



(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)



1-0043540

(51)^{2019.01} **G10L 19/008**; H04S 7/00; G10L
21/0272; G06F 3/01

(13) **B**

(21) 1-2020-00766

(22) 13/07/2018

(86) PCT/EP2018/069145 13/07/2018

(87) WO 2019/012133 17/01/2019

(30) 17181484.1 14/07/2017 EP

(45) 25/02/2025 443

(43) 25/05/2020 386A

(73) Fraunhofer-Gesellschaft zur Foerderung der angewandten Forschung e.V. (DE)
Hansastrasse 27c, 80686 Muenchen, Germany (DE)

(72) HERRE, Juergen (DE); HABETS, Emanuel (NL).

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

(54) THIẾT BỊ VÀ PHƯƠNG PHÁP TẠO MÔ TẢ TRƯỜNG ÂM THANH ĐƯỢC
BIẾN ĐỔI

(21) 1-2020-00766

(57) Sáng chế đề cập đến thiết bị phương pháp tạo mô tả trường âm thanh được biến đổi. Thiết bị tạo mô tả trường âm thanh nâng cao bao gồm: bộ tạo trường âm thanh (100) để tạo ít nhất hai mô tả lớp trường âm thanh biểu thị các trường âm thanh đối với ít nhất một vị trí tham chiếu; và bộ tạo siêu dữ liệu (110) để tạo siêu dữ liệu liên quan đến thông tin không gian của các trường âm thanh, trong đó các mô tả trường âm thanh và siêu dữ liệu tạo thành mô tả trường âm thanh nâng cao. Siêu dữ liệu có thể là thông tin hình học cho mỗi lớp chẳng hạn như khoảng cách biểu diễn đến vị trí tham chiếu.

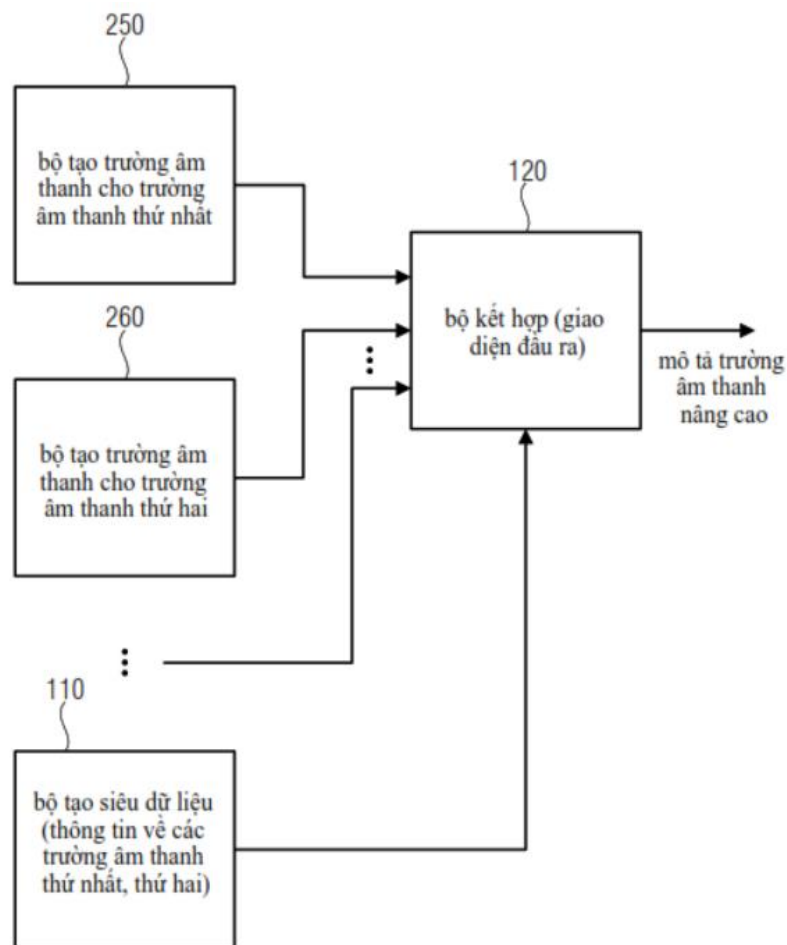


Fig. 2

Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến việc xử lý âm thanh và, đặc biệt là, xử lý âm thanh liên quan đến các trường âm thanh mà được xác định đối với vị trí tham chiếu chẳng hạn như vị trí micrô hoặc micrô ảo.

Tình trạng kỹ thuật của sáng chế

Các tín hiệu Ambisonics (âm thanh bao quanh) bao gồm phép phân tích sóng hài hình cầu bị cắt vát của trường âm thanh. Ambisonics có nhiều kiểu khác nhau. Có Ambisonics 'truyền thống' [31] mà ngày nay được biết đến là 'Ambisonics cấp một' (First-Order Ambisonics - FOA) và bao gồm bốn tín hiệu (nghĩa là, một tín hiệu đẳng hướng và lên đến ba tín hiệu định hướng hình số tám). Nhiều biến thể Ambisonics hơn gần đây được biết đến là 'Ambisonics cấp cao' (Higher-Order Ambisonics - HOA) và cung cấp độ phân giải không gian nâng cao và vùng điểm ngọt người nghe lớn hơn với tùy chỉnh mang nhiều tín hiệu hơn. Nói chung, phép biểu diễn HOA cấp N được xác định đầy đủ bao gồm $(N+1)^2$ tín hiệu.

Liên quan đến ý tưởng Ambisonics, phép biểu diễn mã hóa âm thanh định hướng (Directional Audio Coding - DirAC) được nhận thức để biểu diễn cảnh âm thanh FOA hoặc HOA theo kiểu tham số gọn hơn. Cụ thể hơn, cảnh âm thanh không gian được biểu diễn bởi một (hoặc nhiều) kênh âm thanh được truyền mà biểu diễn sự trộn giảm của cảnh âm thanh và được liên kết với thông tin phụ của hướng và sự khuếch tán trong mỗi ngăn tần số-thời gian (time-frequency - TF). Nhiều thông tin hơn về DirAC có thể được tìm thấy trong các tài liệu [32, 33].

DirAC [32] có thể được sử dụng với các hệ thống micrô khác nhau và với các thiết lập loa phóng thanh tùy ý. Mục đích của hệ thống DirAC là tái tạo ấn tượng không gian của môi trường âm thanh hiện có chính xác nhất có thể bằng cách sử dụng hệ thống đa kênh/loa phóng thanh 3D. Trong môi trường được chọn, các phản hồi (âm thanh liên tục hoặc các phản hồi xung) được đo bằng micrô đẳng hướng (W) và bằng bộ micrô cho phép đo hướng đến của âm thanh và sự khuếch tán của âm thanh. Phương pháp phổ biến là áp dụng ba micrô hình số tám (X, Y, Z) được xếp hàng tương

ứng theo trục tọa độ Đề-các [34]. Cách để thực hiện phương pháp này là sử dụng micrô trường âm thanh, trực tiếp mang lại các phản hồi mong muốn. Các tín hiệu W, X, Y, và Z cũng có thể được tính toán từ bộ các micrô đẳng hướng rời rạc.

Trong DirAC, tín hiệu âm thanh trước tiên được chia thành các kênh tần số. Hướng và sự khuếch tán âm thanh được đo phụ thuộc vào thời gian tại mỗi kênh tần số. Trong phép truyền, một hoặc nhiều kênh âm thanh được gửi, cùng với dữ liệu về hướng và sự khuếch tán được phân tích. Trong quá trình tổng hợp, âm thanh mà được áp dụng cho các loa phóng thanh có thể ví dụ là kênh đẳng hướng W, hoặc âm thanh cho mỗi loa phóng thanh có thể được tính toán là tổng của W, X, Y, và Z, mà tạo thành tín hiệu có các đặc điểm định hướng nhất định cho mỗi loa phóng thanh. Mỗi kênh âm thanh được chia thành các kênh tần số, mà sau đó được chia tùy ý để khuếch tán và đến các dòng không khuếch tán tùy thuộc vào sự khuếch tán được phân tích. Dòng khuếch tán được tái tạo với kỹ thuật, mà tạo ra cảm quan khuếch tán của cảnh âm thanh, ví dụ, các kỹ thuật giải tương quan được sử dụng trong mã hóa tín hiệu âm lập thể [35-37]. Âm thanh không khuếch tán được tái tạo với kỹ thuật nhằm tạo ra nguồn ảo giống như điểm theo dữ liệu hướng (ví dụ VBAP [38]).

Ba kỹ thuật điều hướng trong 6DoF với mức độ tự do hạn chế được đề xuất trong tài liệu [39]. Đưa ra tín hiệu Ambisonics duy nhất, tín hiệu Ambisonics duy nhất được tính toán bằng cách sử dụng: 1) mô phỏng phát lại HOA và chuyển động người nghe trong dãy loa phóng thanh ảo, 2) tính toán và tịnh tiến dọc theo các sóng phẳng, và 3) mở rộng lại trường âm thanh về người nghe.

Hơn nữa, tham chiếu đến công nghệ DirAC như được mô tả, ví dụ, trong công bố "Directional Audio Coding – Perception-Based Reproduction of Spatial Sound", V. Pulkki và cộng sự, Hội thảo quốc tế về các nguyên tắc và ứng dụng của thính giác không gian, 11-13 tháng 11, 2009, Zao, Miyagi, Nhật Bản. Sự tham chiếu này mô tả việc lập trình âm thanh định hướng làm ví dụ cho vị trí tham chiếu liên quan đến việc xử lý trường âm thanh đặc biệt là kỹ thuật được thúc đẩy một cách cảm quan cho việc xử lý âm thanh không gian. Nó có các ứng dụng trong việc thu, mã hóa, và tổng hợp lại âm thanh không gian, trong hội thảo từ xa, trong bộ lọc định hướng, và trong các môi trường thính giác ảo.

Sự tái tạo của cảnh âm thanh thường tập trung vào các thiết lập loa phóng thanh, vì đây là sự tái tạo điển hình riêng ví dụ như phòng khách, và bối cảnh chuyên nghiệp ví dụ như các rạp chiếu phim. Ở đây, mối quan hệ của cảnh với dạng hình tái tạo là tĩnh vì nó đi kèm cùng với hình ảnh hai chiều mà buộc người nghe phải nhìn theo hướng phía trước. Sau đó, mối quan hệ không gian của các đối tượng âm thanh và hình ảnh được xác định và cố định tại thời điểm sản xuất.

Trong thực tế ảo (virtual reality - VR), sự nhập vai đạt được rõ ràng bằng cách cho phép người dùng di chuyển tự do trong cảnh. Do đó, cần phải theo dõi chuyển động của người dùng và điều chỉnh sự tái tạo thị giác và thính giác tới vị trí của người dùng. Thông thường, người dùng đeo màn hình gắn trên đầu (head-mounted display - HMD) và tai nghe. Để có trải nghiệm tuyệt vời với tai nghe, âm thanh phải được lập thể. Sự lập thể là mô phỏng về cách mà đầu, tai, và thân trên của con người thay đổi âm thanh của nguồn tùy thuộc vào hướng và khoảng cách của nó. Điều này đạt được bằng cách tích chập các tín hiệu với các hàm truyền liên quan đến đầu (head-related transfer function - HRTF) cho hướng tương đối của chúng [1, 2]. Sự lập thể cũng làm cho âm thanh dường như đến từ cảnh chứ không phải từ bên trong đầu [3]. Trường hợp phổ biến đã được giải quyết thành công là tái tạo video 360° [4, 5]. Ở đây, người dùng đang hoặc đeo HMD hoặc cầm máy tính bảng hoặc điện thoại trên tay. Bằng cách di chuyển đầu của người đó hoặc thiết bị, người dùng có thể nhìn xung quanh theo hướng bất kỳ. Đây là trường hợp ba độ tự do (three-degrees-of-freedom - 3DoF), vì người dùng có ba độ chuyển động (dọc, ngang, xoay). Một cách trực quan, điều này được nhận ra bằng cách chiếu video trên hình cầu xung quanh người dùng. Âm thanh thường được ghi lại bằng micrô không gian [6], ví dụ, Ambisonics cấp một (FOA), gắn với máy quay video. Trong miền Ambisonics, sự xoay đầu của người dùng được thích ứng theo cách đơn giản [7]. Âm thanh sau đó được ví dụ kết xuất đến các loa phóng thanh được đặt xung quanh người dùng. Các tín hiệu loa phóng thanh ảo sau đó được lập thể.

Các ứng dụng VR hiện đại cho phép sáu độ tự do (six-degrees-of-freedom - 6DoF). Ngoài sự xoay đầu, người dùng có thể di chuyển xung quanh dẫn đến chuyển tiếp vị trí của người đó theo ba chiều không gian. Sự tái tạo 6DoF bị giới hạn bởi kích

thước tổng thể của vùng đi bộ. Trong nhiều trường hợp, vùng này khá nhỏ, ví dụ, phòng khách thông thường. 6DoF thường gặp trong các trò chơi VR. Ở đây, toàn bộ cảnh được tổng hợp với hình ảnh do máy tính tạo ra (computer-generated imagery - CGI). Âm thanh thường được tạo bằng cách sử dụng việc kết xuất dựa trên đối tượng trong đó mỗi đối tượng âm thanh được kết xuất với độ khuếch đại tăng phụ thuộc vào khoảng cách và hướng tương đối từ người dùng dựa trên dữ liệu theo dõi. Hiện thực có thể được nâng cao bằng tiếng vang và nhiễu xạ nhân tạo [8, 9, 10].

Về nội dung được ghi lại, có một số thách thức khác biệt để thuyết phục tái tạo 6DoF nghe nhìn. Ví dụ ban đầu của thao tác âm thanh không gian trong miền chuyển tiếp không gian là các kỹ thuật 'phóng đại âm thanh' [11, 12]. Ở đây, vị trí người nghe hầu như được di chuyển vào cảnh thị giác được ghi lại, tương tự như phóng to hình ảnh. Người dùng lựa chọn một hướng hoặc một phần hình ảnh và sau đó có thể nghe điều này từ điểm được tịnh tiến. Điều này đòi hỏi rằng tất cả các hướng đến (direction of arrival - DoA) đang thay đổi so với tái tạo ban đầu, không phóng đại.

Các phương pháp để tái tạo 6DoF của nội dung được ghi lại đã sử dụng các vị trí ghi lại được phân bố theo không gian được đề xuất. Đối với video, các dây của các camera có thể được dùng để tạo kết xuất trường ánh sáng [13]. Đối với âm thanh, thiết lập tương tự dùng các dây micro được phân bố hoặc các micro Ambisonics. Nó đã được chứng minh rằng có thể tạo ra tín hiệu của 'micro ảo' được đặt tại vị trí tùy ý từ các bản ghi như vậy [14].

Để nhận ra các sửa đổi âm thanh không gian như vậy một cách thuận tiện về mặt kỹ thuật, có thể dùng các kỹ thuật lập trình hoặc xử lý âm thanh tham số (xem tài liệu [15] để tham khảo tổng quát). Mã hóa âm thanh định hướng (Directional audio coding - DirAC) [16] là phương pháp phổ biến để biến đổi bản ghi thành phép biểu diễn mà bao gồm phổ âm thanh và thông tin phụ tham số về hướng âm thanh và sự khuếch tán. Nó được sử dụng cho phóng đại âm thanh [11] và các ứng dụng micro ảo [14].

Phương pháp được đề xuất ở đây cho phép sự tái tạo 6DoF từ bản ghi của micro FOA. Các bản ghi từ vị trí không gian đơn được sử dụng cho sự tái tạo 3DoF hoặc phóng đại âm thanh. Nhưng, theo sự hiểu biết của các tác giả sáng chế, cho đến nay,

không có phương pháp nào để tái tạo hoàn toàn 6DoF từ dữ liệu như vậy được đề xuất. Người ta nhận ra tái tạo 6DoF bằng cách tích hợp thông tin về khoảng cách của các nguồn âm thanh trong bản ghi. Thông tin khoảng cách này được hợp nhất vào phép biểu diễn tham số của DirAC, sao cho phối cảnh thay đổi của người nghe được ánh xạ chính xác.

Không có biểu diễn trường âm thanh Ambisonics nào (có thể là Ambisonics FOA hoặc HOA thông thường hoặc là phép biểu diễn trường âm thanh tham số theo kiểu DirAC) cung cấp đủ thông tin để cho phép dịch chuyển vị trí của người nghe vì nó cần cho các ứng dụng 6DoF vì không phải khoảng cách đối tượng cũng không phải các vị trí đối tượng tuyệt đối trong cảnh âm thanh được xác định trong các định dạng này. Cần lưu ý rằng sự dịch chuyển vị trí của người nghe có thể được chuyển thành sự dịch chuyển tương đương của cảnh âm thanh theo hướng ngược lại.

Vấn đề diễn hình khi di chuyển trong 6DoF được minh họa trên Fig.1b. Giả sử rằng cảnh âm thanh được mô tả tại vị trí A bằng cách sử dụng Ambisonics. Trong trường hợp này, các âm thanh từ nguồn A và nguồn B đến từ cùng một hướng, tức là, chúng có cùng hướng đến (DOA). Trong trường hợp một người di chuyển đến vị trí B, DOA của nguồn A và nguồn B là khác nhau. Sử dụng mô tả Ambisonics tiêu chuẩn của trường âm thanh, nghĩa là, không có thông tin bổ sung, không thể tính toán các tín hiệu Ambisonics tại vị trí B cho các tín hiệu Ambisonics tại vị trí A.

