý trái với khi biến đổi β ngay lập tức (trong đó sự nhất quán liên quan đến tác dụng của việc thay đổi tham số trên kết quả xuyên suốt khoảng giá trị tham số).

Tiếp theo, sự phức tạp tính toán và những yêu cầu về bộ nhớ được mô tả một cách ngắn gọn.

Sự phức tạp tính toán và các yêu cầu về bộ nhớ chia tỷ lệ với số băng của giàn lọc và phụ thuộc vào việc thực hiện xử lý-sau bổ sung các trọng số phổ. Việc thực hiện phương pháp với chi phí thấp có thể đạt được khi đặt $\beta = 1$, $\gamma \in \mathbb{N}$, tính toán các trọng số phổ theo Phương trình (12) hoặc (14), và khi không áp dụng bộ lọc PDC. Việc tính toán của SDR chỉ sử dụng một hàm phi tuyến nghịch đảo cho mỗi băng khi $\beta \in \mathbb{N}$. Đối với $\beta = 1$, chỉ hai bộ đệm để ước lượng PSD được yêu cầu, trong khi các phương pháp tiên hành sử dụng một cách hoàn hảo ICC, ví dụ trong tài liệu tham khảo [7, 10, 20, 21, 23], yêu cầu ít nhất ba bộ đệm.

Tiếp theo, việc thực hiện các khái niệm đã có được mô tả bằng các ví dụ.

Thứ nhất, việc xử lý được áp dụng với hỗn hợp được quét-biên độ của 5 bản ghi nhạc cụ (tiếng trống, tiếng trầm, các âm điệu, 2 đàn ghita) được thử ở 44100 Hz thuộc đoạn trích có độ dài 3 giây được hiển thị hóa. Tiếng trống, tiếng trầm và các âm điệu

được quét đến trung tâm, một đàn ghita được quét đến kênh bên trái và đàn ghita thứ hai được quét đến kênh bên phải, cả hai có $|ICLD| = 20\text{dB}$. Buồng vang tích chập có các đáp ứng xung âm lập thể với RT60 khoảng 1,4 giây cho mỗi kênh đầu vào được sử dụng để tạo ra các thành phần tín hiệu ở môi trường xung quanh. Tín hiệu phản xạ được bổ sung tỷ lệ thẳng-đến-môi trường xung quanh bằng khoảng 8dB sau khi gán trọng số K trong tài liệu tham khảo [29].

Các hình vẽ từ Fig.6a đến Fig.6e minh họa các ảnh phổ của các tín hiệu nguồn thẳng và tín hiệu kênh bên trái và tín hiệu kênh bên phải của tín hiệu hỗn hợp. Các ảnh phổ được tính toán sử dụng STFT với chiều dài 2048 mẫu, sự chồng lấp 50%, kích thước khung là 1024 mẫu và cửa sổ hàm sin. Lưu ý là đối với mục tiêu chỉ làm rõ độ lớn của các hệ số phổ tương ứng với các tần số lên đến 4 kHz được hiển thị. Cụ thể, Fig.6a-e minh họa các tín hiệu đầu vào cho ví dụ âm nhạc.

Cụ thể, các hình vẽ từ Fig.6a đến Fig.6e minh họa trên Fig.6a là các tín hiệu nguồn trong đó các tiếng trống, tiếng trầm và các âm điệu được quét đến trung tâm, trên Fig.6b là các tín hiệu nguồn, trong đó đàn ghita 1, trong hỗn hợp được quét về bên trái, trên Fig.6c là các tín hiệu nguồn trong đó đàn ghita 2, trong hỗn hợp được quét về bên phải, trên Fig.6d là kênh bên trái của tín hiệu hỗn hợp, và trên Fig.6e là các kênh bên phải của tín hiệu hỗn hợp.

Fig.7 thể hiện tín hiệu đầu vào và tín hiệu đầu ra để lấy ra tín hiệu trung tâm được thu bằng cách áp dụng $G_{c2}(m, k; 1, 3)$. Cụ thể, Fig.7 là ví dụ về việc lấy ra trung tâm, trong đó các tín hiệu thời gian đầu vào (màu đen) và các tín hiệu thời gian đầu ra (được phủ màu xám) được minh họa, trong đó Fig.7, đồ thị ở trên minh họa kênh bên trái, và trong đó Fig.7, đồ thị ở dưới minh họa kênh bên phải.

Hằng số thời gian để lấy trung bình đệ quy trong ước lượng PSD ở đây và sau đó được đặt là 200 mili giây.

Fig.8 minh họa các ảnh phổ của tín hiệu đầu ra. Việc quan sát bằng mắt cho biết các tín hiệu nguồn được quét lệch tâm (được thể hiện trên Fig.6b và Fig.6c) phần lớn bị làm suy giảm trong các ảnh phổ đầu ra. Cụ thể, Fig.8 minh họa ví dụ về việc lấy ra trung tâm, các ảnh phổ đặc biệt hơn của các tín hiệu đầu ra. Các ảnh phổ đầu ra cũng thể hiện rằng các thành phần tín hiệu ở môi trường xung quanh bị làm suy giảm.

Fig.9 thể hiện tín hiệu đầu vào và tín hiệu đầu ra để làm suy giảm tín hiệu trung tâm thu được bằng cách áp dụng $G_{s2}(m, k; 1, 3)$. Các tín hiệu thời gian minh họa rằng các âm thanh nhất thời từ các tiếng trống bị làm suy giảm bằng cách xử lý. Cụ thể, Fig.9 minh họa ví dụ về việc làm suy giảm trung tâm, trong đó các tín hiệu thời gian đầu vào (màu đen) và các tín hiệu thời gian đầu ra (được phủ màu xám) được minh họa.

Fig.10 minh họa các ảnh phổ của tín hiệu đầu ra. Có thể quan sát thấy các tín hiệu được quét đến trung tâm được làm suy giảm, ví dụ khi quan sát các thành phần âm thanh nhất thời và các âm được duy trì liên tục trong phạm vi tần số thấp hơn dưới 600Hz và so sánh với Fig.6a. Các âm thanh nổi bật trong tín hiệu đầu ra tương ứng với các dụng cụ và tiếng vang được quét lệch tâm. Cụ thể, Fig.10 minh họa ví dụ về việc làm suy giảm trung tâm, cụ thể hơn, các ảnh phổ của các tín hiệu đầu ra.

Việc nghe không chính thức thông qua các tai nghe bộc lộ rằng sự suy giảm của các thành phần tín hiệu là hữu hiệu. Khi nghe với tín hiệu trung tâm được lấy ra, các thành phần lạ xử lý trở thành có thể nghe được như các biến điệu nhẹ trong các nốt của đàn ghita 2, tương tự như bơm trong nén dải động. Có thể lưu ý rằng tiếng vang bị làm giảm và sự suy giảm đó hữu hiệu với các tần số ngắn hơn là so với các tần số dài. Dù nó bị gây ra bởi tỷ lệ thăng đến môi trường xung quanh lớn hơn trong các tần số thấp hơn, phổ tần suất của các nguồn âm thanh hoặc cảm nhận chủ quan do hiện tượng không chắc có thể không thể được đáp lại mà không có sự phân tích tỷ mỉ hơn.

Khi nghe tín hiệu đầu ra trong đó trung tâm bị làm suy giảm, chất lượng âm tổng cộng tốt hơn không đáng kể khi so sánh với kết quả lấy ra trung tâm. Các thành phần lạ xử lý có thể nghe được như các dịch chuyển không đáng kể của các nguồn được quét hướng đến trung tâm khi các nguồn được định tâm ưu thế là hoạt động, tương đương với việc bơm khi lấy ra trung tâm. Các âm thanh tín hiệu đầu ra ít thăng hơn là kết quả của số lượng môi trường xung quanh trong tín hiệu đầu ra được tăng lên.

Để minh họa việc lọc PDC, các hình vẽ từ Fig.11a đến Fig.11d thể hiện hai tín hiệu tiếng nói mà được trộn để thu được các tín hiệu đầu vào có và không có ICTD. Cụ thể, các hình vẽ từ Fig.11a đến Fig.11d minh họa các tín hiệu nguồn đầu vào để

minh họa PDC, trong đó Fig.11a minh họa tín hiệu nguồn 1; trong đó Fig.11b minh họa tín hiệu nguồn 2; trong đó Fig.11c minh họa kênh bên trái của tín hiệu hỗn hợp; và trong đó Fig.11d minh họa kênh bên phải của tín hiệu hỗn hợp.

Tín hiệu hỗn hợp hai kênh được tạo ra bằng cách trộn các tín hiệu nguồn tiếng nói với các khuếch đại bằng nhau cho mỗi kênh và bằng cách bổ sung nhiễu trắng với SNR là 10 dB (được gán trọng số-K) với tín hiệu.