Tài liệu tham khảo

- [1] Liitola, T., “Headphone sound externalization”, luận án Tiến sĩ, Đại học Công nghệ Helsinki, Phòng Kỹ thuật Điện và Truyền thông Phòng thí nghiệm âm thanh và xử lý tín hiệu âm thanh, 2006.
- [2] Blauert, J., “Spatial Hearing - Revised Edition: The Psychophysics of Human Sound Localization”, nhà xuất bản MIT, 1996, ISBN 0262024136.
- [3] Zhang, W., Samarasinghe, P. N., Chen, H., and Abhayapala, T. D., “Surround by Sound: A Re-view of Spatial Audio Recording and Reproduction,” *Applied Sciences*, 7(5), trang 532, 2017.
- [4] Bates, E. và Boland, F., “Spatial Music, Virtual Reality, and 360 Media,” trong

Audio Eng. Soc. Int. Conf. on Audio for Virtual and Augmented Reality, Los Angeles, CA, Mỹ, 2016.

- [5] Anderson, R., Gallup, D., Barron, J. T., Kontkanen, J., Snavely, N., Esteban, C. H., Agarwal, S., and Seitz, S. M., “Jump: Virtual Reality Video,” *ACM Transactions on Graphics*, 35(6), trang 198, 2016.
- [6] Merimaa, J., *Analysis, Synthesis, and Perception of Spatial Sound: Binaural Localization Modeling and Multichannel Loudspeaker Reproduction*, luận án Tiến sĩ, Đại học Công nghệ Helsinki, 2006.
- [7] Kronlachner, M. and Zotter, F., “Spatial Trans-formations for the Enhancement of Ambisonics Recordings,” trong *2nd International Conference on Spatial Audio*, Erlangen, Đức, 2014.
- [8] Tsingos, N., Gallo, E., và Drettakis, G., “Perceptual Audio Rendering of Complex Virtual Environments,” *ACM Transactions on Graphics*, 23(3), pp. 249–258, 2004.
- [9] Taylor, M., Chandak, A., Mo, Q., Lauterbach, C., Schissler, C., và Manocha, D., “Guided multi-view ray tracing for fast auralization,” *IEEE Trans. Visualization & Comp. Graphics*, 18, pp. 1797– 1810, 2012.
- [10] Rungta, A., Schissler, C., Rewkowski, N., Mehra, R., và Manocha, D., “Diffraction Kernels for Interactive Sound Propagation in Dynamic Environments,” *IEEE Trans. Visualization & Comp. Graphics*, 24(4), pp. 1613–1622, 2018.
- [11] Thiergart, O., Kowalczyk, K., và Habets, E. A. P., “An Acoustical Zoom based on Informed Spatial Filtering,” trong *Int. Workshop on Acoustic Signal Enhancement*, pp. 109–113, 2014.
- [12] Khaddour, H., Schimmel, J., và Rund, F., “A Novel Combined System of Direction Estimation và Sound Zooming of Multiple Speakers,” *Radioengineering*, 24(2), 2015.
- [13] Ziegler, M., Keinert, J., Holzer, N., Wolf, T., Jaschke, T., op het Veld, R., Zakeri, F. S., và Foessel, S., “Immersive Virtual Reality for Live-Action Video using Camera Arrays,” trong *IBC*, Amsterdam, Netherlands, 2017.

- [14] Thiergart, O., Galdo, G. D., Taseska, M., và Habets, E. A. P., “Geometry-Based Spatial Sound Acquisition using Distributed Microphone Arrays,” *IEEE Trans. Audio, Speech, Language Process.*, 21(12), pp. 2583–2594, 2013.
- [15] Kowalczyk, K., Thiergart, O., Taseska, M., Del Galdo, G., Pulkki, V., và Habets, E. A. P., “Parametric Spatial Sound Processing: A Flexible and Efficient Solution to Sound Scene Acquisition, Modification, and Reproduction,” *IEEE Signal Process. Mag.*, 32(2), pp. 31–42, 2015.
- [16] Pulkki, V., “Spatial Sound Reproduction with Directional Audio Coding,” *J. Audio Eng. Soc.*, 55(6), pp. 503–516, 2007.
- [17] International Telecommunication Union, “ITU-R BS.1534-3, Method for the subjective assessment of intermediate quality level of audio systems,” 2015.
- [18] Thiergart, O., Del Galdo, G., Kuech, F., và Prus, M., “Three-Dimensional Sound Field Analysis with Directional Audio Coding Based on Signal Adaptive Parameter Estimators,” trong *Audio Eng. Soc. Conv. Spatial Audio: Sense the Sound of Space*, 2010.
- [19] Kuttruff, H., *Room Acoustics*, Taylor & Francis, 4 edition, 2000.
- [20] Borß, C., “A polygon-based panning method for 3D loudspeaker setups,” trong *Audio Eng. Soc. Conv.*, pp. 343–352, Los Angeles, CA, Mỹ, 2014.
- [21] Rummukainen, O., Schlecht, S., Plinge, A., và Habets, E. A. P., “Evaluating Binaural Reproduction Systems from Behavioral Patterns in a Virtual Reality – A Case Study with Impaired Binaural Cues and Tracking Latency,” trong *Audio Eng. Soc. Conv. 143*, New York, NY, Mỹ, 2017.
- [22] Engelke, U., Darcy, D. P., Mulliken, G. H., Bosse, S., Martini, M. G., Arndt, S., Antons, J.-N., Chan, K. Y., Ramzan, N., và Brunnström, K., “Psychophysiology-Based QoE Assessment: A Survey,” *IEEE Selected Topics in Signal Processing*, 11(1), pp. 6–21, 2017.
- [23] Schlecht, S. J. và Habets, E. A. P., “Sign-Agnostic Matrix Design for Spatial Artificial Reverberation with Feedback Delay Networks,” trong *Proc. Audio Eng. Soc. Conf.*, pp. 1–10– accepted, Tokyo, Nhật Bản, 2018.

- [31] M. A. Gerzon, "Periphony: With-height sound reproduction," *J. Acoust. Soc. Am.*, vol. 21, no. 1, pp. 2-10, 1973.
- [32] V. Pulkki, "Directional audio coding in spatial sound reproduction and stereo upmixing," trong *Proc. of the 28th AES International Conference*, 2006.
- [33] —, "Spatial sound reproduction with directional audio coding," *Journal Audio Eng. Soc.*, vol. 55, no. 6, pp. 503-516, Jun. 2007.
- [34] C. G. và G. M., "Coincident microphone simulation covering three dimensional space and yielding various directional outputs," Bằng sáng chế Mỹ số 4 042 779, 1977.
- [35] C. Faller và F. Baumgarte, "Binaural cue coding - part ii: Schemes and applications," *IEEE Trans. Speech Audio Process.*, vol. 11, no. 6, tháng 11 năm 2003.
- [36] C. Faller, "Parametric multichannel audio coding: Synthesis of coherence cues," *IEEE Trans. Speech Audio Process.*, vol. 14, no. 1, tháng 01 năm 2006.
- [37] H. P. J. E. E. Schuijers, J. Breebaart, "Low complexity parametric stereo coding," trong *Proc. of the 116th AES Convention*, Berlin, Đức, 2004.
- [38] V. Pulkki, "Virtual sound source positioning using vector base amplitude panning," *J. Acoust. Soc. Am.*, vol. 45, no. 6, pp. 456-466, tháng 06 năm 1997.
- [39] J. G. Tylka và E. Y. Choueiri, "Comparison of techniques for binaural navigation of higher- order ambisonics sound fields," trong *Proc. of the AES International Conference on Audio for Virtual and Augmented Reality*, New York, tháng 09 năm 2016.

Bản chất kỹ thuật của sáng chế

Mục đích của sáng chế là một mặt đề xuất mô tả trường âm thanh nâng cao hoặc tạo mô tả trường âm thanh được biến đổi, mặt khác cho phép quá trình xử lý được cải thiện hoặc linh hoạt hoặc hiệu quả.

Mục đích này đạt được bởi thiết bị tạo mô tả trường âm thanh nâng cao theo điểm 1, thiết bị tạo mô tả trường âm thanh được biến đổi theo điểm 12, phương pháp tạo mô tả của trường âm thanh nâng cao theo điểm 39, phương pháp tạo mô tả trường âm thanh được biến đổi theo điểm 40, chương trình máy tính theo điểm 41 hoặc mô tả trường âm thanh theo điểm 43.

Sáng chế dựa trên việc phát hiện rằng các mô tả trường âm thanh điển hình mà liên quan đến vị trí tham chiếu cần thông tin bổ sung để các mô tả trường âm thanh này có thể được xử lý để mô tả trường âm thanh được biến đổi mà không liên quan đến vị trí tham chiếu ban đầu nhưng liên quan đến vị trí tham chiếu khác có thể được tính toán. Để đạt được mục đích này, siêu dữ liệu liên quan đến thông tin không gian của trường âm thanh được tạo và siêu dữ liệu cùng với mô tả trường âm thanh tương ứng với mô tả trường âm thanh nâng cao mà có thể, ví dụ, được truyền hoặc được lưu trữ. Để tạo mô tả trường âm thanh được biến đổi từ mô tả trường âm thanh và siêu dữ liệu và, cụ thể, siêu dữ liệu liên quan đến thông tin không gian của mô tả trường âm thanh, trường âm thanh được biến đổi được tính toán bằng cách sử dụng thông tin âm thanh này, mô tả trường âm thanh và thông tin tịnh tiến biểu thị phép tịnh tiến từ vị trí tham chiếu đến vị trí tham chiếu khác. Do đó, mô tả trường âm thanh nâng cao bao gồm mô tả trường âm thanh và siêu dữ liệu liên quan đến thông tin không gian của trường âm thanh này nằm dưới mô tả trường âm thanh được xử lý để thu được mô tả trường âm thanh được biến đổi liên quan đến vị trí tham chiếu khác được định rõ bởi thông tin tịnh tiến bổ sung mà có thể, ví dụ, được cung cấp hoặc được sử dụng tại phía bộ giải mã.

Tuy nhiên, sáng chế không chỉ đề cập đến trường hợp bộ mã hóa/bộ giải mã, mà còn có thể được áp dụng trong ứng dụng mà tại đó cả việc tạo mô tả trường âm thanh nâng cao và tạo mô tả trường âm thanh được biến đổi diễn ra trên cơ bản một và cùng một địa điểm. Mô tả trường âm thanh được biến đổi có thể, ví dụ, là mô tả của chính trường âm thanh được biến đổi hoặc thực tế trường âm thanh được biến đổi trong các tín hiệu kênh, các tín hiệu lập thể hoặc, một lần nữa, trường âm thanh liên quan đến vị trí tham chiếu mà, tuy nhiên, liên quan đến vị trí tham chiếu mới hoặc khác thay vì vị trí tham chiếu ban đầu. Chẳng hạn, ứng dụng như vậy sẽ ở trong kịch bản thực tế ảo ở

đó mô tả trường âm thanh cùng với siêu dữ liệu tồn tại và ở đó người nghe di chuyển ra khỏi vị trí tham chiếu mà trường âm thanh được đưa ra và di chuyển đến vị trí tham chiếu khác và ở đó, sau đó, trường âm thanh cho người nghe di chuyển xung quanh trong khu vực ảo được tính toán tương ứng với trường âm thanh nhưng bây giờ tại vị trí tham chiếu ở đó người dùng đã di chuyển tới.

Trong phương án, bộ tạo trường âm thanh có thể tạo ra, cho mỗi lớp có số lượng ít nhất là hai lớp, mô tả DirAC của trường âm thanh có một hoặc nhiều tín hiệu trộn giảm và dữ liệu hướng riêng lẻ và dữ liệu khuếch tán tùy ý cho các ngăn thời gian-tần số khác nhau. Trong ngữ cảnh này, bộ tạo siêu dữ liệu được tạo cấu hình để tạo thêm thông tin về khoảng cách hoặc độ sâu cho mỗi lớp, vì siêu dữ liệu là giống nhau cho tất cả các nguồn có trong lớp và cho tất cả các ngăn tần số liên quan đến lớp. Phương án được ưu tiên sẽ có khoảng cách hoặc phạm vi khoảng cách liên quan đến lớp. Ngoài ra, bản đồ độ sâu có thể được cung cấp cho mỗi lớp. Đặc biệt, và trong phương án khác, siêu dữ liệu liên quan đến thông tin không gian sẽ là bản đồ độ sâu liên quan đến khoảng cách nhất định với thông tin vị trí nhất định chẳng hạn như thông tin hướng đến.

Trong các phương án, hướng đến được đưa ra chỉ bằng góc nâng hoặc chỉ bằng góc phương vị hoặc cả hai góc và bản đồ độ sâu sau đó sẽ liên kết, với mỗi nguồn trong lớp có cùng thông tin khoảng cách chẳng hạn như khoảng cách tính bằng mét hoặc khoảng cách tương đối hoặc khoảng cách tuyệt đối hoặc tương đối được lượng tử hóa hoặc bất kỳ thông tin khoảng cách nào khác, từ đó, cuối cùng, khoảng cách liên quan đến vị trí tham chiếu khác hoặc mới, có thể xuất phát từ trường âm thanh.

Sau đây, các triển khai được ưu tiên khác được phác thảo.

Ambisonics trở thành một trong những định dạng được sử dụng phổ biến nhất cho âm thanh 3D trong bối cảnh của các ứng dụng thực tế ảo, tăng cường và hỗn hợp. Một loạt các công cụ tiếp nhận và sản xuất âm thanh được phát triển để tạo ra tín hiệu đầu ra ở định dạng Ambisonics. Để diễn tả nội dung được mã hóa Ambisonics trong những ứng dụng thực tế ảo (VR) tương tác, định dạng Ambisonics được tinh chỉnh thành tín hiệu lập thể hoặc các kênh để tái tạo. Trong các ứng dụng được nói ở trên, người nghe thường có thể thay đổi một cách tương hỗ sự định hướng của người đó trong

cảnh được diễn tả đến mức người đó có thể xoay đầu trong cảnh âm thanh cho phép ba độ tự do (3DoF, nghĩa là, dọc, ngang, xoay) và vẫn trải nghiệm chất lượng âm thanh phù hợp. Điều này được thực hiện bằng cách xoay cảnh âm thanh trước khi kết xuất theo định hướng đầu, mà có thể được thực hiện với độ phức tạp tính toán thấp và là lợi thế của phép biểu diễn Ambisonics. Tuy nhiên, trong các ứng dụng mới nổi, chẳng hạn như VR, mong muốn cho phép người dùng di chuyển tự do trong cảnh âm thanh thay vì chỉ thay đổi hướng (được gọi là 'sáu độ tự do' hay 6DoF). Do đó, việc xử lý tín hiệu là cần thiết để thay đổi phối cảnh của cảnh âm thanh (nghĩa là để di chuyển gần như trong cảnh âm thanh dọc theo các trục x, y, hoặc z). Tuy nhiên, nhược điểm lớn của Ambisonics là định dạng mô tả trường âm thanh từ góc nhìn duy nhất trong cảnh âm thanh. Cụ thể, nó không chứa thông tin về địa điểm thực tế của các nguồn âm thanh trong cảnh âm thanh sẽ cho phép dịch chuyển cảnh âm thanh ('tịnh tiến') như được yêu cầu cho 6DoF. Phần mô tả của sáng chế này cung cấp những sự mở rộng của Ambisonics để khắc phục vấn đề này và cũng tạo điều kiện cho sự tịnh tiến, và do đó cho phép 6DoF chính xác.

Những bản ghi Ambisonics cấp một (FOA) có thể được xử lý và được tái tạo qua tai nghe. Chúng có thể được xoay để tính đến sự định hướng đầu của người nghe. Tuy nhiên, các hệ thống thực tế ảo (VR) cho phép người nghe di chuyển trong sáu độ tự do (6DoF), nghĩa là, ba vòng xoay cộng với ba độ tịnh tiến tự do. Ở đây, các góc và khoảng cách rõ ràng của các nguồn âm thanh phụ thuộc vào vị trí của người nghe. Kỹ thuật để tạo điều kiện cho 6DoF được mô tả. Cụ thể, bản ghi FOA được mô tả sử dụng mô hình tham số, mà được biến đổi dựa trên vị trí người nghe và thông tin về các khoảng cách tới các nguồn. Phương pháp được đánh giá bằng bài kiểm tra nghe, so sánh các kết xuất lập thể khác nhau của cảnh âm thanh tổng hợp trong đó người nghe có thể di chuyển tự do.

Trong các phương án ưu tiên hơn nữa, mô tả trường âm thanh nâng cao được đưa ra bởi giao diện đầu ra để tạo tín hiệu đầu ra để truyền hoặc lưu trữ, trong đó tín hiệu đầu ra bao gồm, đối với khung thời gian, một hoặc nhiều tín hiệu âm thanh phát ra từ trường âm thanh và thông tin không gian cho khung thời gian. Đặc biệt, bộ tạo trường âm thanh trong các phương án khác thích ứng để thu được dữ liệu hướng từ

trường âm thanh, dữ liệu hướng liên quan đến hướng đến của âm thanh trong một khoảng thời gian hoặc ngăn tần số và bộ tạo siêu dữ liệu được tạo cấu hình để thu được thông tin không gian như các mục dữ liệu liên kết thông tin khoảng cách với dữ liệu hướng.

Đặc biệt, trong phương án như vậy, giao diện đầu ra được tạo cấu hình để tạo các tín hiệu đầu ra sao cho các mục dữ liệu đối với các khung thời gian được liên kết với các dữ liệu hướng cho các ngăn tần số khác nhau.