Các hình vẽ từ Fig.12a đến Fig.12c thể hiện các trọng số phổ được tính toán từ hàm khuếch đại (13). Cụ thể, các hình vẽ từ Fig.12a đến Fig.12c minh họa các trọng số phổ $G_{c2}(m, k; 1, 3)$ để làm rõ việc lọc PDC, trong đó Fig.12a minh họa các trọng số phổ cho các tín hiệu đầu vào mà không có ICTD, PDC không có hiệu lực; Fig.12b minh họa các trọng số phổ cho các tín hiệu đầu vào mà có ICTD, PDC không có hiệu lực; và Fig.12c minh họa các trọng số phổ cho các tín hiệu đầu vào mà có ICTD, PDC có hiệu lực.

Các trọng số phổ trong đồ thị ở trên gần bằng 0dB khi tiếng nói là hoạt động và giả định là giá trị tối thiểu trong các dải thời gian-tần số với SNR thấp. Đồ thị thứ hai thể hiện các trọng số phổ cho tín hiệu đầu vào trong đó tín hiệu tiếng nói thứ nhất (Fig.11a) được trộn với ICTD có 26 mẫu. Các đặc trưng bộ lọc kiểu răng lược được minh họa trên Fig.12b. Fig.12c thể hiện các trọng số phổ khi PDC có hiệu lực. Các thành phần lạ lọc kiểu răng lược phần lớn bị giảm, mặc dù việc bù không giống hoàn toàn các tần số rãnh ở 848Hz và 2544Hz.

Việc nghe không chính thức thể hiện rằng nhiều âm cộng phần lớn bị suy giảm. Khi các tín hiệu xử lý không có ICTD, các tín hiệu đầu ra có bit thuộc đặc trưng âm thanh ở môi trường xung quanh mà có lẽ thu được từ sự không nhất quán pha được đưa vào bởi nhiễu âm cộng. Khi các tín hiệu xử lý có ICTD, tín hiệu tiếng nói thứ nhất (Fig.11a) phần lớn bị suy giảm và các thành phần lạ lọc- kiểu răng lược mạnh có thể nghe thấy khi không áp dụng lọc PDC. Với việc lọc PDC bổ sung, các thành phần lạ lọc kiểu răng lược vẫn có thể nghe thấy không đáng kể, nhưng gây khó chịu nhiều hơn. Việc nghe không chính thức vật chất khác bộc lộ các thành phần lạ yếu, mà có thể bị làm giảm hoặc bằng cách giảm γ , bằng cách tăng β , hoặc bằng cách bổ sung kiểu được định tỷ lệ cho tín hiệu đầu vào không được xử lý đến đầu ra. Nói chung, các thành

phần lạ có thể nghe thấy ít hơn khi làm suy giảm tín hiệu trung tâm và có thể nghe thấy nhiều hơn khi lấy ra tín hiệu trung tâm. Các biến dạng của hình ảnh không gian có trước là rất nhỏ. Điều này có thể được cho là do thực tế các trọng số phổ là đồng nhất cho tất cả các tín hiệu kênh và không tác động đến các ICLD. Các thành phần lạ lọc kiểu răng lược hầu như không nghe được khi xử lý các ghi chép tự nhiên mô tả đặc trưng âm lập thể thời gian-đến cho trộn giảm đơn không là đối tượng để có thể nghe thấy rõ ràng các thành phần lạ lọc kiểu răng lược. Đối với việc lọc PDC có thể lưu ý rằng các giá trị hằng số thời gian nhỏ của việc lấy trung bình đệ quy (cụ thể việc bù tức thời các chênh lệch pha khi tính toán X_d) đưa tương quan vào các tín hiệu được sử dụng để trộn giảm. Tiếp theo, việc xử lý là bất khả thi liên quan đến sự khuếch tán của tín hiệu đầu vào. Khi hằng số thời gian được tăng lên, nó có thể được quan sát là (1) hiệu quả của PDC đối với các tín hiệu đầu vào với âm lập thể chênh lệch biên độ giảm xuống và (2) hiệu quả lọc kiểu răng lược trở thành có thể nghe thấy nhiều hơn ở các điểm bắt đầu của nốt âm khi các nguồn âm thẳng được căn chỉnh-thời gian giữa các kênh đầu vào.

Các khái niệm về định tỷ lệ tín hiệu trung tâm trong các ghi chép âm thanh bằng cách áp dụng các trọng số phổ được định giá trị thực mà được tính toán từ các hàm đơn điệu của SDR được cung cấp. Lý do căn bản là việc định tỷ lệ tín hiệu trung tâm cần đưa vào trong cả hai phép tính, dịch chuyển ngang của các nguồn thẳng và lượng khuếch tán, và các đặc tính này được thu hoàn toàn bởi SDR. Việc xử lý có thể được điều chỉnh bởi các tham số người dùng có nghĩa về mặt ngữ nghĩa và so sánh với các kỹ thuật miền tần số khác của phức hợp tính toán và tải bộ nhớ ít. Các khái niệm được đề xuất mang đến các kết quả tốt khi các tín hiệu đầu vào xử lý mô tả đặc trưng âm lập thể chênh lệch biên độ, nhưng có thể tùy thuộc vào các thành phần lạ lọc kiểu răng lược khi các nguồn âm thanh thẳng không được căn chỉnh-thời gian giữa các kênh đầu vào. Cách tiếp cận thứ nhất để giải quyết đó là để bù cho pha khác không trong hàm truyền liên kênh.

Từ trước đến nay, các khái niệm thuộc các phương án được thử nghiệm bằng cách nghe không chính thức. Đối với các ghi chép có tính thương mại điển hình, các kết quả là chất lượng âm thanh tốt nhưng cũng phụ thuộc vào lực phân tách mong muốn.

Mặc dù một số khía cạnh đã được mô tả trong bối cảnh thiết bị, rõ ràng là các khía cạnh này cũng thể hiện sự mô tả cho phương pháp tương ứng, trong đó khối hoặc thiết bị tương ứng với bước thuộc phương pháp hoặc đặc điểm của bước thuộc phương pháp. Tương tự, các khía cạnh khác được mô tả trong nội dung của bước thuộc phương pháp cũng thể hiện sự mô tả của khối hoặc phần hoặc đặc điểm tương ứng của thiết bị tương ứng.

Tín hiệu đã được tách ra theo sáng chế có thể được lưu giữ trên vật ghi lưu giữ dạng số hoặc có thể được truyền trên vật ghi truyền dẫn như phương tiện truyền dẫn không dây hoặc phương tiện truyền dẫn có dây như mạng Internet.

Phụ thuộc vào các yêu cầu thực hiện nhất định, các phương án theo sáng chế có thể được thực hiện dưới dạng phần cứng hoặc phần mềm. Phương án có thể được thực hiện bằng cách sử dụng phương tiện lưu trữ dạng số, ví dụ, đĩa mềm, DVD, CD, ROM, PROM, EPROM, EEPROM hoặc bộ nhớ FLASH, có các tín hiệu điều chỉnh 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 thể lập trình được sao cho phương pháp tương ứng được thực hiện.

Một số phương án theo sáng chế bao gồm vật mang dữ liệu chuyển tiếp có các tín hiệu điều chỉnh 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 các phương pháp được mô tả ở đây được thực hiện.

Thông thường, các phương án theo sáng chế có thể được thực hiện dưới dạng 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 giữ trên vật mang có thể đọc được bằng máy.

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

Do đó, phương án khác của phương pháp theo sáng chế là luồng dữ liệu hoặc chuỗi các tín hiệu đại diện cho chương trình máy tính để thực hiện một trong số các phương pháp được mô tả ở đây. Luồng dữ liệu hoặc chuỗi các 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 mạng Internet.

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

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

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

Các phương án được mô tả ở trên chỉ đơn thuần minh họa cho các nguyên lý của sáng chế. Cần hiểu rằng các sự biến thể và biến đổi của các phương án và các chi tiết được mô tả ở đây sẽ là hiển nhiên với người có trình độ trong lĩnh vực kỹ thuật tương ứng. Do đó, mục đích 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. Thiết bị tạo ra tín hiệu âm thanh được biến đổi bao gồm hai hoặc nhiều hơn hai kênh âm thanh được biến đổi từ tín hiệu đầu vào âm thanh bao gồm hai hoặc nhiều hơn hai kênh đầu vào âm thanh, trong đó thiết bị bao gồm:

bộ tạo thông tin (110) để tạo ra thông tin tín hiệu trên trộn giảm, trong đó bộ tạo thông tin (110) được làm thích ứng để tạo ra thông tin tín hiệu bằng cách kết hợp giá trị phổ của mỗi kênh đầu vào âm thanh trong số hai hoặc nhiều hơn hai kênh đầu vào âm thanh theo cách thứ nhất, trong đó bộ tạo thông tin (110) được làm thích ứng để tạo ra thông tin trộn giảm bằng cách kết hợp giá trị phổ của mỗi kênh đầu vào âm thanh trong số hai hoặc nhiều hơn hai kênh đầu vào âm thanh theo cách thứ hai mà khác so với cách thứ nhất, và trong đó bộ tạo thông tin (110) được làm thích ứng để kết hợp thông tin tín hiệu và thông tin trộn giảm để thu được thông tin tín hiệu trên trộn giảm, và

bộ suy giảm tín hiệu (120) để làm suy giảm hai hoặc nhiều hơn hai kênh đầu vào âm thanh phụ thuộc vào thông tin tín hiệu trên trộn giảm để thu được hai hoặc nhiều hơn hai kênh âm thanh được biến đổi,

trong đó bộ tạo thông tin (110) được tạo cấu hình để xử lý giá trị phổ của mỗi kênh đầu vào âm thanh trong số hai hoặc nhiều hơn hai kênh đầu vào âm thanh bằng cách nhân giá trị phổ đã nêu với liên hợp phức của giá trị phổ đã nêu để thu được mật độ phổ công suất tự động của giá trị phổ đã nêu cho mỗi kênh đầu vào âm thanh trong số hai hoặc nhiều hơn hai kênh đầu vào âm thanh,

trong đó bộ tạo thông tin (110) được tạo cấu hình để kết hợp giá trị phổ của mỗi kênh đầu vào âm thanh trong số hai hoặc nhiều hơn hai kênh đầu vào âm thanh để thu được giá trị kết hợp, và trong đó bộ tạo thông tin (110) được tạo cấu hình để xử lý giá trị kết hợp bằng cách xác định mật độ phổ công suất của giá trị kết hợp,

đặc trưng ở chỗ, bộ tạo thông tin (110) được tạo cấu hình để tạo thông tin tín hiệu $s(m, k, \beta)$ theo công thức:

$$s(m, k, \beta) = \sum_{i=1}^N \Phi_{i,i}(m, k)^\beta,$$

trong đó N biểu thị số lượng các kênh đầu vào âm thanh của tín hiệu đầu vào âm thanh,

trong đó $\Phi_{i,i}(m, k)$ biểu thị mật độ phổ công suất tự động của giá trị phổ của kênh tín hiệu âm thanh thứ i ,

trong đó β là số thực với $\beta > 0$,

trong đó m biểu thị chỉ số thời gian, và trong đó k biểu thị chỉ số tần số.

2. Thiết bị theo điểm 1, trong đó bộ tạo thông tin (110) được tạo cấu hình để kết hợp thông tin tín hiệu và thông tin trộn giảm sao cho thông tin tín hiệu trên trộn giảm chỉ ra tỷ lệ của thông tin tín hiệu trên thông tin trộn giảm.

3. Thiết bị theo điểm 1 hoặc điểm 2, trong đó số kênh âm thanh được biến đổi bằng số kênh đầu vào âm thanh, hoặc trong đó số kênh âm thanh được biến đổi nhỏ hơn số kênh đầu vào âm thanh.

4. Thiết bị theo một điểm trong số các điểm nêu trên,

trong đó bộ tạo thông tin (110) được tạo cấu hình để xử lý giá trị phổ của mỗi kênh đầu vào âm thanh trong số hai hoặc nhiều hơn hai kênh đầu vào âm thanh để thu được hai hoặc nhiều hơn hai giá trị được xử lý, và trong đó bộ tạo thông tin (110) được tạo cấu hình để kết hợp hai hoặc nhiều hơn hai giá trị được xử lý để thu được thông tin tín hiệu, và

trong đó bộ tạo thông tin (110) được tạo cấu hình để xử lý giá trị kết hợp để thu được thông tin trộn giảm.

5. Thiết bị theo một điểm trong số các điểm nêu trên,

trong đó bộ tạo thông tin (110) được tạo cấu hình để xác định tỷ lệ tín hiệu trên trộn giảm làm thông tin tín hiệu trên trộn giảm theo công thức $R(m, k, \beta)$

$$R(m, k, \beta) = \left(\frac{\sum_{i=1}^N \Phi_{i,i}(m, k)^\beta}{\Phi_d(m, k)^\beta} \right)^{\frac{1}{2\beta-1}},$$

trong đó $\Phi_d(m, k)$ biểu thị mật độ phổ công suất của giá trị kết hợp, và

trong đó $\Phi_d(m, k)^\beta$ là thông tin trộn giảm.

6. Thiết bị theo điểm 5, trong đó bộ suy giảm tín hiệu (120) được làm thích ứng để làm suy giảm hai hoặc nhiều hơn hai kênh đầu vào âm thanh tùy thuộc vào hàm khuếch đại $G(m, k)$ theo công thức

$$\mathbf{Y}(m, k) = G(m, k)\mathbf{X}(m, k),$$

trong đó hàm khuếch đại $G(m, k)$ phụ thuộc vào thông tin tín hiệu trên trộn giảm, và trong đó hàm khuếch đại $G(m, k)$ là hàm tăng đơn điệu của thông tin tín hiệu trên trộn giảm hoặc là hàm giảm đơn điệu của thông tin tín hiệu trên trộn giảm,

trong đó $\mathbf{X}(m, k)$ biểu thị tín hiệu đầu vào âm thanh,

trong đó $\mathbf{Y}(m, k)$ biểu thị tín hiệu âm thanh được biến đổi,

trong đó m biểu thị chỉ số thời gian, và

trong đó k biểu thị chỉ số tần số.

7. Thiết bị theo điểm 6,

trong đó hàm khuếch đại $G(m, k)$ là hàm thứ nhất $G_{c1}(m, k, \beta, \gamma)$, hàm thứ hai $G_{c2}(m, k, \beta, \gamma)$, hàm thứ ba $G_{s1}(m, k, \beta, \gamma)$ hoặc hàm thứ tư $G_{s2}(m, k, \beta, \gamma)$,

trong đó

$$G_{c1}(m, k, \beta, \gamma) = (1 + R_{\min} - R(m, k, \beta))^\gamma,$$

trong đó

$$G_{c2}(m, k, \beta, \gamma) = \left(\frac{R_{\min}}{R(m, k, \beta)} \right)^\gamma,$$

trong đó

$$G_{s1}(m, k, \beta, \gamma) = R(m, k, \beta)^\gamma,$$

trong đó

$$G_{s2}(m, k, \beta, \gamma) = \left(1 + R_{\min} - \frac{R_{\min}}{R(m, k, \beta)} \right)^\gamma,$$

trong đó β là số thực với $\beta > 0$,

trong đó γ là số thực với $\gamma > 0$, và

trong đó $R_{\min}$ là cực tiểu của R .

8. Hệ thống tạo ra tín hiệu âm thanh được biến đổi, trong đó hệ thống bao gồm:

bộ bù pha (210) để tạo ra tín hiệu âm thanh được bù pha bao gồm hai hoặc nhiều hơn hai kênh âm thanh được bù pha từ tín hiệu âm thanh không được xử lý bao gồm hai hoặc nhiều hơn hai kênh âm thanh không được xử lý, và

thiết bị (220) theo một điểm trong số các điểm trên để nhận tín hiệu âm thanh được bù pha làm tín hiệu đầu vào âm thanh và để tạo ra tín hiệu âm thanh được biến đổi bao gồm hai hoặc nhiều hơn hai kênh âm thanh được biến đổi từ tín hiệu đầu vào âm thanh bao gồm hai hoặc nhiều hơn hai kênh âm thanh được bù pha làm hai hoặc nhiều hơn hai kênh đầu vào âm thanh,

trong đó một trong số hai hoặc nhiều hơn hai kênh âm thanh không được xử lý là kênh tham chiếu,

trong đó bộ bù pha (210) được làm thích ứng để ước lượng cho mỗi kênh âm thanh không được xử lý trong số hai hoặc nhiều hơn hai kênh âm thanh không được xử lý mà không là kênh tham chiếu hàm chuyển pha giữa kênh âm thanh không được xử lý và kênh tham chiếu đã nêu, và

trong đó bộ bù pha (210) được làm thích ứng để tạo ra tín hiệu âm thanh được bù pha bằng cách biến đổi mỗi kênh âm thanh không được xử lý trong số các kênh âm thanh không được xử lý mà không là kênh tham chiếu phụ thuộc vào hàm chuyển pha của kênh âm thanh không được xử lý đã nêu.