Trong phương án khác, bộ tạo trường âm thanh cũng được tạo cấu hình để tạo thông tin khuếch tán cho nhiều ngăn tần số của khung thời gian của trường âm thanh, trong đó bộ tạo siêu dữ liệu được tạo cấu hình để chỉ tạo thông tin khoảng cách cho ngăn tần số khác với giá trị được định trước, hoặc khác với vô hạn hoặc để tạo giá trị khoảng cách cho ngăn tần số, khi giá trị khuếch tán thấp hơn ngưỡng được định trước hoặc thích ứng. Do đó, đối với các ngăn thời gian/tần số mà có sự khuếch tán cao, mọi giá trị khoảng cách không được tạo hoặc giá trị khoảng cách được định trước được tạo mà được làm sáng tỏ bởi bộ giải mã theo cách nhất định. Do đó, chắc chắn rằng đối với các ngăn thời gian/tần số có sự khuếch tán cao, mọi kết xuất liên quan đến khoảng cách không được thực hiện, vì sự khuếch tán cao biểu thị rằng, đối với ngăn thời gian/tần số như vậy, âm thanh không đến từ nguồn được định vị nhất định mà đến từ bất kỳ hướng nào và, do đó, là giống nhau bất kể trường âm thanh được cảm nhận tạo vị trí tham chiếu ban đầu hay vị trí tham chiếu khác hoặc mới.

Đối với bộ tính toán trường âm thanh, các phương án được ưu tiên bao gồm giao diện tịnh tiến để cung cấp thông tin tịnh tiến hoặc thông tin xoay biểu thị sự xoay của người nghe dự định đến trường âm thanh được biến đổi, nguồn cung cấp siêu dữ liệu để cung cấp siêu dữ liệu tới bộ tính toán trường âm thanh và nguồn cung cấp trường âm thanh để cung cấp mô tả trường âm thanh tới bộ tính toán trường âm thanh và, ngoài ra, giao diện đầu ra để đưa ra trường âm thanh được biến đổi bao gồm mô tả trường âm thanh và siêu dữ liệu được biến đổi, siêu dữ liệu được biến đổi được thu từ siêu dữ liệu bằng cách sử dụng thông tin tịnh tiến, hoặc giao diện đầu ra đưa ra nhiều kênh loa phóng thanh, mỗi kênh loa phóng thanh liên quan đến vị trí loa phóng thanh được định

rõ trước, hoặc giao diện đầu ra đưa ra phép biểu diễn lập thể của trường âm thanh được biến đổi.

Trong phương án, mô tả trường âm thanh bao gồm nhiều thành phần trường âm thanh. Nhiều thành phần trường âm thanh bao gồm thành phần đẳng hướng và ít nhất một thành phần định hướng. Ví dụ, mô tả trường âm thanh là mô tả trường âm thanh Ambisonics cấp một có thành phần đẳng hướng và ba thành phần định hướng X, Y, Z hoặc trường âm thanh như vậy là mô tả Ambisonics cấp cao bao gồm thành phần đẳng hướng, ba thành phần định hướng đối với các hướng X, Y, và Z và, ngoài ra, các thành phần định hướng khác nữa mà có liên quan đến các hướng khác ngoài các hướng X, Y, Z.

Trong phương án, thiết bị bao gồm bộ phân tích để phân tích các thành phần trường âm thanh để suy ra, đối với các ngăn thời gian hoặc tần số khác nhau, thông tin hướng đến. Thiết bị còn có thêm bộ biến đổi tịnh tiến để tính toán thông tin DoA được biến đổi về ngăn tần số hoặc thời gian bằng cách sử dụng thông tin DoA và siêu dữ liệu, ở đó siêu dữ liệu liên quan đến bản đồ độ sâu kết hợp khoảng cách với mô tả lớp, nghĩa là với tất cả các nguồn của lớp, cho tất cả các ngăn tần số của khung thời gian. Do đó, "bản đồ độ sâu" tương đối đơn giản là đủ cho mỗi lớp. Bản đồ độ sâu cho một lớp chỉ yêu cầu ít nhất một khoảng cách hoặc phạm vi khoảng cách cho lớp này như trên Fig.4c.

Hơn nữa, bộ tính toán trường âm thanh có bộ bù khoảng cách để tính toán trường âm thanh được biến đổi bằng cách sử dụng thông tin bù khoảng cách tùy thuộc vào khoảng cách được cung cấp từ siêu dữ liệu là giống nhau cho mỗi ngăn tần số hoặc thời gian của nguồn từ lớp, và từ khoảng cách mới được liên kết với ngăn thời gian hoặc tần số, khoảng cách mới có liên quan đến thông tin DoA được biến đổi.

Trong phương án, bộ tính trường âm thanh tính vector thứ nhất chỉ từ vị trí tham chiếu tới nguồn âm thanh thu được bằng cách phân tích trường âm thanh. Ngoài ra, bộ tính trường âm thanh tính vector thứ hai chỉ từ vị trí tham chiếu khác đến nguồn âm thanh và phép tính này được thực hiện bằng cách sử dụng vector thứ nhất và thông tin tịnh tiến, ở đó thông tin tịnh tiến xác định vector tịnh tiến từ vị trí tham chiếu đến vị trí

tham chiếu khác. Và, sau đó, khoảng cách từ vị trí tham chiếu khác tới nguồn âm thanh được tính toán bằng cách sử dụng vector thứ hai.

Hơn nữa, bộ tính toán trường âm thanh được tạo cấu hình để nhận, ngoài thông tin tịnh tiến, thông tin xoay biểu thị phép xoay của đầu người nghe theo một trong ba hướng xoay dọc, ngang, xoay. Bộ tính toán trường âm thanh sau đó được tạo cấu hình để thực hiện biến đổi xoay để xoay dữ liệu hướng đến được biến đổi cho trường âm thanh bằng cách sử dụng thông tin xoay, trong đó dữ liệu hướng đến được biến đổi được suy ra từ phép xoay của dữ liệu nhận thu được bởi phép phân tích âm thanh của mô tả trường âm thanh và thông tin tịnh tiến.

Trong phương án, bộ tính trường âm thanh được tạo cấu hình để xác định các tín hiệu nguồn từ mô tả trường âm thanh và các hướng của các tín hiệu nguồn liên quan đến vị trí tham chiếu bởi phép phân tích âm thanh.

Sau đó, các hướng mới của các nguồn âm thanh được tính toán mà liên quan đến vị trí tham chiếu khác và điều này được thực hiện bằng cách sử dụng siêu dữ liệu, và sau đó thông tin khoảng cách của các nguồn âm thanh liên quan đến vị trí tham chiếu khác được tính toán và, sau đó, trường âm thanh được biến đổi được tổng hợp bằng cách sử dụng thông tin khoảng cách và các hướng mới của các nguồn âm thanh.

Trong phương án, việc tổng hợp trường âm thanh được thực hiện bằng cách quét các tín hiệu nguồn âm thanh sang hướng được cung cấp bởi thông tin hướng mới liên quan đến thiết lập tái tạo, và việc chia tỷ lệ của các tín hiệu nguồn âm thanh được thực hiện bằng cách sử dụng thông tin khoảng cách trước khi thực hiện thao tác quét. Các tín hiệu nguồn có thể cần được liên kết với lớp mới nếu khoảng cách thay đổi. Sau đó, hệ số tỷ lệ khác với 1 được sử dụng, nếu mô tả trường âm thanh được tạo. Nếu, theo hướng bộ chuyển mã, mô tả đa lớp mới được tạo ra, thì chỉ một thay đổi từ lớp này sang lớp khác là đủ để tính đến sự thay đổi khoảng cách. "Quét" có thể được thực hiện như được nêu trên Fig.4i để tạo mô tả trường âm thanh mới dưới dạng trường liên quan đến vị trí cụ thể. Tuy nhiên, để tạo các tín hiệu loa phóng thanh, có thể thực hiện việc quét bằng cách sử dụng các độ khuếch đại quét được suy ra từ DoA mới.

Trong phương án khác, phép tổng hợp trường âm thanh được thực hiện bằng cách xoay các tín hiệu nguồn âm thanh FOA hoặc HOA trong lớp theo hướng được

cung cấp bởi thông tin hướng mới liên quan đến việc thiết lập tái tạo, chia tỷ lệ các tín hiệu nguồn âm thanh, và cuối cùng là liên kết tín hiệu nguồn âm thanh với lớp bằng cách sử dụng thông tin khoảng cách. Điều này có thể thay thế "quét" như đã thảo luận.

Trong phương án khác, phần khuếch tán của tín hiệu nguồn âm thanh được bổ sung vào phần hướng của tín hiệu nguồn âm thanh, phần hướng được biến đổi bởi thông tin khoảng cách trước khi được bổ sung vào phần khuếch tán.

Trong một phương án khác, phần khuếch tán của tín hiệu nguồn âm thanh được thêm vào lớp chuyên dụng. Các tín hiệu liên quan đến lớp chuyên dụng này được xoay dựa trên thông tin hướng mới liên quan đến việc thiết lập tái tạo.

Đặc biệt, nên ưu tiên thực hiện tổng hợp nguồn âm thanh trong phép biểu diễn phổ ở đó thông tin hướng mới được tính toán cho mỗi ngăn tần số, ở đó thông tin khoảng cách được tính toán cho mỗi ngăn tần số, và ở đó phép tổng hợp hướng cho mỗi ngăn tần số bằng cách sử dụng tín hiệu âm thanh cho ngăn tần số được thực hiện bằng cách sử dụng tín hiệu âm thanh cho ngăn tần số, độ khuếch đại quét cho ngăn tần số được suy ra từ thông tin hướng mới và hệ số tỷ lệ cho ngăn tần số được suy ra từ thông tin khoảng cách cho ngăn tần số được thực hiện.

Hơn nữa, phép tổng hợp khuếch tán được thực hiện bằng cách sử dụng tín hiệu âm thanh khuếch tán được suy ra từ tín hiệu âm thanh từ ngăn tần số và bằng cách sử dụng tham số khuếch tán được suy ra từ phép phân tích tín hiệu cho ngăn tần số và, sau đó, tín hiệu hướng và tín hiệu khuếch tán được kết hợp để thu được tín hiệu âm thanh được tổng hợp cho ngăn thời gian hoặc tần số, sau đó, phép chuyển đổi tần số-thời gian được thực hiện bằng cách sử dụng các tín hiệu âm thanh cho các ngăn thời gian/tần số để thu được tín hiệu âm thanh tổng hợp miền thời gian làm trường âm thanh được biến đổi.

Do đó, nói chung, bộ tính trường âm thanh được tạo cấu hình để tổng hợp, cho mỗi nguồn âm thanh, và trường âm thanh liên quan đến vị trí tham chiếu khác bằng cách, ví dụ, xử lý, đối với mỗi nguồn, tín hiệu nguồn bằng cách sử dụng hướng mới cho tín hiệu âm thanh để thu được mô tả trường âm thanh của tín hiệu nguồn liên quan đến vị trí tham chiếu khác/mới. Hơn nữa, tín hiệu nguồn được biến đổi trước khi xử lý tín hiệu nguồn hoặc ngay sau khi xử lý tín hiệu nguồn bằng cách sử dụng thông tin

hướng. Và, cuối cùng, các mô tả trường âm thanh đối với các nguồn được bổ sung vào với nhau để thu được trường âm thanh được biến đổi liên quan đến vị trí tham chiếu khác.

Trong phương án khác, bộ tính trường âm thanh thực hiện, ngoài phép phân tích DirAC hoặc bất kỳ phép phân tích nguồn âm thanh nào khác, thuật toán tách nguồn. Cuối cùng, thuật toán tách nguồn dẫn đến, trong các tín hiệu nguồn âm thanh, ví dụ, trong miền thời gian hoặc trong miền tần số. Sau đó, tín hiệu khuếch tán được tính bằng cách trừ các tín hiệu nguồn âm thanh từ trường âm thanh ban đầu sao cho trường âm thanh ban đầu được phân tách thành tín hiệu khuếch tán và một số tín hiệu nguồn âm thanh, trong đó mỗi tín hiệu nguồn âm thanh có liên quan theo hướng nhất định.

Theo các phương án được ưu tiên, bộ tạo trường âm thanh tạo ra mô tả trường âm thanh thứ nhất đối với vị trí tham chiếu trong đó mô tả trường âm thanh thứ nhất chỉ bao gồm dữ liệu âm thanh từ các nguồn âm thanh nằm trong mô tả trường âm thanh thứ nhất xung quanh vị trí tham chiếu và bộ tạo trường âm thanh còn tạo thêm mô tả trường âm thanh thứ hai đối với vị trí tham chiếu trong đó mô tả trường âm thanh thứ hai chỉ có dữ liệu âm thanh từ nguồn thứ hai nằm trong âm lượng thứ hai xung quanh vị trí tham chiếu, âm lượng thứ hai khác với âm lượng thứ nhất. Mỗi âm lượng bao gồm một và tốt hơn là nhiều hơn một nguồn âm thanh, nhưng siêu dữ liệu được tạo cấu hình để cung cấp mô tả không gian của âm lượng thứ nhất và/hoặc âm lượng thứ hai mà sau đó có thể áp dụng cho tất cả các nguồn âm thanh trong âm lượng. Ví dụ, khi mô tả không gian là khoảng cách biểu diễn của lớp nhất định đến vị trí tham chiếu, thì khoảng cách này được sử dụng theo cùng một cách cho tất cả các nguồn trong lớp để xác định, cùng với sự ước lượng hướng đến tương ứng, vị trí của nguồn trong lớp. Tuy nhiên, vị trí trong lớp chỉ được xác định, đối với vị trí tham chiếu theo cách xuyên tâm, trong khi khoảng cách của mỗi nguồn trong cùng một lớp tương ứng với mô tả không gian hoặc, trong phương án cụ thể, khoảng cách biểu diễn liên quan đến âm lượng/lớp này. Do đó, phép biểu diễn siêu dữ liệu rất hiệu quả và nhỏ gọn thu được bao gồm, thông thường, trong một giá trị duy nhất cho số lượng nguồn tùy ý trong lớp và, ngoài ra, giống nhau cho tất cả các băng con của tín hiệu âm thanh. Ví

dụ, khi phép phân tích âm thanh của trường âm thanh lớp được sử dụng, thì thông tin khoảng cách sẽ là giống nhau cho mỗi băng con.

Ở đây cần lưu ý rằng mỗi âm lượng được biểu diễn bằng một lớp và tốt hơn là bởi một lớp riêng biệt để các âm lượng không trùng nhau.

Do đó, sáng chế đề cập đến mô tả trường âm thanh đa lớp có thông tin phụ rất nhỏ gọn trên một mặt, nhưng đối với mô tả trường âm thanh đơn lẻ, tốn nhiều chi phí hơn, vì mô tả trường âm hoàn chỉnh được cung cấp cho từng lớp. Ví dụ, khi một lớp đơn lẻ được biểu diễn bằng các tín hiệu định dạng B, thì cần bốn thành phần âm thanh cho mỗi lớp, tức là, thành phần đẳng hướng và ba thành phần định hướng. Do đó, khi trường âm thanh được biểu diễn bằng ba lớp, sau đó, nhìn chung, mười hai thành phần âm thanh, tức là, cần bốn thành phần trên mỗi lớp khi mỗi lớp được biểu diễn bằng các tín hiệu Ambisonics bậc nhất hoặc các tín hiệu định dạng B. Đương nhiên, các mô tả trường âm thanh khác nhau có thể được sử dụng cho các lớp khác nhau chẳng hạn như mô tả trường âm thanh bậc cao hơn cho lớp có số lượng nguồn lớn hơn so với mô tả trường âm thanh có số lượng thành phần nhỏ hơn cho lớp có số lượng nguồn thấp hơn.

Tuy nhiên, mặt khác, chỉ cần một thông tin hình học duy nhất làm siêu dữ liệu cho mỗi lớp trường âm thanh.

Trong các phương án được ưu tiên, các âm lượng là các hình cầu hoặc các vỏ hình cầu xung quanh vị trí tham chiếu trong khi, thông thường, lớp thấp nhất là hình cầu xung quanh vị trí tham chiếu và lớp cao hơn là vỏ hình cầu kéo dài xung quanh lớp thứ nhất được biểu diễn bằng hình cầu.

Tuy nhiên, âm lượng được biểu diễn bằng trường âm thanh không nhất thiết phải là hình cầu. Thay vào đó, âm lượng có thể là hình lập phương, phần tử song song hoặc bất kỳ dạng hình học ba chiều điển hình nào khác. Tuy nhiên, sáng chế cũng có thể được áp dụng trong tình huống hai chiều, do đó âm lượng được biểu diễn bằng khu vực và, thông thường, phần mở rộng nhỏ vô hạn trong chiều thứ ba. Do đó, thuật ngữ "âm lượng" không chỉ đề cập đến âm lượng ba chiều thực sự mà còn đề cập đến trường hợp hai chiều trong đó âm lượng trong trường hợp hai chiều là mặt phẳng có phần mở rộng nhỏ vô hạn theo hướng thứ ba. Do đó, "âm lượng" trong trường hợp hai chiều sẽ

là đường tròn xung quanh điểm tham chiếu đối với lớp thứ nhất và là vòng tròn về "âm lượng thứ nhất" có bán kính biểu diễn cao hơn bán kính biểu diễn của lớp thứ nhất.