9. Phương pháp tạo ra tín hiệu âm thanh được biến đổi bao gồm hai hoặc nhiều hơn hai kênh âm thanh được biến đổi từ tín hiệu đầu vào âm thanh bao gồm hai hoặc nhiều hơn hai kênh đầu vào âm thanh, trong đó phương pháp bao gồm các bước:

tạo ra thông tin tín hiệu bằng cách kết hợp giá trị phổ của mỗi kênh đầu vào âm thanh trong số hai hoặc nhiều hơn hai kênh đầu vào âm thanh theo cách thứ nhất,

tạo ra thông tin trộn giảm bằng cách kết hợp giá trị phổ của mỗi kênh đầu vào âm thanh trong số hai hoặc nhiều hơn hai kênh đầu vào âm thanh theo cách thứ hai khác cách thứ nhất,

tạo ra thông tin tín hiệu trên trộn giảm bằng cách kết hợp thông tin tín hiệu và thông tin trộn giảm, và

làm suy giảm hai hoặc nhiều hơn hai kênh đầu vào âm thanh tùy thuộc vào thông tin tín hiệu trên trộn giảm để thu được hai hoặc nhiều hơn hai kênh âm thanh được biến đổi,

trong đó phương pháp còn bao gồm các bước:

xử lý giá trị phổ của mỗi kênh đầu vào âm thanh trong số hai hoặc nhiều hơn hai kênh đầu vào âm thanh bằng cách nhân giá trị phổ đã nêu với liên hợp phức của giá trị phổ đã nêu để thu được mật độ phổ công suất tự động của giá trị phổ đã nêu cho mỗi kênh đầu vào âm thanh trong số hai hoặc nhiều hơn hai kênh đầu vào âm thanh,

kết hợp giá trị phổ của mỗi kênh đầu vào âm thanh trong số hai hoặc nhiều hơn hai kênh đầu vào âm thanh để thu được giá trị kết hợp, và xử lý giá trị kết hợp bằng cách xác định mật độ phổ công suất của giá trị kết hợp,

phương pháp được đặc trưng bởi việc tạo thông tin tín hiệu $s(m, k, \beta)$ theo công thức:

$$s(m, k, \beta) = \sum_{i=1}^N \Phi_{i,i}(m, k)^\beta,$$

trong đó N biểu thị số lượng các kênh đầu vào âm thanh của tín hiệu đầu vào âm thanh,

trong đó $\Phi_{i,i}(m, k)$ biểu thị mật độ phổ công suất tự động của giá trị phổ của kênh tín hiệu âm thanh thứ i ,

trong đó β là số thực với $\beta > 0$,

trong đó m biểu thị chỉ số thời gian, và trong đó k biểu thị chỉ số tần số.

10. Vật ghi đọc được bằng máy tính chứa chương trình máy tính để thực hiện phương pháp theo điểm 9 khi được chạy trên máy tính hoặc bộ xử lý tín hiệu.

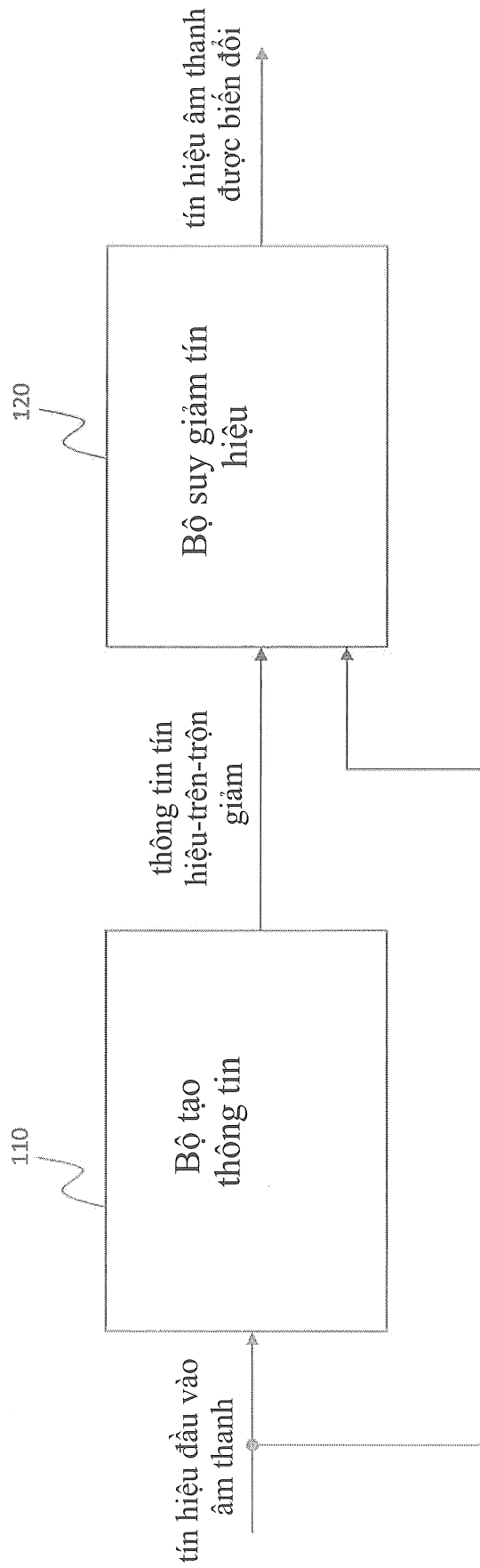


FIG 1

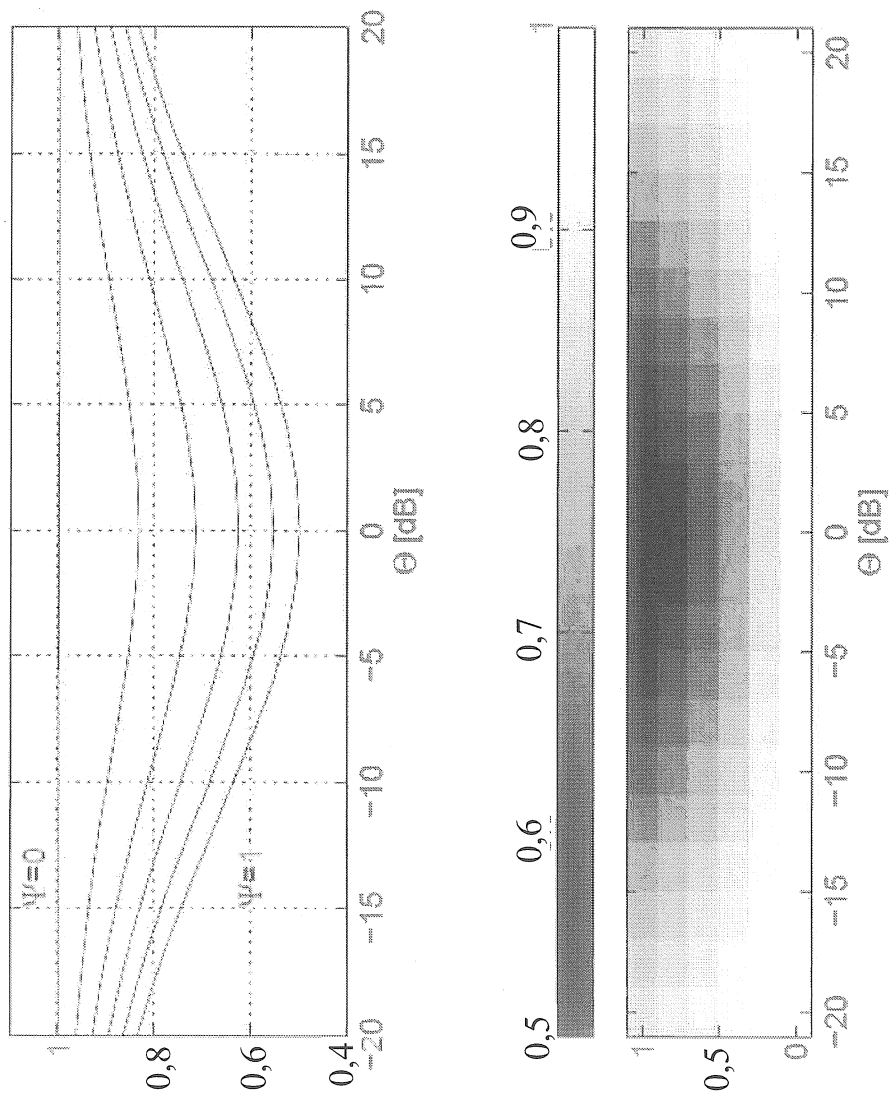


FIG 2

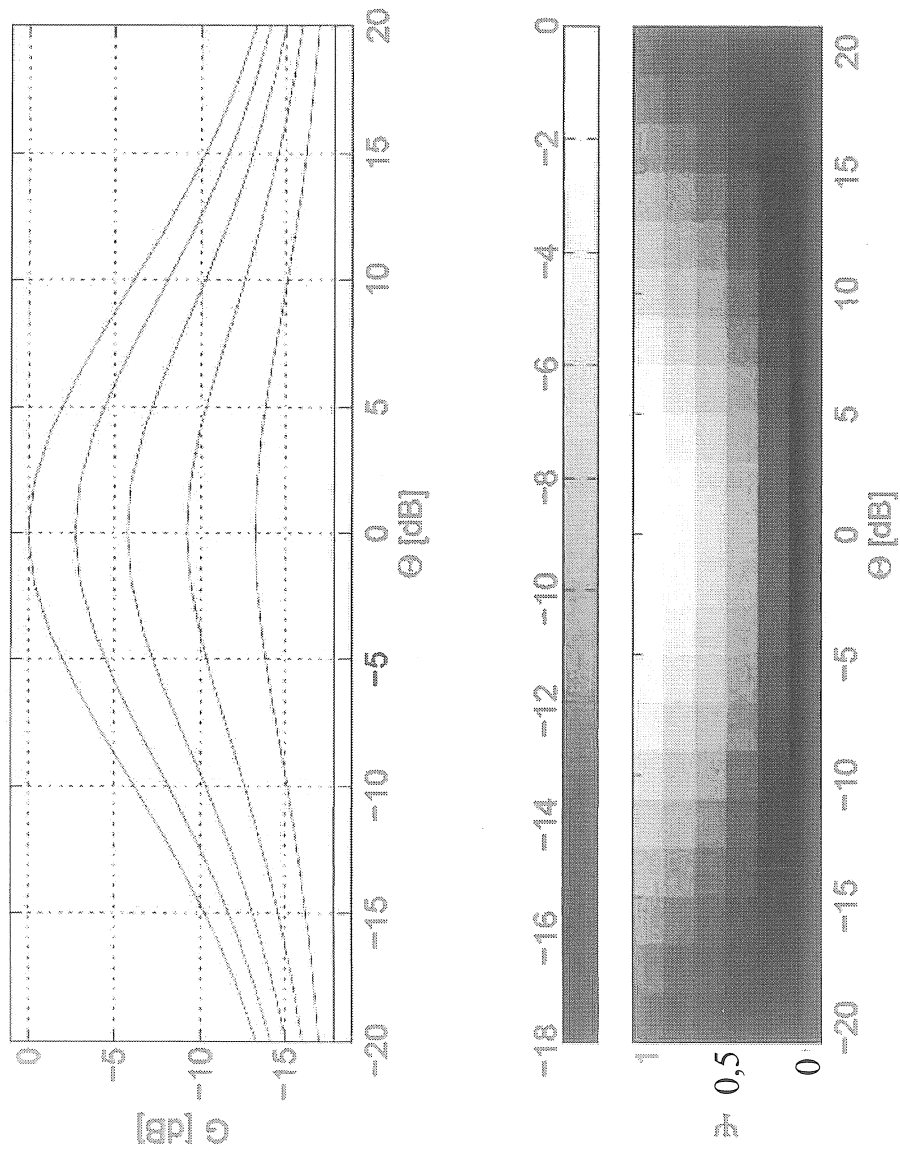


FIG 3

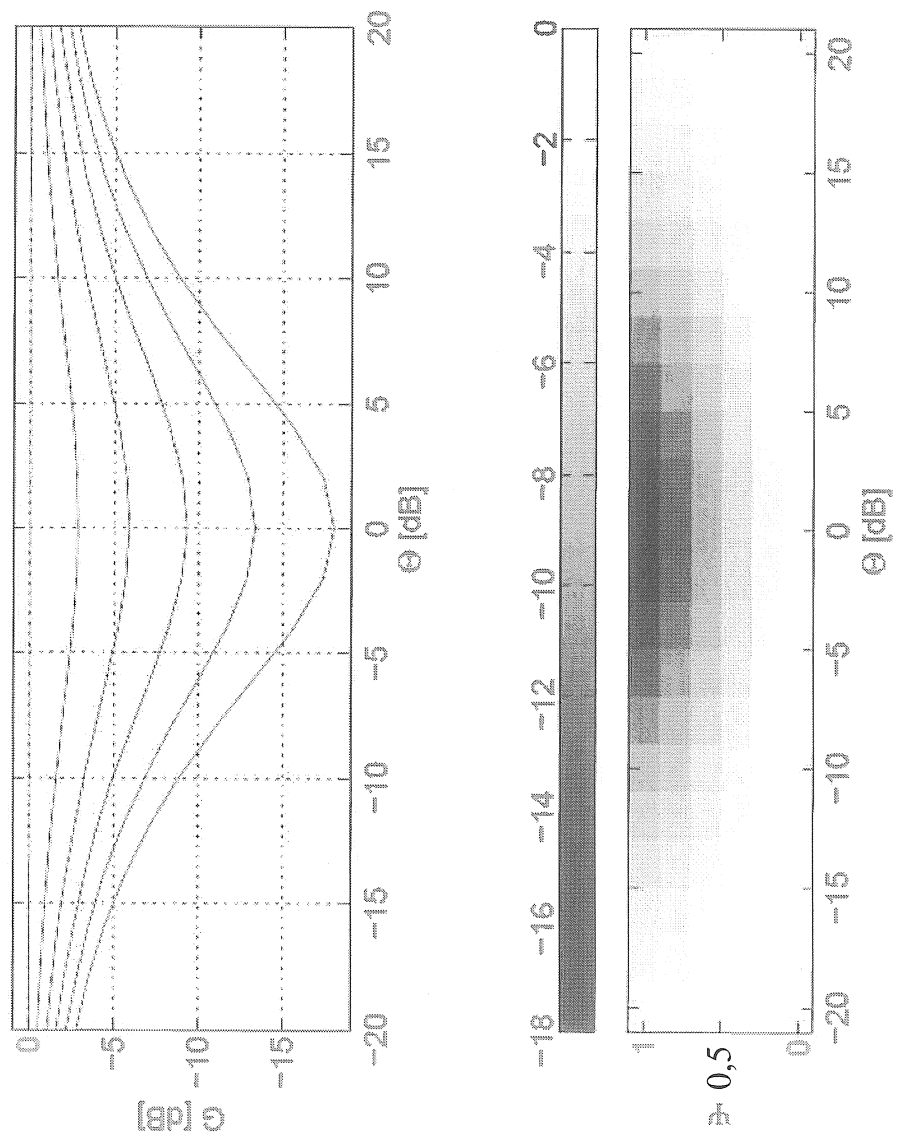


FIG 4

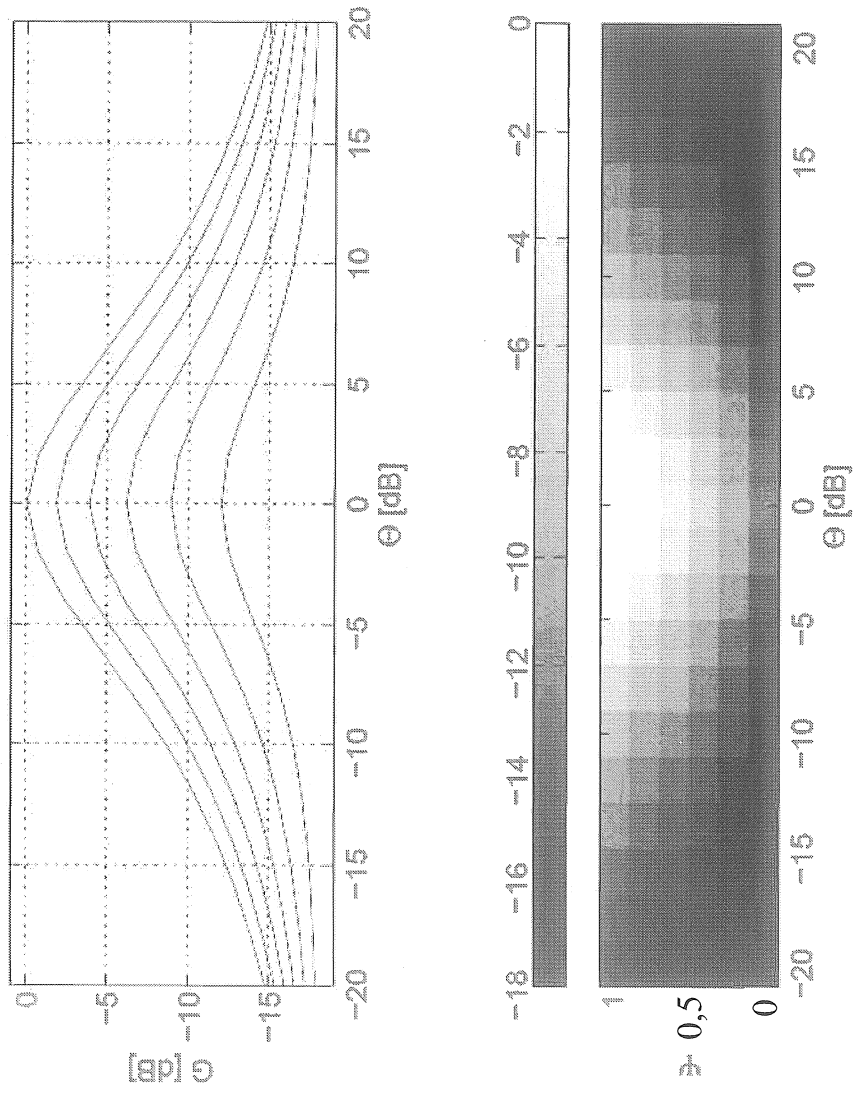


FIG 5

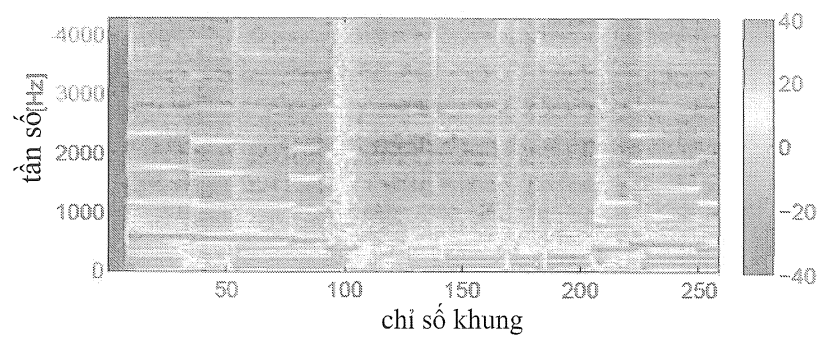