Ngoài ra, thiết bị tạo mô tả trường âm thanh nâng cao không thể chỉ được tạo cấu hình như một loại bộ mã hóa tạo ra, từ trường âm thanh ban đầu, hai hoặc nhiều mô tả trường âm thanh được phân lớp và liên kết các mô tả không gian của âm lượng thứ nhất và/hoặc âm lượng thứ hai với các mô tả trường âm thanh này. Trong các phương án khác, thiết bị tạo mô tả trường âm thanh nâng cao cũng có thể được thực hiện như bộ chuyển mã nhận mô tả lớp với siêu dữ liệu và tạo mô tả được phân lớp mới với siêu dữ liệu mới. Ví dụ, khi siêu dữ liệu cho mỗi lớp được biểu diễn bằng khoảng cách biểu diễn đến điểm tham chiếu, và khi mô tả trường âm thanh nâng cao được chuyển mã có các lớp có cùng khoảng cách biểu diễn đến điểm tham chiếu (mới) khác nhau, thì siêu dữ liệu cho trường âm thanh nâng cao được tạo bởi bộ chuyển mã sẽ giống như siêu dữ liệu ban đầu, nhưng bộ chuyển mã sẽ tạo ra mô tả trường âm thanh được biến đổi cho mỗi lớp trong đó các hướng mới của các nguồn riêng lẻ được tính và, trong đó, ngoài ra, khoảng cách mới của nguồn âm thanh đến vị trí tham chiếu được tính bằng cách di chuyển nguồn âm thanh từ lớp này sang lớp khác và, tất nhiên, bằng cách làm giảm hoặc khuếch đại tín hiệu nguồn âm thanh cho nguồn âm thanh. Đặc biệt, sự suy giảm sẽ được cung cấp cho tín hiệu nguồn âm thanh, khi nguồn âm thanh chuyển từ lớp thấp hơn sang lớp cao hơn hoặc, thay vào đó, sự khuếch đại cho tín hiệu nguồn âm thanh được cung cấp, khi nguồn âm thanh chuyển từ lớp cao hơn sang lớp thấp hơn, nghĩa là gần hơn với vị trí tham chiếu mới.

Mỗi mô tả trường âm thanh cho mỗi lớp có thể được phân tích bằng bất kỳ kỹ thuật tách nguồn nào khác, ví dụ, có thể là bất kỳ kỹ thuật tách nguồn toàn dải nào mà không chỉ tạo ra tín hiệu nguồn từ mô tả lớp mà ngoài ra còn xác định hướng đến của nguồn này. Ngoài ra, mô tả trường âm thanh của nó cũng có thể được phân tích bởi bộ phân tích DirAC thực hiện việc phân tách nguồn được chọn theo tần số sao cho đối với mỗi ngăn thời gian/tần số, tín hiệu âm thanh nguồn cùng với, thông thường, giá trị độ khuếch tán được tính toán.

Tuy nhiên, khi tính toán khoảng cách của nguồn nhất định đến vị trí tham chiếu mới, thông tin khoảng cách thu được từ siêu dữ liệu cho mỗi lớp là giống nhau cho

mỗi nguồn được xác định từ mô tả lớp nhất định. Do đó, trong trường hợp phân tích bằng thông rộng, ví dụ, trong đó, hai hoặc nhiều nguồn có các hướng đến khác nhau đã được xác định từ mô tả lớp, thông tin khoảng cách là giống nhau cho mỗi nguồn.

Ngoài ra, khi mô tả trường âm thanh cho lớp được phân tích bởi bộ phân tích DirAC, thì thông tin khoảng cách cho mỗi ngăn thời gian/tần số sẽ lại giống nhau, nghĩa là sẽ bằng khoảng cách tham chiếu cho lớp tương ứng.

Trong trường hợp sáng chế được áp dụng như một bộ giải mã, nghĩa là, trong đó bộ tính trường âm thanh tính toán trường âm thanh được biến đổi ở dạng, ví dụ, phép biểu diễn thành phần định hướng đầy đủ chẳng hạn như phép biểu diễn Ambisonics, siêu dữ liệu chỉ được yêu cầu cho việc chia tỷ lệ các tín hiệu âm thanh tương ứng tùy thuộc vào khoảng cách cũ/mới đến vị trí tham chiếu. Sau đó, người ta có thể tính toán, cho mỗi nguồn âm thanh của mỗi lớp, phép biểu diễn Ambisonics nhất định và phép biểu diễn Ambisonics này được tính bằng cách sử dụng hướng đến mới được xác định từ hướng cũ của thông tin đến và thông tin tịnh tiến từ vị trí tham chiếu cũ sang vị trí tham chiếu mới và sau đó, mỗi tín hiệu nguồn sẽ được chia tỷ lệ để tính khoảng cách của nguồn từ vị trí tham chiếu trước đó đến vị trí tham chiếu mới và sau đó, theo tỷ lệ tương ứng, phép biểu diễn Ambisonics riêng của các nguồn có thể được đặt chồng lên nhau để có phép biểu diễn Ambisonics đầy đủ của trường âm thanh. Do đó, bộ giải mã như vậy sẽ được tạo cấu hình để chuyển phép biểu diễn lớp thành mô tả trường âm thanh đơn đối với vị trí tham chiếu mới mà sau đó có thể được xử lý thêm như chuyển đổi thành các tín hiệu loa, v.v.

Ngoài ra, bộ tính trường âm thanh có thể được tạo cấu hình để thực hiện phép biểu diễn loa phóng thanh cho thiết lập loa phóng thanh thực hoặc ảo dự định từ mỗi lớp riêng lẻ, ví dụ, tổng hợp DirAC của mô tả trường âm thanh của một lớp và sau đó, cho các tín hiệu loa phóng thanh riêng lẻ từ các lớp khác nhau có thể được thêm vào với nhau để cuối cùng tạo thành phép biểu diễn loa phóng thanh mà sau đó có thể được kết xuất bằng thiết lập loa phóng thanh được xác định trước này hoặc thậm chí có thể được chuyển đổi thành phép biểu diễn âm lập thể bằng bộ kết xuất âm lập thể.

Do đó, sáng chế có thể được sử dụng để tạo mô tả trường âm thanh phân lớp với siêu dữ liệu liên quan đến vị trí tham chiếu nhất định hoặc để tạo trường âm thanh

nâng cao được chuyển mã một lần nữa bằng cách sử dụng phép biểu diễn lớp, nhưng hiện tại với các lớp có liên quan đến vị trí tham chiếu mới, hoặc sáng chế có thể được áp dụng để giải mã phép biểu diễn lớp cộng với siêu dữ liệu thành mô tả trường âm thanh nhất định mà không có siêu dữ liệu liên quan đến vị trí tham chiếu mới.

Mô tả vắn tắt các hình vẽ

Các phương án được ưu tiên của sáng chế được mô tả ngay sau đây với việc tham chiếu đến các hình vẽ đi kèm, trong đó:

Fig.1a là phương án được ưu tiên của thiết bị tạo mô tả trường âm thanh nâng cao;

Fig.1b minh họa giải thích vấn đề mẫu trong sáng chế;

Fig.2 là việc triển khai được ưu tiên của thiết bị tạo mô tả trường âm thanh nâng cao;

Fig.3a minh họa mô tả trường âm thanh nâng cao bao gồm dữ liệu âm thanh, và thông tin phụ cho dữ liệu âm thanh;

Fig.3b minh họa thêm trường âm thanh nâng cao bao gồm dữ liệu âm thanh và siêu dữ liệu liên quan đến thông tin không gian chẳng hạn như thông tin hình học cho mỗi mô tả lớp;

Fig.4a minh họa việc triển khai thiết bị tạo mô tả trường âm thanh được biến đổi;

Fig.4b minh họa việc triển khai khác nữa của thiết bị tạo mô tả trường âm thanh được biến đổi;

Fig.4c minh họa trường hợp đa lớp;

Fig.4d minh họa phương án của bộ giải mã hoặc tùy chọn bộ chuyển mã 1;

Fig.4e minh họa bộ kết xuất cho đối tượng đơn lẻ;

Fig.4d minh họa phương án của bộ giải mã hoặc tùy chọn bộ chuyển mã 2;

Fig.4g minh họa một phần của bộ giải mã/bộ kết xuất;

Fig.4h minh họa một phần của bộ chuyển mã;

Fig.4i minh họa thiết bị mẫu để tạo mô tả trường âm thanh từ tín hiệu âm thanh chẳng hạn như tín hiệu đơn âm và dữ liệu hướng đến;

Fig.5 minh họa sáu tái tạo DoF của âm thanh không gian;

Fig.6 minh họa phương án được ưu tiên của tái tạo DoF thứ sáu bao gồm thiết bị tạo mô tả trường âm thanh nâng cao và thiết bị tạo mô tả trường âm thanh được biến đổi trên mỗi lớp;

Fig.7 minh họa việc triển khai được ưu tiên để tính toán DoA mới và khoảng cách mới của nguồn âm thanh đối với vị trí tham chiếu mới/khác nhau;

Fig.8 minh họa phương án xác định và áp dụng hệ số tỷ lệ;

Fig.9 minh họa một phương án được ưu tiên khác của tái tạo DoF thứ sáu bao gồm thiết bị tạo mô tả trường âm thanh nâng cao và thiết bị tạo mô tả trường âm thanh được biến đổi trên mỗi lớp trong ngữ cảnh DirAC trên mỗi lớp;

Fig.10 minh họa việc triển khai được ưu tiên của thiết bị tạo mô tả trường âm thanh được biến đổi;

Fig.11 minh họa việc triển khai được ưu tiên khác nữa của thiết bị tạo mô tả trường âm thanh được biến đổi;

Fig.11 minh họa việc triển khai được ưu tiên khác nữa của thiết bị tạo mô tả trường âm thanh được biến đổi;

Fig.12a minh họa việc triển khai phân tích DirAC đã có trong tình trạng kỹ thuật; và

Fig.12b minh họa việc triển khai tổng hợp DirAC đã có trong tình trạng kỹ thuật.

Mô tả chi tiết sáng chế

Để kích hoạt các ứng dụng 6DoF cho các phép biểu diễn Ambisonics/DirAC đã đề cập, cần phải mở rộng các phép biểu diễn này theo cách cung cấp thông tin còn thiếu để xử lý tịnh tiến. Cần lưu ý rằng sự mở rộng này có thể, ví dụ, 1) bổ sung khoảng cách hoặc các vị trí của các đối tượng vào phép biểu diễn cảnh hiện có,

và/hoặc 2) bổ sung thông tin mà có thể tạo thuận lợi cho quá trình phân tách các đối tượng riêng lẻ.

Ngoài ra, mục tiêu của các phương án là bảo toàn/tái sử dụng cấu trúc của các hệ thống Ambisonics (không tham số hoặc theo tham số) hiện có để cung cấp khả năng tương thích ngược với các phép biểu diễn/hệ thống này theo nghĩa là

- các phép biểu diễn được mở rộng có thể được tịnh tiến thành các phép biểu diễn không mở rộng hiện có (ví dụ, để kết xuất), và
- cho phép sử dụng lại các triển khai phần mềm và phần cứng hiện có khi làm việc với phép biểu diễn được mở rộng.

Sau đây, một số cách tiếp cận được mô tả, cụ thể là một cách tiếp cận hạn chế (nhưng rất đơn giản) và ba định dạng Ambisonics được mở rộng khác nhau để kích hoạt 6DoF.

Thay vì phép biểu diễn Ambisonics thông thường, nhiều tín hiệu Ambisonics (nghĩa là, các tập hợp đa tín hiệu) được xác định. Mỗi tín hiệu Ambisonics tương ứng với phạm vi khoảng cách cụ thể của cảnh âm thanh. Các phạm vi khoảng cách có thể bao gồm các phân vùng đồng nhất (ví dụ 0-1 mét, 1-2 mét, 2-3 mét, ...) hoặc các phân vùng không đồng nhất (ví dụ 0-1 mét, 1-3 mét, hơn 3 mét). Các phạm vi khoảng cách không chồng lấp có thể được xác định tĩnh hoặc động khi phản hồi lại các thuộc tính thực tế của cảnh âm thanh và được xác định theo định dạng Ambisonics nhiều lớp. Ngoài ra, các phạm vi khoảng cách chồng chéo cùng với các hàm cửa sổ có thể được xác định. Một trong các lớp có thể được phân bổ riêng cho các âm thanh khuếch tán/xung quanh mà không yêu cầu mô tả khoảng cách chính xác nhưng được phân phối trong toàn bộ cảnh âm thanh.

Khái niệm về phép biểu diễn Ambisonics đa lớp được áp dụng cho cả Ambisonics truyền thống và tham số (kiểu DirAC). Ví dụ với ba lớp được thể hiện trên Fig.4c.

Phần sau đây mô tả cách sử dụng Ambisonics đa lớp được đề xuất để triển khai các phép tịnh tiến.

Tùy chọn I (Fig.4d): Các đối tượng trong mỗi lớp Ambisonics được tạo bằng cách áp dụng kỹ thuật tách nguồn. Tín hiệu Ambisonics xung quanh / khuếch tán / dư cũng có thể được tạo ra. DOA thu được bằng cách sử dụng việc định vị nguồn 1D/2D, và khoảng cách được đưa ra bởi siêu dữ liệu của lớp. Lưu ý rằng trong nhiều trường hợp, DOA cũng có thể được trích xuất từ các bộ lọc tách nguồn, và các phương pháp tồn tại để thực hiện phân tách nguồn chung và ước lượng DOA.

Mỗi đối tượng kênh đơn sau đó được kết xuất thành một hoặc nhiều lớp như trên Fig.4e. Thứ nhất, DOA và khoảng cách sau khi tịnh tiến, được biểu diễn tương ứng là DOA' và Khoảng cách', được tính dựa trên thông tin tịnh tiến (ví dụ được biểu diễn dưới dạng vector) và khoảng cách lớp (ví dụ khoảng cách trung bình hoặc biểu diễn của lớp đó). Thứ hai, tín hiệu Ambisonics được tạo cho đối tượng dựa trên DOA' . Thứ ba, đối tượng được kết xuất tới (các) lớp thích hợp dựa trên Khoảng cách' và khoảng cách lớp mà từ đó đối tượng được trích xuất. Sau khi tạo các tín hiệu Ambisonics được biến đổi cho tất cả các đối tượng, lớp Ambisonics đã biến đổi thứ i ($i \in \{1, \dots, L\}$) được tính bằng cách thêm đầu ra Ambisonics đã biến đổi của Lớp i của tất cả các đối tượng. Tín hiệu Ambisonics xung quanh / khuếch tán / dư cho lớp thứ i được thêm trực tiếp vào đầu ra Ambisonics được biến đổi thứ i .

Ngoài ra, cái gọi là bộ lọc bù khoảng cách có thể được áp dụng để bù cho sự thay đổi khoảng cách. Bộ lọc có thể được áp dụng trực tiếp vào đối tượng dựa trên Khoảng cách' và khoảng cách lớp.

Tùy chọn II (Fig.4f): Các đối tượng được tạo bằng cách sử dụng tất cả các lớp L bằng cách áp dụng kỹ thuật tách nguồn. Cũng có thể tạo tín hiệu Ambisonics xung quanh / khuếch tán / dư đơn.

Mỗi đối tượng kênh đơn sau đó được kết xuất thành một hoặc nhiều lớp như được thể hiện trên Fig.4e. Thứ nhất, DOA và khoảng cách sau khi tịnh tiến, được biểu diễn tương ứng là DOA' và Khoảng cách', được tính dựa trên thông tin tịnh tiến (ví dụ được biểu diễn dưới dạng vector) và khoảng cách lớp (ví dụ khoảng cách trung bình hoặc biểu diễn của lớp đó). Thứ hai, tín hiệu Ambisonics được tạo cho đối tượng dựa trên DOA' . Thứ ba, đối tượng được kết xuất tới (các) lớp thích hợp dựa trên Khoảng cách' và khoảng cách lớp mà từ đó đối tượng được trích xuất. Ngoài ra, cái gọi là bộ

lọc bù khoảng cách có thể được áp dụng cho mỗi đối tượng để bù cho sự thay đổi khoảng cách. Sau khi tạo các tín hiệu Ambisonics được biến đổi cho tất cả các đối tượng, lớp Ambisonics đã biến đổi thứ i ($i \in \{1, \dots, L\}$) được tính bằng cách thêm đầu ra Ambisonics đã biến đổi của Lớp i của tất cả các đối tượng. Tín hiệu Ambisonics xung quanh / khuếch tán / dư cho lớp thứ i được thêm trực tiếp vào đầu ra Ambisonics được biến đổi thứ i .

Ngoài ra, cái gọi là bộ lọc bù khoảng cách có thể được áp dụng để bù cho sự thay đổi khoảng cách. Bộ lọc có thể được áp dụng trực tiếp vào đối tượng dựa trên 'Khoảng cách' và khoảng cách lớp.

Việc tạo tín hiệu Ambisonics đa lớp rất đơn giản cho nội dung do máy tính tạo và sản xuất. Việc chuyển đổi bản ghi tự nhiên thông qua các mảng micrô hoặc các micrô không gian (ví dụ micrô định dạng B) thành các tín hiệu Ambisonics đa lớp sẽ đơn giản hơn.

Các lớp có thể được chuyển đổi thành tín hiệu Ambisonics truyền thống bằng cách chiếu, tổng hợp, hoặc trộn giảm. Chuyển đổi đơn giản và có hiệu quả tính toán được thể hiện trên Fig.4g.

Có thể nhận thấy sự thay đổi trong định hướng bằng cách áp dụng việc xoay lớp độc lập cho từng tín hiệu Ambisonics đa lớp hoặc bằng cách áp dụng việc xoay đơn cho tín hiệu Ambisonics truyền thống. Việc xoay lớp độc lập có thể được thực hiện trước hoặc sau khi tịnh tiến.

Fig.1a minh họa thiết bị tạo mô tả trường âm thanh nâng cao bao gồm bộ tạo (mô tả) trường âm thanh 100 để tạo ít nhất một mô tả trường âm thanh biểu thị trường âm thanh đối với ít nhất một vị trí tham chiếu. Ngoài ra, thiết bị bao gồm bộ tạo siêu dữ liệu 110 để tạo siêu dữ liệu liên quan đến thông tin không gian của trường âm thanh. Siêu dữ liệu nhận, dưới dạng đầu vào, trường âm thanh hoặc thay vào đó hoặc ngoài ra, thông tin riêng biệt về các nguồn âm thanh.