FIG 6A

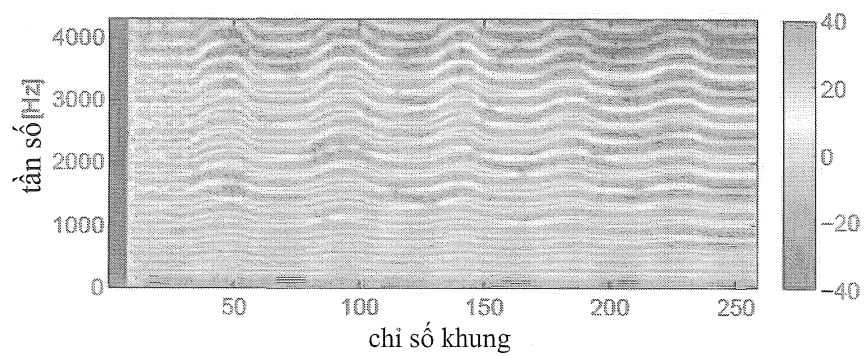


FIG 6B

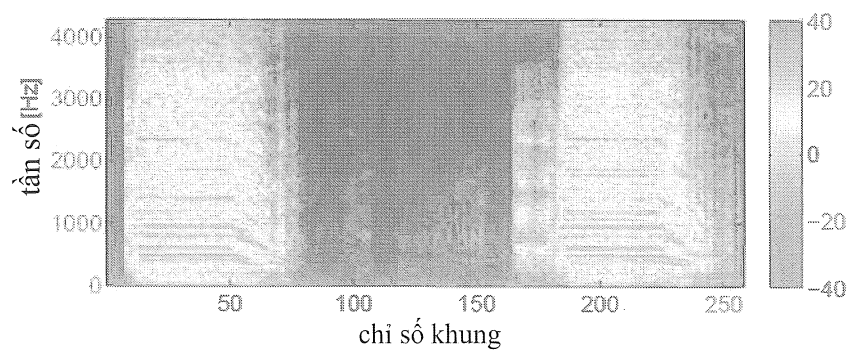


FIG 6C

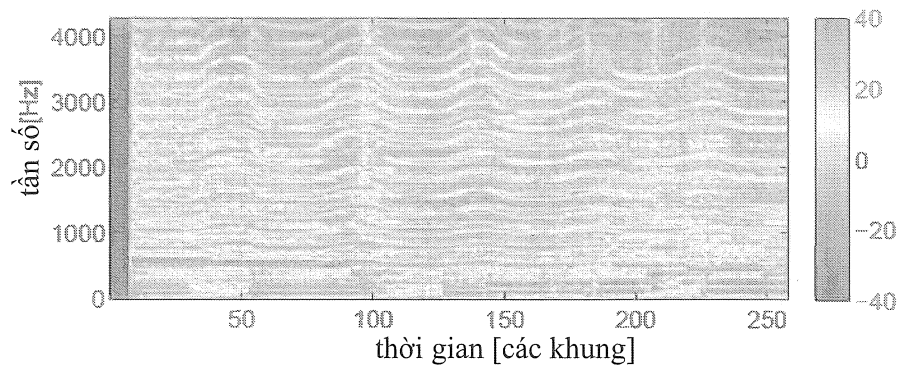


FIG 6D

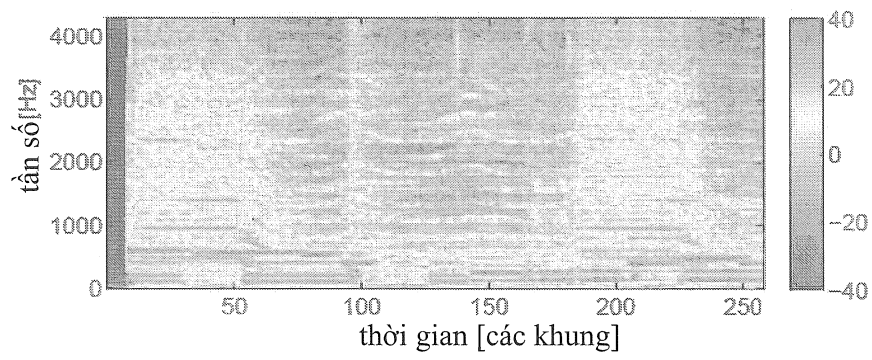


FIG 6E

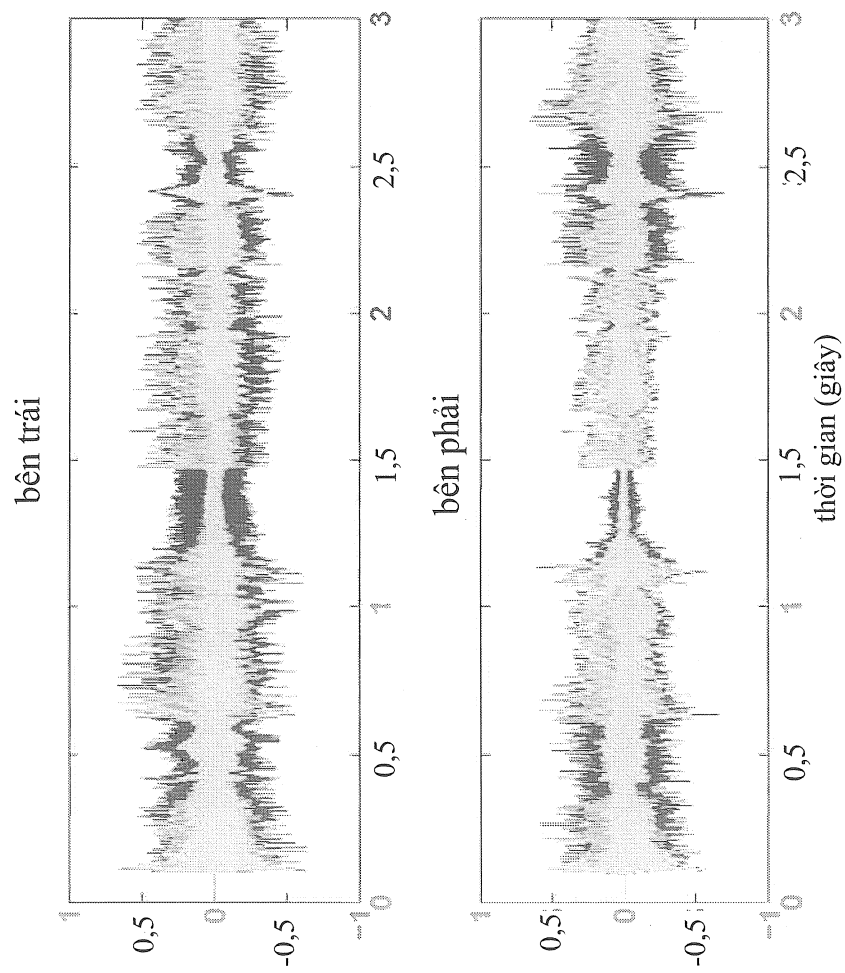


FIG 7