Cả đầu ra của bộ tạo mô tả trường âm thanh 100 và bộ tạo siêu dữ liệu 110 tạo thành mô tả trường âm thanh nâng cao. Trong phương án, cả đầu ra của bộ tạo mô tả trường âm thanh 100 và bộ tạo siêu dữ liệu 110 có thể được kết hợp trong bộ kết hợp

120 hoặc giao diện đầu ra 120 để thu được mô tả trường âm thanh nâng cao mà bao gồm siêu dữ liệu không gian hoặc thông tin không gian của trường âm thanh như được tạo bởi bộ tạo siêu dữ liệu 110.

Fig.1b minh họa tình huống được giải quyết theo sáng chế. Ví dụ, vị trí A là ít nhất một vị trí tham chiếu và trường âm thanh được tạo bởi nguồn A và nguồn B và micro thực nhất định hoặc, ví dụ, micro ảo được đặt tại vị trí A phát hiện âm thanh từ nguồn A và nguồn B. Âm thanh là sự chồng chất của âm thanh đến từ các nguồn âm thanh phát ra. Điều này biểu diễn mô tả trường âm thanh như được tạo bởi bộ tạo mô tả trường âm thanh.

Ngoài ra, bộ tạo siêu dữ liệu sẽ, bằng các triển khai nhất định có được thông tin không gian đối với nguồn A và thông tin không gian khác đối với nguồn B chẳng hạn như các khoảng cách của các nguồn này tới vị trí tham chiếu chẳng hạn như vị trí A.

Đương nhiên, vị trí tham chiếu có thể, thay vào đó, là vị trí B. Sau đó, micro thực hoặc ảo sẽ được đặt ở vị trí B và mô tả trường âm thanh sẽ là trường âm thanh, ví dụ, được biểu diễn bằng các thành phần Ambisonics cấp một hoặc các thành phần Ambisonics cấp cao hơn hoặc bất kỳ thành phần âm thanh nào khác có khả năng mô tả trường âm thanh đối với ít nhất một vị trí tham chiếu, nghĩa là, vị trí B.

Sau đó, bộ tạo siêu dữ liệu có thể tạo ra, dưới dạng thông tin về các nguồn âm thanh, khoảng cách của nguồn âm A đến vị trí B hoặc khoảng cách của nguồn B đến vị trí B. Tất nhiên, thông tin thay thế về các nguồn âm thanh có thể là vị trí tuyệt đối hay tương đối đối với vị trí tham chiếu. Vị trí tham chiếu có thể là gốc của hệ tọa độ chung hoặc có thể được đặt trong mối quan hệ xác định với gốc của hệ tọa độ chung.

Siêu dữ liệu khác có thể là vị trí tuyệt đối của một nguồn âm thanh và vị trí tương đối của nguồn âm thanh khác đối với nguồn âm thanh thứ nhất, v.v.

Fig.2 minh họa thiết bị tạo mô tả trường âm thanh nâng cao, trong đó bộ tạo trường âm thanh bao gồm bộ tạo trường âm thanh 250 cho trường âm thanh thứ nhất, bộ tạo trường âm thanh 260 cho trường âm thanh thứ hai và, số lượng bộ tạo trường âm thanh tùy ý cho một hoặc nhiều trường âm thanh chẳng hạn như trường âm thanh thứ ba, thứ tư, v.v. Ngoài ra, siêu dữ liệu được tạo cấu hình để tính toán và chuyển tiếp

đến bộ kết hợp 120 thông tin về trường âm thanh thứ nhất và trường âm thanh thứ hai. Tất cả các thông tin này được sử dụng bởi bộ kết hợp 120 để tạo mô tả trường âm thanh nâng cao. Do đó, bộ kết hợp 120 cũng được tạo cấu hình thành giao diện đầu ra để tạo mô tả trường âm thanh nâng cao.

Fig.3a minh họa mô tả trường âm thanh nâng cao dưới dạng dòng dữ liệu bao gồm mô tả trường âm thanh thứ nhất 330, mô tả trường âm thanh thứ hai 340 và, kèm theo đó, siêu dữ liệu 350 bao gồm thông tin về mô tả trường âm thanh thứ nhất và mô tả trường âm thanh thứ hai. Mô tả trường âm thanh thứ nhất có thể, ví dụ, là mô tả định dạng B hoặc mô tả cấp cao hơn hoặc bất kỳ mô tả nào khác mà cho phép xác định sự phân phối hướng của các nguồn âm thanh trong phép biểu diễn toàn dải hoặc trong phép biểu diễn tần số được chọn. Do đó, mô tả trường âm thanh thứ nhất 330 và mô tả trường âm thanh thứ hai 340 cũng có thể, ví dụ, là những mô tả trường âm thanh tham số cho các vị trí tham chiếu khác nhau có, ví dụ, tín hiệu trộn giảm và dữ liệu hướng đến cho các ngăn thời gian/tần số khác nhau.

Tuy nhiên, thông tin hình học 350 cho các mô tả trường âm thanh thứ nhất và thứ hai là giống nhau cho tất cả các nguồn có trong mô tả trường âm thanh thứ nhất 330 hoặc, cho các nguồn trong mô tả trường âm thanh thứ hai 340, tương ứng. Do đó, ví dụ, khi tồn tại ba nguồn trong mô tả trường âm thanh thứ nhất 330 và thông tin hình học trên mô tả trường âm thanh thứ nhất thì thông tin hình học này giống nhau cho ba nguồn trong mô tả trường âm thanh thứ nhất. Tương tự, ví dụ, khi tồn tại năm nguồn trong mô tả trường âm thanh thứ hai, thì thông tin hình học cho trường âm thanh thứ hai có trong siêu dữ liệu 350 là giống nhau cho tất cả các nguồn trong mô tả trường âm thanh thứ hai.

Fig.3b minh họa cấu trúc kiểu mẫu của siêu dữ liệu 350 trên Fig.3a. Trong phương án, điểm tham chiếu 351 có thể có trong siêu dữ liệu. Tuy nhiên, điều này không nhất thiết là trường hợp trong thông tin điểm tham chiếu 351 cũng có thể được bỏ qua.

Đối với trường âm thanh thứ nhất, thông tin hình học thứ nhất được đưa ra mà, ví dụ, có thể là bán kính giữa hoặc bán kính biểu diễn của lớp thứ nhất, ví dụ, là giá trị 0,5 mét theo phương án mẫu của Fig.4c sẽ được mô tả sau.

Trường âm thanh thứ hai được mô tả bởi thông tin hình học thứ hai 353 mà, ví dụ, tương ứng với bán kính giữa của lớp thứ hai, chẳng hạn như hai mét cho phương án theo Fig.4c, vì lớp thứ hai kéo dài từ một mét đến ba mét.

Trường âm thanh thứ ba sẽ được mô tả bằng thông tin hình học thứ ba 354 và khoảng cách biểu diễn, ví dụ, là "bán kính giữa" của lớp thứ ba, chẳng hạn như bốn mét, v.v. Đương nhiên, mỗi mô tả trường âm thanh tốt hơn là bao gồm nhiều hơn một nguồn, nhưng có trường hợp cũng có thể là mô tả trường âm thanh của lớp nhất định chỉ bao gồm một nguồn đơn.

Ví dụ, khi Fig.4c được coi là minh họa cho phương pháp Ambisonics đa lớp được ưu tiên, lớp thứ nhất kéo dài từ 0 hoặc khoảng cách tối thiểu, ví dụ 0,5m đến 1m. Lớp thứ hai kéo dài từ 1m đến 3m và lớp thứ ba bao gồm tất cả các nguồn âm thanh có khoảng cách cao hơn 3m so với điểm tham chiếu thứ nhất là trung tâm của ba vòng tròn được minh họa trên Fig.4c.

Ngoài ra, Fig.4c minh họa là hai nguồn âm thanh 1, 2 có trong lớp 1, hai nguồn âm thanh 3, 4 có trong lớp 2 và các nguồn âm thanh 5 và 6 có trong lớp 3.

Như đã phác thảo, Fig.3a minh họa ví dụ về dòng dữ liệu hoặc dòng dữ liệu chung bao gồm mô tả trường âm thanh nâng cao. Dòng dữ liệu có thể bao gồm các khung thời gian i , $i+1$, v.v. và thông tin phụ có liên quan cho khung thời gian tương ứng.

Fig.4b minh họa một triển khai khác trong đó bộ phân tích trường âm thanh 210 trên Fig.2 không thực sự tạo ra trộn giảm, mà tạo ra phép biểu diễn đầy đủ ở định dạng B hoặc định dạng A hoặc định dạng bất kỳ khác chẳng hạn như phép biểu diễn bậc cao trong khoảng thời gian nhất định.

Fig.4c cũng minh họa vị trí người nghe mới. Khi các lớp tương tự được vẽ xung quanh vị trí người nghe mới, rõ ràng là nguồn âm thanh 1 di chuyển, do sự tịnh tiến sang vị trí người nghe mới, từ lớp 1 sang lớp 3 và do đó, phải bị suy yếu (mạnh).

Hơn nữa, nguồn âm thanh 2 di chuyển từ lớp 1 sang lớp 2 và do đó cũng nên bị suy yếu (yếu).

Hơn nữa, nguồn âm thanh 3 di chuyển từ lớp 2 trước đó sang lớp 3 mới và do đó, cũng phải bị suy yếu (yếu).

Ngoài ra, nguồn âm thanh 4 vẫn ở lớp thứ hai và do đó, không yêu cầu suy yếu. Hơn nữa, nguồn âm thanh 5 cũng vẫn nằm trong cùng một lớp và do đó, không cần phải được định tỷ lệ. Cuối cùng, nguồn âm thanh 6 được chuyển từ lớp thứ ba trước đó sang lớp thứ nhất mới và do đó, cần phải được khuếch đại (mạnh).

Do đó, nói chung, khi nguồn âm thanh di chuyển, do sự tịnh tiến từ điểm tham chiếu sang điểm tham chiếu (mới) khác nhau, thì hệ số tỷ lệ nhất định được xác định. Khi, "chuyển động" của nguồn âm thanh đi từ lớp cao hơn sang lớp thấp hơn, thì tỷ lệ là sự suy yếu và khi "chuyển động" đi từ lớp thấp hơn đến lớp cao hơn, thì tỷ lệ là sự suy yếu. Hơn nữa, khi "chuyển động" đi từ lớp này sang lớp kế tiếp, thì hệ số tỷ lệ sẽ là hệ số tỷ lệ yếu như suy giảm yếu hoặc khuếch đại yếu và khi việc di chuyển không đi từ lớp này sang lớp kế tiếp mà đi từ lớp này lớp này sang lớp khác không phải là lớp liền kề với một lớp, nghĩa là từ lớp thứ nhất sang lớp thứ ba hoặc ngược lại, khi đó hệ số tỷ lệ sẽ là hệ số tỷ lệ sao cho suy yếu hoặc khuếch đại mạnh hơn được thực hiện.

Ví dụ, khi nguồn di chuyển từ lớp thứ ba sang lớp thứ hai, và khi các giá trị mẫu của Fig.3b được xem xét, thì hệ số tỷ lệ sẽ nằm trong hệ số khuếch đại là 2,0, tức là, 4m chia cho 2m. Tuy nhiên, khi nguồn âm thanh di chuyển từ lớp thứ hai sang lớp thứ ba, thì hệ số tỷ lệ sẽ là 0,5, tức là, 2m chia cho 4m.

Thay vào đó, khi nguồn âm thanh di chuyển từ lớp thứ nhất sang lớp thứ ba, và giả sử các số liệu mẫu của Fig.3b, thì hệ số tỷ lệ sẽ là từ 0,1 đến 5 thu được bằng cách chia 0,5m cho 5m.

Thay vào đó, khi, nguồn âm thanh di chuyển từ trường âm thanh thứ ba sang trường âm thanh thứ nhất, chẳng hạn như nguồn âm thanh 6 theo phương án của Fig.4c, thì hệ số tỷ lệ sẽ là 8,0 thu được bằng cách chia 4m cho 0,5m.

Đương nhiên, mặc dù các ví dụ này đã được mô tả liên quan đến trường hợp "âm lượng" hai chiều được minh họa trên Fig.4c, cũng có các cân nhắc tương tự cho trường hợp âm lượng ba chiều thực khi các vòng tròn trên Fig.4c được coi là để biểu diễn hình cầu cho lớp 1 và vỏ hình cầu cho lớp 2 hoặc lớp 3.

Nói chung, Fig.4d minh họa việc triển khai bộ tính trường âm thanh cho thiết bị để tạo mô tả trường âm thanh được biến đổi hoặc, thay vào đó, cho thiết bị tạo mô tả trường âm thanh nâng cao trong bối cảnh biểu diễn bộ chuyển mã. Phép biểu diễn lớp thứ nhất được minh họa rõ ràng là lớp Ambisonics 1, mô tả trường âm thanh thứ hai được minh họa là lớp Ambisonics 2 và mô tả trường âm thanh bổ sung tùy chọn được minh họa ở lớp Ambisonics L được đưa vào các phân tách nguồn khác nhau và các khối ước lượng DoA 422, 422a, 422b. Ngoài ra, bộ kết xuất 450 được sử dụng để tạo cấu hình để hiển thị các đối tượng được tìm thấy bởi các thủ tục trong các khối 422, 422a, 422b cho các lớp "mới" tương ứng. Hơn nữa, các khối phân tích 422, 422a, 422b không chỉ được tạo cấu hình để trích xuất các tín hiệu đối tượng và hướng đến của các tín hiệu đối tượng mà còn được tạo cấu hình để tạo, từ mỗi mô tả trường âm thanh, tín hiệu xung quanh/khuếch tán/dur. Ví dụ, tín hiệu này có thể thu được bằng cách trừ tổng của tất cả các tín hiệu nguồn riêng lẻ có được từ thủ tục phân tách nguồn từ phép biểu diễn riêng lẻ. Ngoài ra, khi phân tách nguồn 422, 422a, 422b được tạo cấu hình, ví dụ, dưới dạng phân tích DirAC, thì tín hiệu khuếch tán/xung quanh/dur sẽ được biểu diễn, theo cách theo tham số, bằng các tham số khuếch tán.

Ngoài ra, Fig.4d minh họa bộ cộng 601, 602, 603 để thêm vào phép biểu diễn Ambisonics được biến đổi của lớp nhất định, tín hiệu khuếch tán tương ứng cho lớp.

Điều này được thực hiện cho từng lớp 1, lớp 2, lớp L và đầu ra trên Fig.4d một lần nữa biểu diễn cho mô tả trường âm nâng cao đã được chuyển mã, do các phép biểu diễn được biến đổi khác nhau được xuất ra bởi khối 601, 602, 603, siêu dữ liệu nhất định sẽ liên quan đến việc chỉ ra khoảng cách biểu diễn của lớp 1, lớp 2 và lớp L.

Fig.4e minh họa, đối với trường hợp trên Fig.4d, bộ kết xuất cho đối tượng hoặc nguồn đơn. Ví dụ, đối tượng thu được bằng khối 422 cho lớp thứ nhất hoặc khối 422a cho lớp thứ hai hoặc 422b cho lớp thứ L được nhập vào khối 430 để tính toán tín hiệu Ambisonics và ví dụ, khối 430 có thể được tạo cấu hình như minh họa trên Fig.4i. Để tính toán tín hiệu Ambisonics, DoA' mới, tức là giá trị hướng đến đối với nguồn được tạo bởi khối 423, 423a, 423b bởi DoA cũ, thông tin tịnh tiến từ vị trí tham chiếu ban đầu sang vị trí người nghe mới.

Do đó, tín hiệu Ambisonics mới cho đối tượng liên quan đến DoA' mới được tính toán và xuất ra dưới dạng đối tượng. Do đó, đầu ra của khối 430 của Fig.4i, chẳng hạn, sẽ là tín hiệu của đối tượng. Hơn nữa, khối 423, 423a, 423b sẽ tính toán/xác định khoảng cách mới như đã thảo luận trước đó đối với Fig.4c và do đó, ví dụ, sẽ tính toán các hệ số tỷ lệ như đã thảo luận trước đó đối với Fig.4c và ví dụ trên Fig.4b và sau đó, kết xuất cho các lớp khối 810 có thể được tạo cấu hình để chia tỷ lệ tín hiệu của đối tượng thu được từ khối 430 với hệ số tỷ lệ nhất định được suy ra từ khoảng cách cũ (khoảng cách lớp) và khoảng cách mới, sau đó, tất cả các tín hiệu thuộc về lớp 1, 2 hoặc L nhất định xung quanh vị trí người nghe mới của Fig.4c có thể được thêm vào tương ứng với nhau để một lần nữa có phép biểu diễn lớp. Do đó, đối với phương án của Fig.4c, chỉ có nguồn hiện tại ở lớp 1 sẽ là nguồn ban đầu 6, trong khi các nguồn ban đầu 5, 3, 1 sẽ được thêm vào cùng với các phép biểu diễn Ambisonics của chúng sau khi chia tỷ lệ như được xác định riêng lẻ đối với mỗi nguồn để thu được tín hiệu lớp 3 khi $L = 3$ và tương tự, tín hiệu nguồn cho các nguồn 2, 4 sẽ không biểu diễn cho tín hiệu lớp 2 như được minh họa trong bảng trên Fig.4c, nguồn 4 không yêu cầu chia tỷ lệ, nhưng nguồn 2 yêu cầu thao tác chia tỷ lệ.

Do đó, Fig.4e minh họa cách xác định các đối tượng cho mỗi lớp và sau đó thêm tín hiệu xung quanh/khuếch tán/dur bằng bộ cộng 601, 602, 603, thu được phép biểu diễn Ambisonics được biến đổi đầy đủ cho mỗi lớp, và sau đó có thể là đầu ra dưới dạng mô tả trường âm thanh nâng cao/được biến đổi cùng với siêu dữ liệu liên quan như được minh họa bởi Fig.4h.