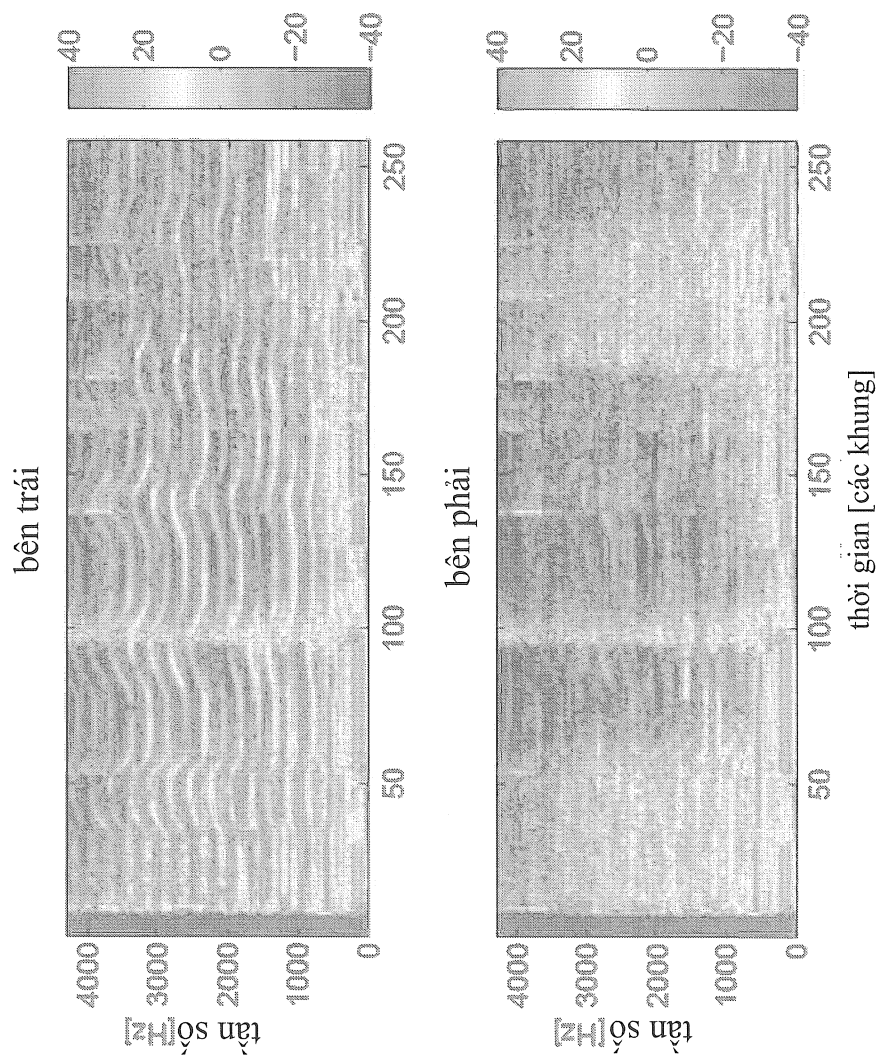


FIG 8

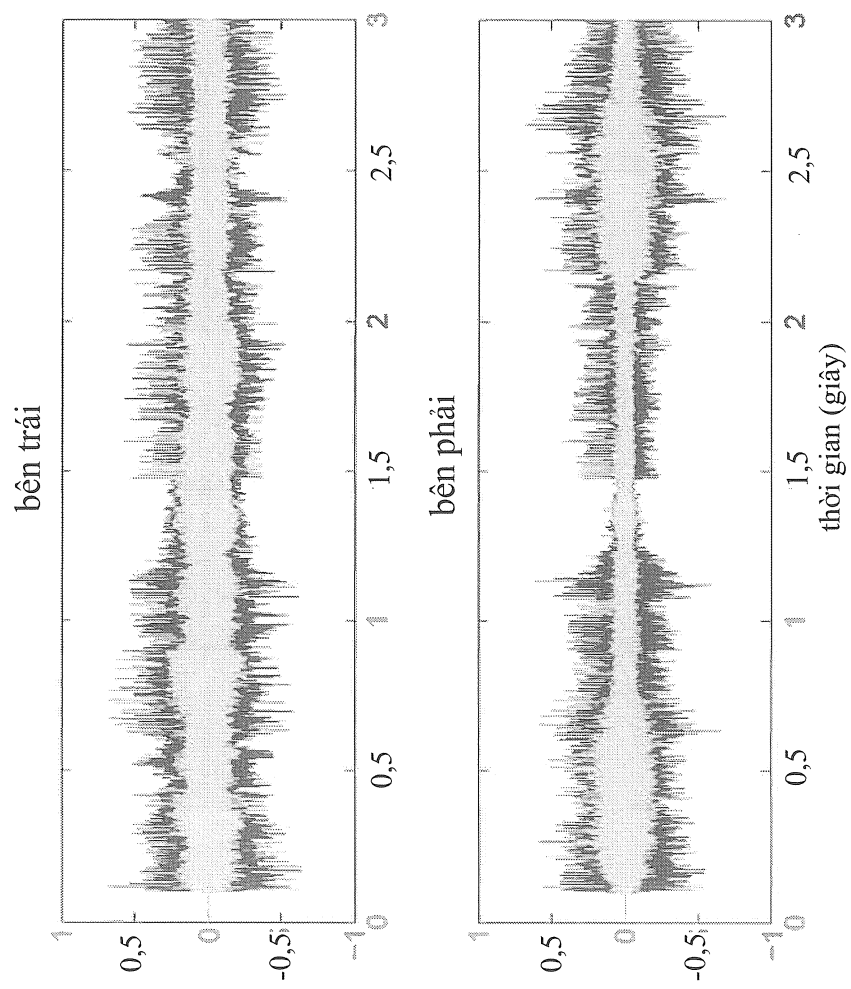


FIG 9

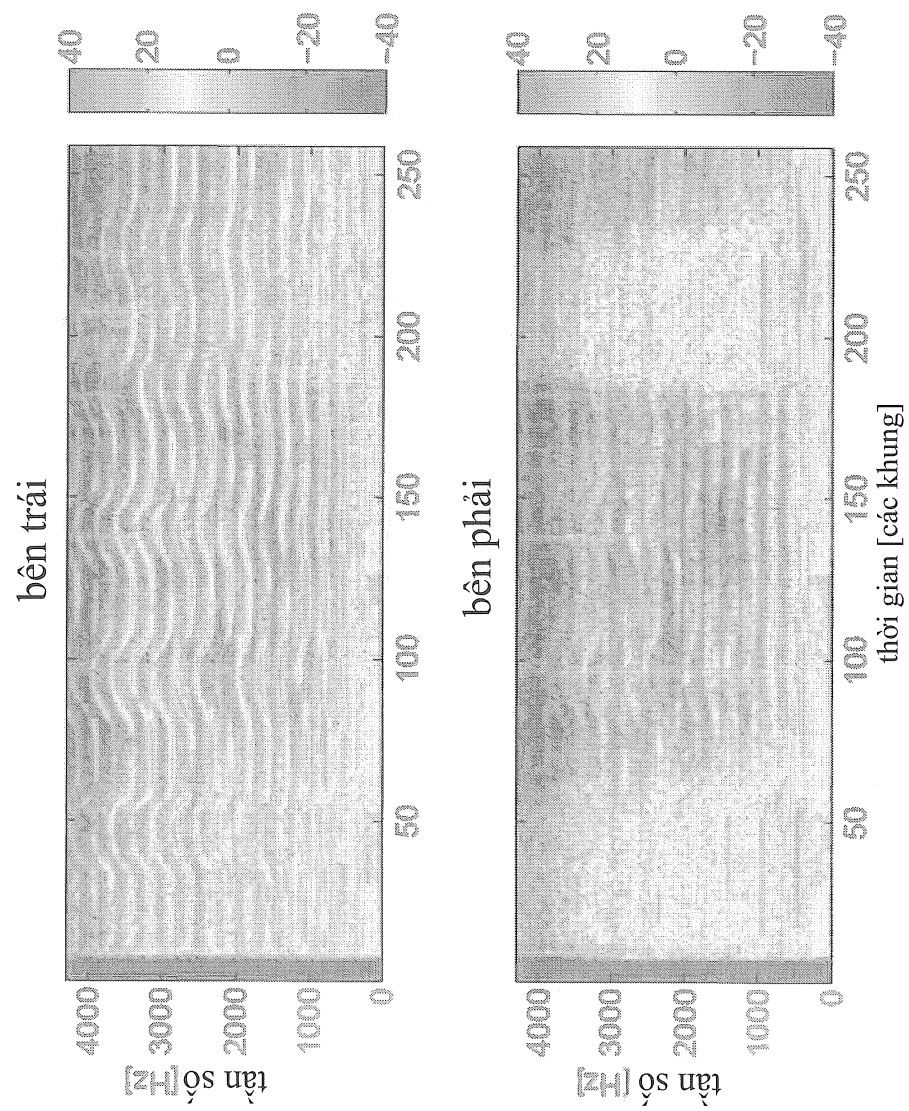


FIG 10

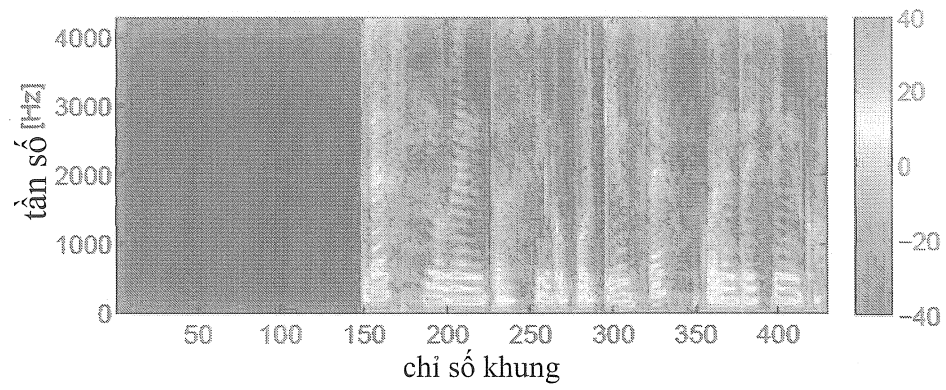


FIG 11A

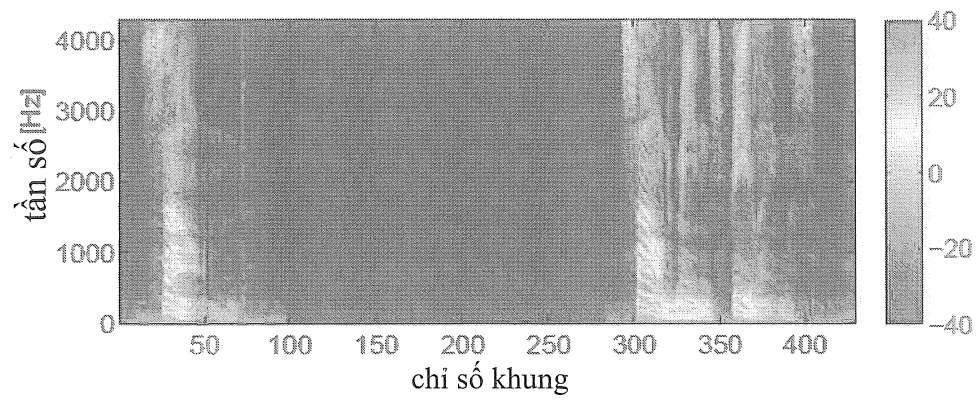


FIG 11B

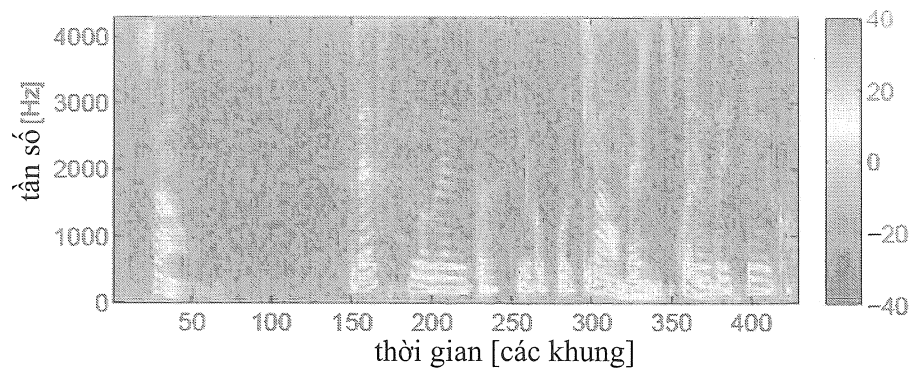