Tuy nhiên, thay vào đó, phép biểu diễn Ambisonics cho mỗi lớp có thể được thêm vào một cách đơn giản bằng bộ cộng 600 được minh họa trên Fig.4g để chuyển đổi phép biểu diễn Ambisonics đa lớp thành phép biểu diễn Ambisonics truyền thống mà sau đó có thể được kết xuất một cách truyền thống thành phép biểu diễn thực hoặc phép biểu diễn loa phóng thanh thành phép biểu diễn âm lập thể.

Tùy chọn II khác được minh họa trên Fig.4f khác với Fig.4e là chỉ có tín hiệu xung quanh/khuếch tán/dur đơn được tạo ra từ tất cả các lớp riêng lẻ, và tín hiệu xung quanh/khuếch tán/dur đơn này chỉ được thêm bởi các bộ cộng 604 vào lớp cao nhất. Chẳng hạn, việc phân tách nguồn và ước tính DoA 422c có thể được thực hiện cho

từng mô tả trường âm thanh cho từng lớp riêng lẻ cho từng lớp như trên Fig.4d. Tuy nhiên, thuật toán phân tách nguồn và ước lượng DoA cũng có thể được triển khai để thêm tất cả các mô tả lớp riêng lẻ và sau đó thu được mô tả lớp tín hiệu và sau đó thực hiện phân tách nguồn và ước lượng DoA trên phép biểu diễn Ambisonics đơn này. Tuy nhiên, nên ưu tiên tiến hành riêng lẻ cho từng lớp và, để tạo tín hiệu khuếch tán đơn, các tín hiệu khuếch tán riêng lẻ thu được từ phương án của Fig.4d có thể được thêm vào trong khối 422c.

Do đó, tùy chọn II có thể dẫn đến các phép biểu diễn Ambisonics không hoàn toàn cho các lớp thấp hơn và chỉ trong phép biểu diễn "ướt" cho lớp cao nhất, tín hiệu khuếch tán đơn chỉ được thêm vào lớp cao nhất. Đương nhiên, quy trình này rất hữu ích, vì các tín hiệu khuếch tán dù sao không bị thu nhỏ bởi quy trình khoảng cách, mà được sử dụng theo cách tương tự như đã được xác định ban đầu bất kể mô tả trường âm thanh có liên quan đến vị trí tham chiếu ban đầu hay vị trí tham chiếu mới tương ứng, ví dụ, với vị trí người nghe của Fig.4c.

Fig.6 minh họa thuật toán tách nguồn cho từng lớp, ví dụ, có thể được tạo cấu hình dưới dạng phân tích DirAC. Sau đó, tín hiệu khuếch tán được minh họa trên Fig.6 được xuất ra bởi khối 422 sẽ không có ở đó, mà tín hiệu khuếch tán này sẽ được biểu diễn bằng tham số khuếch tán, tức là, sẽ được biểu diễn theo tham số. Tương ứng, tín hiệu đối tượng sẽ là, đối với quy trình DirAC, tín hiệu áp suất trên mỗi ngăn thời gian/tần số. Tuy nhiên, trong trường hợp chung, tín hiệu đối tượng cũng có thể là tín hiệu toàn dải.

Lớp thứ nhất được biểu diễn bởi nhánh trên, lớp thứ hai được biểu diễn bởi nhánh ở giữa Fig.6 và lớp thứ ba được biểu diễn bởi nhánh ở dưới cùng của Fig.6.

Trong trường hợp quy trình phân tách nguồn chung, bộ tổng hợp tương ứng 425 sẽ tạo tín hiệu Ambisonics cho mỗi lớp, ví dụ như được minh họa trên Fig.4d hoặc Fig.4f. Sau đó, tất cả các tín hiệu Ambisonics này có thể được kết hợp như đã thảo luận chung tham chiếu đến Fig.4g, và, sau đó, tín hiệu có thể được kết xuất bởi bộ kết xuất 427, ví dụ, trong trường hợp âm lập thể, khi tín hiệu được gửi đến ống nghe được áp dụng vào đầu của người nghe trong môi trường thực tế ảo được minh họa tại 500 trên Fig.6.

Fig.6 cũng chỉ ra thực tế rằng các quy trình phân tách nguồn, tịnh tiến, tạo các tín hiệu Ambisonics hoặc xoay có thể được áp dụng theo nhiều thứ tự khác nhau.

Trong khi Fig.6, lớp trên minh họa rằng việc phân tách nguồn được thực hiện trước khi chuyển đổi xoay và cả trước khi biến đổi tịnh tiến/chia tỷ lệ âm lượng, nhánh trung bình trên Fig.6 thể hiện việc tách nguồn 533a được thực hiện trước khi tịnh tiến 423a và việc tạo các tín hiệu Ambisonics được thực hiện bởi khối 425a, nhưng sau đó, tiếp theo bước này, phép biến đổi xoay được áp dụng trong khối 422a.

Một lần nữa, nhánh thấp hơn trên Fig.6 minh họa rằng biến đổi xoay 424b thậm chí có thể được áp dụng trước khi tách nguồn 422b, nhưng cũng có thể được áp dụng sau khi tách nguồn, vì chuyển động của đầu người nghe, dĩ nhiên, là giống nhau cho tất cả các nguồn của lớp và thậm chí cho tất cả các nguồn của tất cả các lớp. Hơn nữa, biến đổi tịnh tiến 423b cũng có thể được thực hiện trước khi tạo các tín hiệu Ambisonics và thậm chí trước khi chuyển đổi xoay, nhưng không phải trước khi tách nguồn, vì một trong hai phép tịnh tiến yêu cầu hướng đến của mỗi nguồn.

Trong trường hợp các tín hiệu Ambisonics ở đầu vào vào bộ kết hợp 600 và ngoài ra, khi đầu ra của bộ kết hợp 600 là tín hiệu Ambisonics truyền thống như, ví dụ, được thảo luận liên quan đến Fig.4g, thì biến đổi xoay thậm chí có thể được áp dụng sau khối 600 và sau đó, bộ kết xuất 427 sẽ phải kết xuất trực tiếp thành phép biểu diễn âm lập thể hoặc trực tiếp thành phép biểu diễn loa phóng thanh hoặc thành phép biểu diễn âm lập thể thông qua phép biểu diễn loa phóng thanh ảo. Tuy nhiên, mô tả trường âm thanh được biến đổi tịnh tiến đầy đủ và xoay hoàn toàn không còn được phân lớp nữa mà "các đặc điểm" của lớp được bỏ do sự kết hợp trong bộ kết hợp 600.

Trong bối cảnh này, rõ ràng là từ Fig.6 và Fig.4e, trong trường hợp không chỉ chuyển mã tín hiệu mà còn trong trường hợp tạo ra mô tả trường âm thanh đơn, "kết xuất các lớp 810" được thảo luận trong bối cảnh theo bảng trên Fig.4c là không cần thiết theo cách được minh họa đối với Fig.4c. Vì chỉ có trường âm thanh đơn được tạo ra nên không có vấn đề gì, dù nguồn có chuyển từ lớp này sang lớp khác hay không. Thay vào đó, chỉ cần các hệ số tỷ lệ, nhưng các hệ số tỷ lệ này có thể được suy ra trực tiếp từ các khoảng cách biểu diễn cho các lớp hoặc từ khoảng cách cũ được đưa ra bởi thông tin lớp 350 và khoảng cách mới có được bằng cách sử dụng khoảng cách cũ,

DoA của nguồn và thông tin tịnh tiến như đã thảo luận, ví dụ, liên quan đến Fig.7. Do đó, cho dù nguồn thay đổi hay "chuyển" từ một lớp sang lớp khác chỉ là vấn đề đối với phép biểu diễn bộ chuyển mã, mà không phải là vấn đề trong trường hợp của Fig.4g hoặc Fig.6.

Fig.4a minh họa triển khai được ưu tiên của thiết bị tạo mô tả trường âm thanh được biến đổi từ mô tả trường âm thanh và siêu dữ liệu liên quan đến thông tin không gian của mô tả trường âm thanh. Đặc biệt, thiết bị bao gồm bộ tính trường âm thanh 420 mà tạo trường âm thanh được biến đổi sử dụng siêu dữ liệu, mô tả trường âm thanh và thông tin tịnh tiến chỉ ra sự tịnh tiến từ vị trí tham chiếu tới vị trí tham chiếu khác.

Ví dụ, khi trường âm thanh được cung cấp đối với vị trí A trên Fig.1b, thì vị trí tham chiếu khác có thể là vị trí B và thông tin tịnh tiến sẽ là, ví dụ, vector biểu thị việc tịnh tiến vị trí A sang vị trí B. Sau đó, bộ tính trường âm thanh 420 sẽ tính toán trường âm thanh được biến đổi như người nghe cảm nhận được đặt ở vị trí B, và đối với phép tính này, bộ tính trường âm thanh có mô tả trường âm thanh liên quan đến vị trí A và thông tin tịnh tiến và, ngoài ra, siêu dữ liệu liên quan đến vị trí không gian của nguồn A và nguồn B.

Trong phương án, bộ tính trường âm thanh 420 được kết nối với giao diện đầu vào 400 để nhận mô tả trường âm thanh nâng cao như, ví dụ, đã thảo luận liên quan đến Fig.1a hoặc Fig.2 và giao diện đầu vào 400 sau đó phân tách mô tả trường âm thanh trên một mặt, nghĩa là, những gì đã được tạo ra bởi khối 100 trên Fig.1a hoặc khối 210 trên Fig.2. Hơn nữa, giao diện đầu ra 400 phân tách siêu dữ liệu từ mô tả trường âm thanh nâng cao, nghĩa là, mục 350 trên Fig.3a hoặc tùy chọn 351 và từ 352 tới 354 trên Fig.3b.

Hơn nữa, giao diện tịnh tiến 410 thu thông tin tịnh tiến và/hoặc thông tin xoay bổ sung hoặc riêng biệt từ người nghe. Việc triển khai của giao diện tịnh tiến 410 có thể là bộ phận theo dõi đầu không chỉ theo dõi sự xoay của đầu trong môi trường thực tế ảo, mà còn là sự tịnh tiến của đầu từ một vị trí, nghĩa là, vị trí A trên Fig.1b tới vị trí khác, nghĩa là, vị trí B trên Fig.1b.

Fig.4b minh họa triển khai khác tương tự như trên Fig.1a, nhưng không liên quan đến trường hợp bộ mã hóa/bộ giải mã, mà liên quan đến trường hợp chung khi việc cung cấp siêu dữ liệu được chỉ định bởi bộ cung cấp trường âm thanh 402, việc cung cấp trường âm thanh được chỉ định bởi bộ cung cấp trường âm thanh 404 được thực hiện mà không có giao diện đầu vào nhất định phân tách mô tả trường âm thanh được mã hóa hoặc nâng cao, nhưng tất cả đều được thực hiện, ví dụ, trong trường hợp thực tế hiện có, ví dụ, trong ứng dụng thực tế ảo. Tuy nhiên, sáng chế không bị giới hạn bởi các ứng dụng thực tế ảo, mà cũng có thể được triển khai trong bất kỳ ứng dụng thực tế ảo nào khác, trong đó việc xử lý âm thanh không gian của các trường âm thanh mà liên quan đến vị trí tham chiếu là hữu ích để biến đổi trường âm thanh liên quan đến vị trí tham chiếu thứ nhất tới trường âm thanh khác liên quan đến vị trí tham chiếu thứ hai khác.

Bộ tính toán trường âm thanh 420 sau đó tạo mô tả trường âm thanh được biến đổi hoặc, thay vào đó, tạo phép biểu diễn loa phóng thanh (ảo) hoặc tạo phép biểu diễn âm lập thể chẳng hạn như phép biểu diễn hai kênh để tái tạo ống nghe. Do đó, bộ tính toán trường âm thanh 420 có thể tạo, như trường âm thanh được biến đổi, mô tả trường âm thanh được biến đổi, về cơ bản giống như mô tả trường âm thanh ban đầu, nhưng bây giờ đối với vị trí tham chiếu mới. Trong phương án thay thế, phép biểu diễn loa phóng thanh thực hoặc ảo có thể được tạo cho thiết lập loa phóng thanh được xác định trước chẳng hạn như sơ đồ 5.1 hoặc thiết lập loa phóng thanh có nhiều loa phóng thanh hơn và, đặc biệt, có bố trí các loa phóng thanh ba chiều thay vì chỉ bố trí hai chiều, nghĩa là, sự bố trí loa phóng thanh có các loa phóng thanh được nâng lên đối với vị trí người dùng. Các ứng dụng khác mà đặc biệt hữu ích cho các ứng dụng thực tế ảo là các ứng dụng để tái tạo âm lập thể, nghĩa là, cho ống nghe mà có thể được áp dụng cho đầu của người dùng thực tế ảo.

Ví dụ, Fig.6 hoặc Fig.9 được mô tả sau đó minh họa tình huống, trong đó bộ tổng hợp DirAC chỉ hoạt động trên thành phần trộn giảm chẳng hạn như thành phần đẳng hướng hoặc áp lực, trong khi, trong phương án thay thế khác được minh họa đối với Fig.12b, bộ tổng hợp DirAC hoạt động trên toàn bộ dữ liệu trường âm thanh, nghĩa

là, phép biểu diễn thành phần đầy đủ có trong phương án này trên Fig.12b, mô tả trường với thành phần đẳng hướng w và ba thành phần định hướng x, y, z .

Fig.4i minh họa triển khai khác để thực hiện phép tổng hợp khác với bộ tổng hợp DirAC. Ví dụ, khi bộ phân tích trường âm thanh tạo ra, cho mỗi tín hiệu nguồn, tín hiệu đơn riêng biệt S và hướng đến ban đầu và khi, phụ thuộc vào thông tin tĩnh tiến, hướng đến mới được tính toán, sau đó bộ tạo tín hiệu Ambisonics 430 trên Fig.4i, ví dụ, sẽ được sử dụng để tạo mô tả trường âm thanh cho tín hiệu nguồn âm thanh, nghĩa là, tín hiệu đơn S nhưng cho dữ liệu hướng đến (DoA) mới bao gồm góc ngang θ hoặc góc độ cao θ và góc phương vị ϕ . Sau đó, quy trình được thực hiện bởi bộ tính toán trường âm thanh 420 trên Fig.4b sẽ tạo ra, ví dụ, mô tả trường âm thanh Ambisonics cấp một cho mỗi nguồn âm thanh với hướng đến mới và, sau đó, biến đổi khác nữa trên mỗi nguồn âm thanh có thể được thực hiện bằng cách sử dụng hệ số tỷ lệ tùy thuộc vào khoảng cách của nguồn âm thanh tới vị trí tham chiếu mới và, sau đó, tất cả trường âm thanh từ các nguồn riêng biệt có thể chồng lên nhau để cuối cùng thu được trường âm thanh được biến đổi, một lần nữa, ví dụ, trong phép biểu diễn Ambisonics liên quan đến vị trí tham chiếu mới nhất định.

Khi một người giải thích rằng mỗi ngăn thời gian/tần số được xử lý bởi bộ phân tích DirAC 422, 422a, 422b trên Fig.6 hoặc Fig.9 biểu diễn nguồn âm thanh nhất định (giới hạn băng thông), thì có thể sử dụng bộ tạo tín hiệu Ambisonics 430, thay vì bộ tổng hợp DirAC 425, 425a, 425b để tạo, đối với mỗi ngăn thời gian/tần số, phép biểu diễn Ambisonics đầy đủ sử dụng tín hiệu trộn giảm hoặc tín hiệu áp lực hoặc thành phần đẳng hướng cho ngăn thời gian/tần số này như là "tín hiệu đơn S " trên Fig.4i. Sau đó, sự biến đổi tần số-thời gian riêng biệt trong bộ biến đổi tần số-thời gian cho mỗi thành phần trong số các thành phần W, X, Y, Z sau đó sẽ dẫn đến mô tả trường âm thanh khác với những gì được minh họa trên Fig.4c.

Các phương án khác nữa được phác thảo sau đây. Mục tiêu là thu được tín hiệu âm lập thể ảo ở vị trí nghe của người nghe được cung cấp tín hiệu ở vị trí ghi ban đầu và thông tin về các khoảng cách của các nguồn âm thanh từ vị trí ghi. Các nguồn vật lý được giả định là có thể tách rời theo góc của chúng đối với vị trí ghi.

Cảnh được ghi lại từ điểm nhìn (PoV) của micrô là vị trí được sử dụng làm gốc của hệ tọa độ tham chiếu. Cảnh phải được tái tạo từ PoV của người nghe là người được theo dõi trong 6DoF, xem Fig.5. Nguồn âm thanh đơn được hiển thị ở đây để minh họa, mỗi quan hệ giữ cho mỗi ngăn tần số-thời gian.

Fig.5 minh họa bước tái tạo 6DoF của âm thanh không gian. Nguồn âm thanh được ghi lại bởi micrô với DoA $\mathbf{r}_r$ trong khoảng cách d_r tương đối với vị trí và định hướng của các micrô (đường màu đen và đường cung). Nó phải được tái tạo tương ứng với người nghe di chuyển với DoA $\mathbf{r}_l$ và khoảng cách d_l (nét đứt). Điều này phải xem xét phép tịnh tiến $\mathbf{l}$ và phép xoay $\mathbf{o}$ của người nghe (chấm).

Nguồn âm thanh tại tọa độ $\mathbf{d}_r \in \mathbb{R}^3$ được ghi từ hướng đến (DoA) được biểu thị bằng vector đơn vị $\mathbf{r}_r = \mathbf{d}_r / \|\mathbf{d}_r\|$. DoA này có thể được ước lượng từ phép phân tích ghi lại. Nó đến từ khoảng cách $d_r = \|\mathbf{d}_r\|$. Giả sử rằng thông tin này có trong siêu dữ liệu cho mỗi lớp dưới dạng khoảng cách lớp hoặc, nói chung, như mục 352, 353, 354 của Fig.3b, và có thể được biểu thị dưới dạng thông tin khoảng cách theo dạng bản đồ độ sâu $m(l, r)$ ánh xạ mỗi nguồn của một lớp l có bất kỳ hướng r nào từ vị trí ghi đến khoảng cách (ví dụ tính bằng mét hoặc tương tự) của lớp từ điểm tham chiếu.

Người nghe được theo dõi trong 6DoF. Tại một thời điểm nhất định, người đó ở vị trí $\mathbf{l} \in \mathbb{R}^3$ so với micrô và có góc xoay $\mathbf{o} \in \mathbb{R}^3$ so với hệ tọa độ của micrô. Vị trí ghi được chọn làm gốc của hệ tọa độ để đơn giản hóa ký hiệu.

Do đó âm thanh phải được tái tạo với khoảng cách d_l khác dẫn đến âm lượng thay đổi, và khác DoA $\mathbf{r}_l$ mà là kết quả của cả phép tịnh tiến và phép xoay tiếp theo.

Phương pháp để thu tín hiệu ảo từ góc nhìn của người nghe bởi các phép biến đổi chuyên dụng dựa trên phép biểu diễn tham số, như được giải thích trong phần dưới đây, được phác thảo.

Phương pháp được đề xuất dựa trên tiếp cận DirAC cơ bản để mã hóa âm thanh không gian tham số, xem tài liệu [16]. Giả sử rằng có một nguồn có hướng chi phối trên mỗi trường hợp thể hiện thời gian-tần số của phổ được phân tích và chúng có thể được xử lý độc lập. Bản ghi được biến đổi thành phép biểu diễn thời gian-tần số bằng cách sử dụng biến đổi Fourier thời gian ngắn (short time Fourier transform - STFT).

Chỉ số khung thời gian được ký hiệu là n và chỉ số tần số được ký hiệu là k . Bản ghi được biến đổi sau đó được phân tích, ước tính các hướng $r_r(k, n)$ và độ khuếch tán $\psi(k, n)$ cho mỗi ngăn thời gian-tần số của phổ phức tạp $P(k, n)$. Trong quá trình tổng hợp, tín hiệu được chia thành phần có hướng và phần khuếch tán. Ở đây, tín hiệu loa phóng thanh được tính toán bằng cách quét phần có hướng tùy thuộc vào các vị trí loa phóng thanh và bổ sung phần khuếch tán.

Phương pháp để biến đổi bản ghi tín hiệu FOA đến phối cảnh người nghe trong 6DoF có thể được chia thành năm bước, xem Fig.6.

Fig.6 minh họa phương pháp tái tạo 6DoF. Tín hiệu FOA được ghi trong định dạng B được xử lý bởi bộ mã hóa DirAC mà tính toán các giá trị hướng và độ khuếch tán cho mỗi ngăn thời gian-tần số của phổ phức tạp. Vector hướng sau đó được biến đổi bởi vị trí được theo dõi của người nghe và theo thông tin khoảng cách có trong bản đồ khoảng cách cho mỗi lớp. Vector hướng kết quả sau đó được xoay theo sự xoay đầu. Cuối cùng, các tín hiệu cho các kênh loa phóng thanh ảo 8+4 được tổng hợp trong bộ giải mã DirAC. Sau đó chúng được lập thể.

Trong phương án, tín hiệu đầu ra được phân tích trong bộ mã hóa DirAC 422, thông tin khoảng cách được bổ sung từ bản đồ khoảng cách $m(l)$ đưa ra khoảng cách cho mỗi lớp, sau đó phép tịnh tiến và phép xoay được theo dõi của người nghe được áp dụng trong các biến đổi mới 423 và 424. Bộ giải mã DirAC 425 tổng hợp các tín hiệu cho các loa phóng thanh ảo 8+4, lần lượt được lập thể 427 để phát lại tai nghe. Lưu ý rằng do phép xoay của cảnh âm thanh sau phép tịnh tiến được hoạt động độc lập, nó có thể được áp dụng thay thế trong bộ kết xuất lập thể. Tham số duy nhất được biến đổi cho 6DoF là vector hướng. Theo định nghĩa mô hình, phần khuếch tán được giả định là đẳng hướng và đồng nhất và do đó được giữ nguyên.

Đầu vào của bộ mã hóa DirAC là tín hiệu âm thanh FOA trong phép biểu diễn định dạng B. Nó bao gồm bốn kênh, tức là, áp lực âm thanh đẳng hướng và ba gradien không gian cấp một, theo các giả định nhất định tỷ lệ thuận với vận tốc của hạt. Tín hiệu này được mã hóa theo phương pháp tham số, xem tài liệu [18]. Các tham số được suy ra từ áp lực âm thanh phức tạp $P(k, n)$, mà là tín hiệu đẳng hướng được biến đổi và vector vận tốc hạt phức tạp $U(k, n) = [U_x(k, n), U_y(k, n), U_z(k, n)]^T$

tương ứng với các tín hiệu gradient được biến đổi.

Phép biểu diễn DirAC bao gồm tín hiệu $P(k, n)$, độ khuếch tán $\psi(k, n)$ và hướng $\mathbf{r}(k, n)$ của sóng âm thanh tại mỗi ngăn thời gian-tần số. Để suy ra cái sau, đầu tiên, vector cường độ âm thanh hoạt động $I_a(k, n)$ được tính là phần thực (ký hiệu là $\text{Re}(\cdot)$) của sản phẩm của vector áp lực với liên hợp phức tạp (ký hiệu là $(\cdot)^*$) của vector vận tốc [18]:

$$I_a(k, n) = \frac{1}{2} \text{Re}(P(k, n) \mathbf{U}^*(k, n)). \quad (1)$$

Độ khuếch tán được ước tính từ hệ số biến thiên của vector này [18]:

$$\psi(k, n) = \sqrt{1 - \frac{\|E\{I_a(k, n)\}\|^2}{E\{\|I_a(k, n)\|^2\}}} \quad (2)$$

trong đó E biểu thị toán tử kỳ vọng dọc theo các khung thời gian, được triển khai dưới dạng số trung bình động.

Do dự định điều khiển âm thanh bằng cách sử dụng bản đồ khoảng cách dựa trên hướng có khoảng cách cho mỗi nguồn của lớp đến vị trí tham chiếu, phương sai của các ước tính hướng sẽ thấp. Vì các khung thường ngắn, điều này không phải lúc nào cũng đúng. Do đó, số trung bình động được áp dụng để thu được ước tính hướng làm nhãn $\bar{I}_a(k, n)$. DoA của phần trực tiếp của tín hiệu sau đó được tính toán là vector độ dài đơn vị theo hướng ngược lại:

$$\mathbf{r}_r(k, n) = - \frac{\bar{I}_a(k, n)}{\|\bar{I}_a(k, n)\|} \quad (3)$$

Vì hướng được mã hóa dưới dạng vector ba chiều có độ dài đơn vị cho mỗi ngăn thời gian-tần số, nên việc tích hợp thông tin khoảng cách là đơn giản. Các vector chỉ hướng được nhân với mục nhập bản đồ tương ứng của chúng sao cho độ dài vector biểu diễn khoảng cách của nguồn âm thanh tương ứng $d_r(k, n)$:

$$\begin{aligned} \mathbf{d}_r(k, n) &= \mathbf{r}_r(k, n) d_r(k, n) \\ &= \mathbf{r}_r(k, n) m(l, \mathbf{r}_r(k, n)) \end{aligned}$$

(4)

trong đó $\mathbf{d}_r(k, n)$ là vector chỉ từ vị trí ghi của micrô đến nguồn âm thanh hoạt động tại thời điểm n và tần số k .

Vị trí người nghe được đưa ra bởi hệ thống theo dõi cho khung xử lý hiện thời là $\mathbf{l}(n)$. Với phép biểu diễn vector của các vị trí nguồn, có thể trừ đi vector vị trí theo dõi $\mathbf{l}(n)$ để tạo ra vector hướng được tịnh tiến mới $\mathbf{d}_1(k, n)$ với độ dài $d_1(k, n) = \|\mathbf{d}_1(k, n)\|$, xem Fig.7. Các khoảng cách từ PoV của người nghe đến các nguồn âm thanh được suy ra, và các DoA được điều chỉnh theo một bước duy nhất:

$$\mathbf{d}_1(k, n) = \mathbf{d}_r(k, n) - \mathbf{l}(n) \quad (5)$$

Khía cạnh quan trọng của bước tái tạo thực là sự giảm dần khoảng cách. Sự giảm dần được giả định là hàm của khoảng cách giữa nguồn âm thanh và người nghe [19]. Độ dài của các vector chỉ hướng là để mã hóa sự giảm dần hoặc khuếch đại sự tái tạo. Khoảng cách đến vị trí ghi được mã hóa theo $\mathbf{d}_r(k, n)$ theo bản đồ khoảng cách, và khoảng cách được tái tạo được mã hóa theo $\mathbf{d}_1(k, n)$. Nếu bình thường hóa các vector thành chiều dài đơn vị và sau đó nhân với tỷ lệ khoảng cách mới và cũ, sẽ thấy rằng độ dài cần thiết được đưa ra bằng cách chia $\mathbf{d}_1(k, n)$ cho độ dài của vector ban đầu:

$$\mathbf{d}_v(k, n) = \frac{\mathbf{d}_1(k, n)}{\|\mathbf{d}_1(k, n)\|} \frac{\|\mathbf{d}_1(k, n)\|}{\|\mathbf{d}_r(k, n)\|} = \frac{\mathbf{d}_1(k, n)}{\|\mathbf{d}_r(k, n)\|} \quad (6)$$

Những thay đổi cho sự định hướng của người nghe được áp dụng trong bước sau. Sự định hướng được đưa ra bởi sự theo dõi có thể được viết dưới dạng vector bao gồm cường độ, độ chệch và cuộn $\mathbf{o}(n) = [\mathbf{o}_x(n), \mathbf{o}_z(n), \mathbf{o}_y(n)]$ so với vị trí ghi là gốc. Hướng nguồn được xoay theo định hướng người nghe, mà được triển khai bằng cách sử dụng ma trận xoay 2D:

$$\mathbf{d}_p(k, n) = \mathbf{R}_y(\mathbf{o}_y(n))\mathbf{R}_z(\mathbf{o}_z(n))\mathbf{R}_x(\mathbf{o}_x(n))\mathbf{d}_v(k, n) \quad (7)$$

DoA kết quả cho người nghe sau đó được đưa ra bởi vector chuẩn hóa theo chiều dài đơn vị:

$$\mathbf{r}_p(k, n) = \frac{\mathbf{d}_p(k, n)}{\|\mathbf{d}_p(k, n)\|}$$

(8)

Vectơ hướng được biến đổi, độ khuếch tán và phổ phức tạp được sử dụng để tổng hợp các tín hiệu cho thiết lập loa phóng thanh ảo 8+4 được phân bố đồng đều. Tám loa ảo được đặt trong các bước góc 45° trên mặt phẳng người nghe (độ cao 0°), và bốn loa phóng thanh trong dạng chéo 90° ở trên độ cao 45° . Phép tổng hợp được chia thành phần có hướng và phần khuếch tán cho mỗi kênh loa phóng thanh $1 \leq i \leq I$, trong đó $I = 12$ là số lượng loa phóng thanh [16]:

$$Y_i(k, n) = Y_{i,S}(k, n) + Y_{i,D}(k, n) \quad (9)$$

Đối với phần có hướng, phép quét biên độ làm mờ biên (edge fading amplitude panning - EFAP) được áp dụng để tái tạo âm thanh từ hướng đúng cho dạng hình loa phóng thanh ảo [20]. Với vectơ DoA $\mathbf{r}_p(k, n)$, điều này cung cấp độ khuếch đại quét $G_i(\mathbf{r})$ cho mỗi kênh loa phóng thanh ảo i . Độ khuếch đại phụ thuộc vào khoảng cách cho mỗi DoA được suy ra từ chiều dài kết quả của vectơ chỉ hướng $\mathbf{d}_p(k, n)$. Phép tổng hợp hướng cho kênh trở i thành:

$$Y_{i,S}(k, n) = \sqrt{1 - \psi(k, n)} P(k, n) G_i(\mathbf{r}_p(k, n)) (\|\mathbf{d}_p(k, n)\|)^{-\gamma} \quad (10)$$

trong đó số mũ γ là thừa số điều chỉnh thường được đặt thành khoảng 1 [19]. Lưu ý rằng với $\gamma = 0$ độ khuếch đại phụ thuộc vào khoảng cách bị tắt.

Áp lực $P(k, n)$ được sử dụng để tạo I tín hiệu giải tương quan $\tilde{P}_i(k, n)$. Các tín hiệu giải tương quan này được bổ sung vào các kênh loa phóng thanh riêng lẻ dưới dạng thành phần khuếch tán. Điều này tuân theo phương pháp tiêu chuẩn [16]:

$$Y_{i,D}(k, n) = \sqrt{\psi(k, n)} \frac{1}{\sqrt{I}} \tilde{P}_i(k, n) \quad (11)$$

Phần khuếch tán và có hướng của mỗi kênh được bổ sung với nhau, và các tín hiệu được biến đổi trở lại miền thời gian bởi STFT nghịch đảo. Các tín hiệu miền

thời gian kênh này được tích hợp với các HRTF cho tai trái và tai phải tùy thuộc vào vị trí loa để tạo các tín hiệu lập thể.

Fig.8 minh họa triển khai được ưu tiên của việc thực hiện chuyển mã mà hoạt động, ví dụ, không có khoảng cách biểu diễn, mà với các phép xác định chung về các hệ số tỷ lệ không được suy ra từ khoảng cách của các lớp mà chỉ từ việc nguồn có di chuyển từ lớp này sang lớp khác là lớp liền kề hoặc không liền kề với lớp trước đó hay không. Do đó, khối 800 được tạo cấu hình để xác định, liệu nguồn có còn ở trong cùng một lớp hay không. Nếu được xác định rằng là cùng một lớp là kết quả như, ví dụ, đối với các nguồn 4 và 5, thì khối 802 sẽ xác định rằng không cần tỷ lệ cụ thể của tín hiệu đối tượng cho hướng đến mới và "hệ số tỷ lệ" biểu thị không có tỷ lệ nào được đặt thành 1 hoặc, thay vào đó, nó có thể được đánh dấu hoặc được thông báo theo bất kỳ cách nào mà đối với nguồn như vậy, bất kỳ tỷ lệ nào sẽ không được thực hiện.

Tuy nhiên, khi xác định rằng nguồn di chuyển đến lớp thấp hơn, thì khối 804 sẽ xác định hệ số tỷ lệ là lớn hơn 1. Việc xác định này có thể được thực hiện bằng cách sử dụng các khoảng cách biểu diễn của hai lớp, tức là, lớp nguồn và lớp đích. Tuy nhiên, hệ số tỷ lệ nhất định cũng có thể được sử dụng, ví dụ, giữa 1 và 2 khi việc di chuyển diễn ra từ một lớp sang lớp liền kề và, ví dụ, bằng 4 khi nguồn di chuyển qua hai lớp, v.v.

Fig.8 minh họa thêm tình huống tại khối 806 trong đó nguồn di chuyển đến lớp cao hơn, ví dụ như nguồn 1, 2, 3. Sau đó, hệ số tỷ lệ cho nguồn này sẽ được xác định là thấp hơn 1. Một lần nữa, hệ số tỷ lệ có thể được xác định từ khoảng cách biểu diễn như đã được thảo luận trước đó, nhưng cũng có thể, thay vào đó, thu được bằng cách sử dụng các hệ số tỷ lệ cố định phụ thuộc vào số lượng lớp đã được di chuyển do tính tiến từ vị trí tham chiếu ban đầu sang vị trí tham chiếu hoặc vị trí người nghe mới. Ví dụ, khi chuyển động chỉ diễn ra bởi một lớp, thì có thể sử dụng hệ số tỷ lệ vừa phải chẳng hạn như giữa 0,5 và 1 và khi chuyển động diễn ra từ hai lớp trở lên, thì có thể sử dụng hệ số tỷ lệ cao hơn giữa 0,1 và 0,5.

Bộ xử lý/chia tỷ lệ tín hiệu đối tượng 808 sau đó sẽ áp dụng tỷ lệ cho tín hiệu đối tượng trước khi tạo Ambisonics hoặc sẽ áp dụng tỷ lệ cho tất cả các thành phần

của phép biểu diễn Ambisonics để cuối cùng thu được tín hiệu đối tượng được xử lý theo phép biểu diễn đơn trên mỗi nguồn hoặc theo phép biểu diễn Ambisonics.

Fig.9 minh họa phương án khác nữa, tương tự như Fig.6, nhưng trong đó các khối riêng lẻ 422, 422a, 422b theo, ví dụ, Fig.4d được triển khai dưới dạng triển khai bộ mã hóa/giải mã DirAC. Sau đó, kết quả có thể là các tín hiệu kênh loa phóng thanh riêng lẻ hoặc các tín hiệu Ambisonics hoặc bất kỳ phép biểu diễn tín hiệu nào khác, ví dụ, cho bộ chuyển mã. Tuy nhiên, khi đầu ra là các tín hiệu kênh hoặc thậm chí là các tín hiệu âm lập thể, thì các tín hiệu tương ứng đó có thể được thêm vào với nhau trong bộ kết hợp 600 để biểu diễn và mô tả trường âm thanh đơn có thể được kết xuất trong khối 427 bằng bất kỳ quy trình nào khác.