FIG 11C

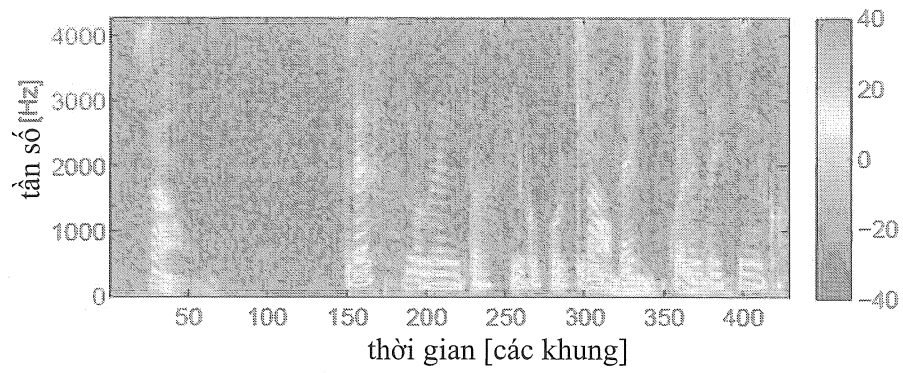


FIG 11D

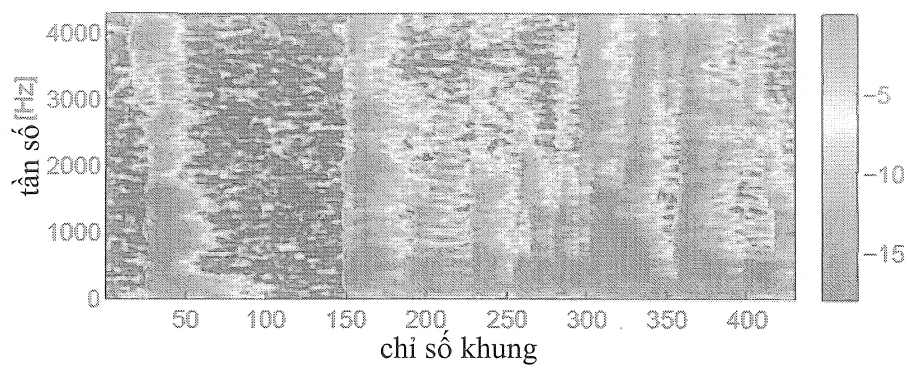


FIG 12A

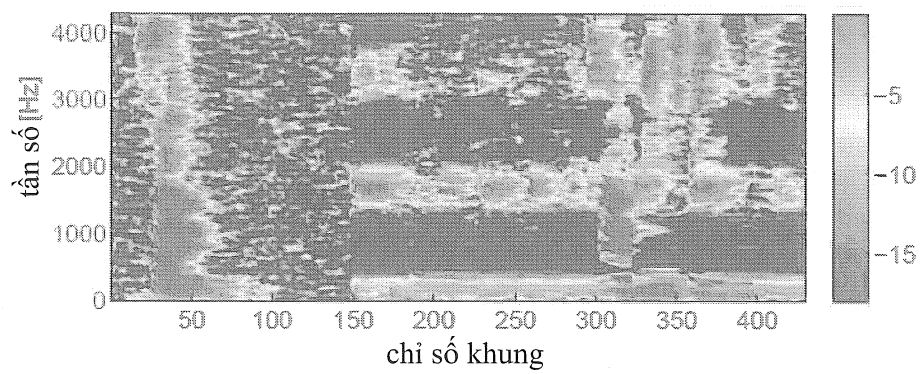


FIG 12B

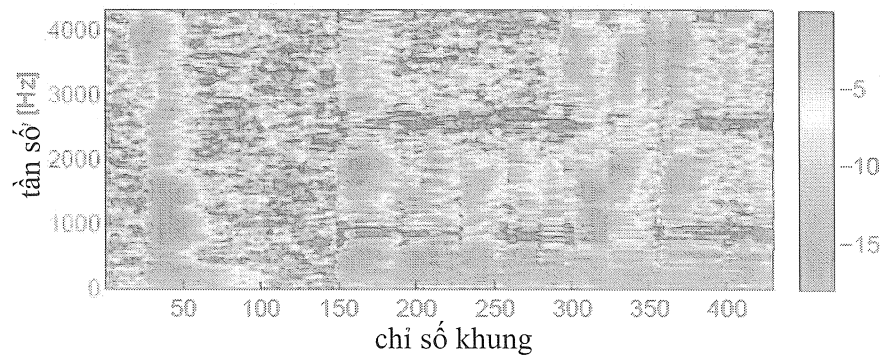


FIG 12C

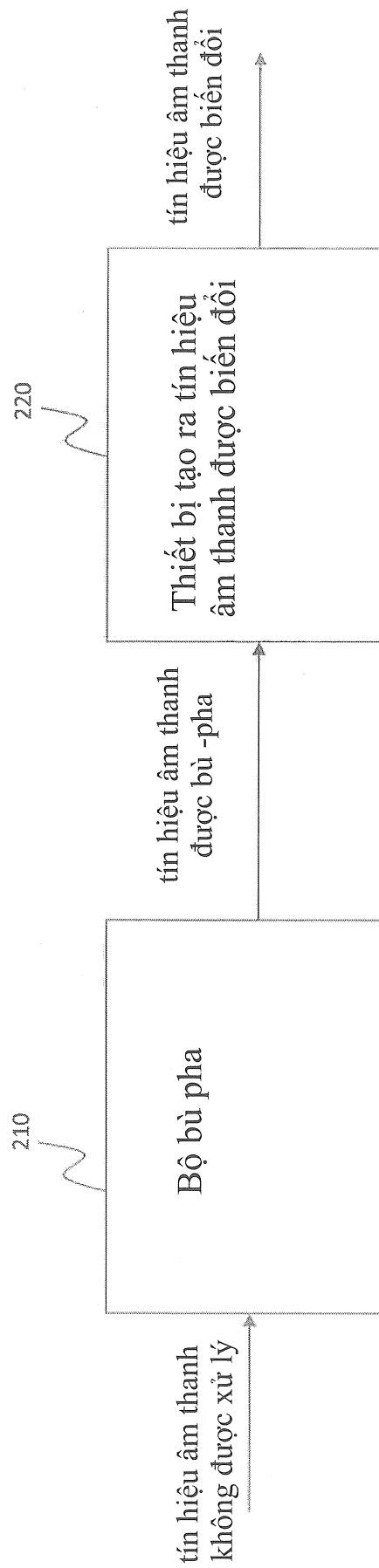


FIG 13