Fig.10 minh họa triển khai được ưu tiên hơn nữa của bộ tính trường âm thanh 420. Quy trình được minh họa trên Fig.10 được thực hiện cho từng lớp riêng biệt. Sự khác biệt duy nhất là đối với mỗi lớp, thông tin lớp 350 khác nhau được sử dụng, và thông tin lớp này là giống nhau cho từng nguồn trong lớp. Thông tin lớp được đọc từ siêu dữ liệu bằng bộ xác định khoảng cách 1120 tạo ra giá trị khoảng cách. Nếu siêu dữ liệu đã bao gồm khoảng cách tính bằng mét hoặc tương tự, thì khối 1120 chỉ cần trích xuất thẳng dữ liệu từ dòng dữ liệu hoặc bất cứ điều gì chuyển tiếp thông tin này đến khối 1140. Do đó, thông tin khoảng cách giống nhau cho mỗi DoA của từng nguồn trong cùng một lớp được tạo và sử dụng bởi khối 1140.

Dựa trên mô tả trường âm thanh, hướng đến dải đầy đủ hoặc hướng đến của mỗi dải được xác định trong 1100. Những thông tin hướng đến này biểu diễn dữ liệu hướng đến của trường âm thanh. Dựa trên dữ liệu hướng đến này, phép biến đổi tịnh tiến được thực hiện trong khối 1110. Để đạt được mục đích này, khối 1120 lấy siêu dữ liệu cho mô tả trường âm thanh cho lớp. Dựa trên dữ liệu, khối 1110 tạo dữ liệu hướng đến mới cho trường âm thanh mà, trong phép triển khai này, chỉ phụ thuộc vào sự tịnh tiến từ vị trí tham chiếu đến vị trí tham chiếu khác. Để đạt được mục đích này, khối 1110 nhận thông tin tịnh tiến được tạo, ví dụ, bằng cách theo dõi trong bối cảnh triển khai thực tế ảo.

Tốt hơn là hoặc thay vào đó, dữ liệu xoay cũng được sử dụng. Để đạt được mục đích này, khối 1130 thực hiện biến đổi xoay bằng cách sử dụng thông tin xoay. Khi cả

phép tịnh tiến và xoay được thực hiện, thì ưu tiên thực hiện biến đổi xoay sau khi tính toán các DoA mới của trường âm thanh mà đã bao gồm thông tin từ phép tịnh tiến và khoảng cách lớp từ khối 1120.

Sau đó, trong khối 1140, mô tả trường âm thanh mới được tạo. Để đạt được mục đích này, mô tả trường âm thanh ban đầu có thể được sử dụng hoặc, thay vào đó, các tín hiệu nguồn mà được tách ra từ mô tả trường âm thanh bởi thuật toán tách nguồn có thể được sử dụng hoặc bất kỳ ứng dụng nào khác có thể được sử dụng. Về cơ bản, mô tả trường âm thanh mới có thể là, ví dụ, mô tả trường âm thanh định hướng như thu được bởi bộ tạo Ambisonics 430 hoặc như được tạo bởi bộ tổng hợp DirAC 425 hoặc có thể là phép biểu diễn âm lập thể được tạo từ phép biểu diễn loa ảo trong kết xuất âm lập thể tiếp theo.

Tốt hơn là, như được minh họa trên Fig.10, khoảng cách trên mỗi hướng đến cũng được sử dụng để tạo mô tả trường âm thanh mới để thích ứng âm lượng hoặc cường độ âm của nguồn âm thanh nhất định với vị trí mới, nghĩa là, vị trí tham chiếu mới hoặc khác.

Mặc dù Fig.10 minh họa tình huống trong đó phép biến đổi xoay được thực hiện sau phép biến đổi tịnh tiến, cần lưu ý rằng thứ tự có thể khác nhau. Đặc biệt, phép biến đổi xoay có thể được áp dụng cho các DoA của trường âm thanh như được tạo bởi khối 1100 và, sau đó, phép biến đổi tịnh tiến bổ sung được áp dụng do phép tịnh tiến của đối tượng từ vị trí tham chiếu đến vị trí tham chiếu khác.

Ngay sau khi các DoA của trường âm thanh được xác định bởi khối 1100, thông tin khoảng cách được lấy từ siêu dữ liệu bằng cách sử dụng khối 1120 và thông tin khoảng cách này sau đó được sử dụng bằng cách tạo mô tả trường âm thanh mới trong khối 1140 để tính toán khoảng cách bị thay đổi và, do đó, cường độ âm bị thay đổi của nguồn nhất định đối với vị trí tham chiếu nhất định. Về cơ bản, có thể nói rằng trong trường hợp khoảng cách trở nên lớn hơn, thì tín hiệu nguồn âm thanh cụ thể có thể được giảm dần trong khi, khi khoảng cách trở nên ngắn hơn, thì tín hiệu nguồn âm thanh được khuếch đại. Đương nhiên, sự giảm dần hoặc khuếch đại của nguồn âm thanh nhất định phụ thuộc vào khoảng cách được thực hiện theo tỷ lệ thay đổi khoảng cách, nhưng, trong các phương án khác, các hoạt động ít phức tạp hơn có thể được áp

dụng cho sự khuếch đại hoặc giảm dần này của các tín hiệu nguồn âm thanh theo mức tăng khá thô. Ngay cả việc triển khai ít phức tạp hơn như vậy mang lại các kết quả vượt trội so với tình huống mà bất kỳ thay đổi khoảng cách nào đều bị bỏ qua hoàn toàn.

Fig.7 minh họa triển khai được ưu tiên của bộ tính trường âm thanh 420. Trong khối 1102, việc phân tách nguồn và tính toán thông tin hướng đến hoặc hướng nói chung cho mỗi nguồn được thực hiện. Sau đó, trong khối 1104, vector hướng đến được nhân với vector thông tin khoảng cách, nghĩa là, ví dụ, vector từ vị trí tham chiếu ban đầu đến nguồn âm thanh, nghĩa là, vector từ mục 520 đến mục 510 trên Fig.5. Sau đó, trong khối 1106, thông tin tịnh tiến, nghĩa là, vector từ mục 520 đến mục 500 trên Fig.5 được tính đến để tính toán vector hướng tịnh tiến mới mà là vector từ vị trí người nghe 500 đến vị trí nguồn âm thanh 510. Sau đó, vector hướng đến mới với độ dài chính xác được biểu thị bởi d_v được tính toán trong khối 1108. Vector này được định hướng theo cùng hướng với d_r , nhưng có độ dài khác nhau, vì độ dài của vector này phản ánh thực tế rằng nguồn âm thanh 510 được ghi trong trường âm thanh ban đầu với âm lượng nhất định và do đó, độ dài d_v biểu thị sự thay đổi cường độ âm nhiều hơn hoặc ít hơn. Điều này có được bằng cách chia vector d_l cho khoảng cách ghi d_r , tức là, độ dài của vector d_r từ micro 520 đến nguồn âm thanh 510.

Như trên Fig.5, khi khoảng cách được tái tạo lớn hơn khoảng cách được ghi, thì độ dài d_v sẽ thấp hơn phần tử đơn vị. Điều này sẽ dẫn đến sự giảm dần của nguồn âm thanh 510 cho việc tái tạo tại vị trí người nghe mới. Tuy nhiên, khi khoảng cách được tái tạo d_l nhỏ hơn khoảng cách được ghi, thì độ dài d_v như được tính toán bởi khối 1108 sẽ lớn hơn 1 và hệ số tỷ lệ tương ứng sẽ dẫn đến khuếch đại nguồn âm thanh.

Fig.11 minh họa triển khai được ưu tiên hơn nữa của bộ tính trường âm thanh.

Trong khối 1200, các nguồn âm thanh riêng lẻ từ trường âm thanh được xác định, ví dụ, cho mỗi dải hoặc toàn dải. Khi phép xác định cho mỗi khung và dải được thực hiện, thì điều này có thể được hoàn thành bằng phép phân tích DirAC. Nếu phép xác định toàn dải hoặc băng con được thực hiện, thì điều này có thể được hoàn thành bằng bất kỳ loại thuật toán tách nguồn toàn dải hoặc băng con nào.

Trong khối 1210, phép tịnh tiến và/hoặc xoay của người nghe được xác định, ví dụ, bằng cách theo dõi đầu.

Trong khối 1220, khoảng cách cũ cho mỗi nguồn được xác định bằng cách sử dụng siêu dữ liệu và, ví dụ, bằng cách sử dụng siêu dữ liệu cho lớp chẳng hạn như khoảng cách biểu diễn. Do đó, mỗi dải được coi là nguồn nhất định (với điều kiện là độ khuếch tán thấp hơn một ngưỡng nhất định), và sau đó, khoảng cách nhất định cho mỗi ngăn thời gian/tần số có giá trị độ khuếch tán thấp được xác định.

Sau đó, trong khối 1230, khoảng cách mới cho mỗi nguồn được thu, ví dụ, bởi phép tính vector cho mỗi dải mà, ví dụ, được thảo luận trong bối cảnh theo Fig.7 hoặc bằng cách sử dụng quy trình theo Fig.8 dựa trên việc phát hiện sự thay đổi của lớp.

Hơn nữa, như được minh họa trong khối 1240, hướng cũ cho mỗi nguồn được xác định, ví dụ, bởi phép tính DoA thu được trong phép phân tích DirAC hoặc bởi hướng đến hoặc phép phân tích thông tin hướng trong thuật toán tách nguồn, ví dụ.

Sau đó, trong khối 1250, hướng mới cho mỗi nguồn được xác định, ví dụ, bằng cách thực hiện phép tính vector cho mỗi dải hoặc toàn dải.

Sau đó, trong khối 1260, trường âm thanh mới được tạo cho người nghe đã tịnh tiến và xoay. Điều này có thể được thực hiện, ví dụ, bằng cách chia tỷ lệ phần hướng cho mỗi kênh trong phép tổng hợp DirAC. Tùy thuộc vào sự triển khai cụ thể, phép biến đổi khoảng cách có thể được thực hiện trong các khối 1270a, 1270b hoặc 1270c ngoài ra hoặc thay thế để thực hiện phép biến đổi khoảng cách trong khối 1260.

Ví dụ, khi đã xác định rằng trường âm thanh chỉ có duy nhất một nguồn âm thanh, thì phép biến đổi khoảng cách có thể sẵn sàng được thực hiện trong khối 1270a.

Thay vào đó, khi các tín hiệu nguồn riêng lẻ được tính toán bởi khối 1200, thì phép biến đổi khoảng cách có thể được thực hiện cho các nguồn riêng lẻ trong khối 1270b, trước khi trường âm thanh mới thực sự được tạo trong khối 1260.

Ngoài ra, khi việc tạo trường âm thanh trong khối 1260, ví dụ, không kết xuất tín hiệu thiết lập loa phóng thanh hoặc tín hiệu âm lập thể, mà mô tả trường âm thanh khác, ví dụ, bằng cách sử dụng bộ mã hóa hoặc bộ tính toán Ambisonics 430, thì phép biến đổi khoảng cách cũng có thể được thực hiện sau khi tạo trong khối 1260, mà có

nghĩa là trong khối 1270c. Tùy thuộc vào việc triển khai, việc biến đổi khoảng cách cũng có thể được phân bố cho một số bộ biến đổi để, đến cuối cùng, nguồn âm thanh nhất định có cường độ âm nhất định được định hướng bởi chênh lệch giữa khoảng cách ban đầu giữa nguồn âm thanh và vị trí tham chiếu và khoảng cách mới giữa nguồn âm thanh và vị trí tham chiếu khác.

Fig.12a minh họa bộ phân tích DirAC như được bộc lộ ban đầu, ví dụ, trong tài liệu tham khảo được trích dẫn trước đó "Directional Audio Coding" từ IWPASH năm 2009.

Bộ phân tích DirAC bao gồm ngân hàng các bộ lọc băng tần 1310, bộ phân tích năng lượng 1320, bộ phân tích cường độ 1330, khối trung bình theo thời gian 1340 và bộ tính độ khuếch tán 1350 và bộ tính hướng 1360.

Trong DirAC, cả phép phân tích và tổng hợp đều được thực hiện trong miền tần số. Có một số phương pháp để phân chia âm thanh thành các dải tần số, trong từng thuộc tính riêng biệt. Các biến đổi tần số được sử dụng phổ biến nhất bao gồm biến đổi biến đổi Fourier thời gian ngắn (short time Fourier transform - STFT), và giàn lọc gương cầu phương (quadrature mirror filter - QMF). Ngoài ra, có toàn quyền tự do để thiết kế giàn lọc với các bộ lọc tùy ý được tối ưu hóa cho bất kỳ mục đích cụ thể nào. Bất kể biến đổi thời gian-tần số đã chọn, mục tiêu thiết kế là bắt chước độ phân giải của thính giác không gian của con người. Mục tiêu của phân tích hướng là ước tính ở mỗi dải tần số hướng đến của âm thanh, cùng với ước tính nếu âm thanh đến từ một hoặc nhiều hướng cùng một lúc. Về nguyên tắc, điều này có thể được thực hiện với một số kỹ thuật, tuy nhiên, phân tích năng lượng của trường âm thanh đã được tìm thấy là phù hợp, được minh họa trên Fig.12a. Phân tích năng lượng có thể được thực hiện, khi tín hiệu áp lực và các tín hiệu vận tốc ở một, hai hoặc ba chiều được thu từ một vị trí duy nhất. Trong các tín hiệu định dạng B cấp một, tín hiệu đẳng hướng được gọi là tín hiệu W, mà được thu nhỏ bởi căn bậc hai của hai. Áp lực âm thanh có thể được ước tính là $P = \sqrt{2 * W}$ được biểu diễn trong miền STFT.

Các kênh X, Y và Z có mẫu định hướng của lưỡng cực được định hướng dọc theo trục Đề-các, tạo thành vector $U = [X, Y, Z]$. Vector ước tính vector vận tốc trường âm thanh, và cũng được biểu diễn trong miền STFT. Năng lượng E của trường âm

thanh được tính toán. Việc thu các tín hiệu định dạng B có thể thu được hoặc bằng cách định vị ngẫu nhiên các micrô có hướng, hoặc với tập hợp các micrô đẳng hướng có khoảng cách gần nhau. Trong một số ứng dụng, các tín hiệu micrô có thể được tạo thành trong miền tính toán, nghĩa là, được mô hình hóa.

Hướng của âm thanh được xác định là hướng ngược lại của vector cường độ I . Hướng được chỉ rõ là góc phương vị và các giá trị độ cao tương ứng trong siêu dữ liệu được truyền. Độ khuếch tán của trường âm thanh cũng được tính toán bằng cách sử dụng toán tử kỳ vọng của vector cường độ và năng lượng. Kết quả của phương trình này là số có giá trị thực nằm giữa 0 và 1, đặc trưng nếu năng lượng âm thanh đến từ một hướng duy nhất (độ khuếch tán bằng 0), hoặc từ tất cả các hướng (độ khuếch tán bằng 1). Quy trình này là phù hợp trong trường hợp có sẵn thông tin vận tốc 3 chiều (3D) đầy đủ hoặc ít hơn.

Fig.12b minh họa phép tổng hợp DirAC, một lần nữa có giàn lọc dải 1370, khối micrô ảo 1400, khối tổng hợp hướng/khuếch tán 1450, và thiết lập loa phóng thanh nhất định hoặc thiết lập loa phóng thanh dự định ảo 1460. Ngoài ra, bộ biến đổi khuếch đại khuếch tán 1380, khối bảng khuếch đại quét biên độ dựa trên vector (vector based amplitude panning - VBAP) 1390, khối bù micrô 1420, khối trung bình khuếch đại loa phóng thanh 1430 và bộ phân phối 1440 cho các kênh khác được sử dụng.

Trong phép tổng hợp DirAC với các loa phóng thanh này, phiên bản chất lượng cao của phép tổng hợp DirAC được thể hiện trên Fig.12b nhận tất cả các tín hiệu định dạng B, trong đó tín hiệu micrô ảo được tính toán cho mỗi hướng loa phóng thanh của thiết lập loa phóng thanh 1460. Mẫu định hướng được sử dụng thường là lưỡng cực. Các tín hiệu micrô ảo sau đó được biến đổi theo kiểu phi tuyến tính, tùy thuộc vào siêu dữ liệu. Phiên bản tốc độ bit thấp của DirAC không được thể hiện trên Fig.12b, tuy nhiên, trong tình huống này, chỉ một kênh của âm thanh được truyền như được minh họa trên Fig.6. Sự khác biệt trong bước xử lý là tất cả các tín hiệu micrô ảo sẽ được thay thế bằng một kênh âm thanh duy nhất nhận được. Các tín hiệu micrô ảo được chia thành hai dòng: dòng khuếch tán và dòng không khuếch tán, mà được xử lý riêng.

Âm thanh không khuếch tán được tái tạo dưới dạng các nguồn điểm bằng cách sử dụng quét biên độ dựa trên vector (VBAP). Khi quét, tín hiệu âm thanh đơn âm

được áp dụng cho tập con của các loa phóng thanh sau khi nhân với các hệ số khuếch đại dành riêng cho loa phóng thanh. Các hệ số khuếch đại được tính toán bằng cách sử dụng thông tin thiết lập loa phóng thanh, và hướng quét được chỉ định. Trong phiên bản tốc độ bit thấp, tín hiệu đầu vào được quét đơn giản đến các hướng được bao hàm bởi siêu dữ liệu. Trong phiên bản chất lượng cao, mỗi tín hiệu micrô ảo được nhân với hệ số khuếch đại tương ứng, tạo ra hiệu ứng tương tự với quét, tuy nhiên nó ít bị ảnh hưởng bởi bất kỳ thành phần lạ phi tuyến nào.

Trong nhiều trường hợp, siêu dữ liệu định hướng phải tuân theo các thay đổi đột ngột theo thời gian. Để tránh các thành phần lạ, các hệ số khuếch đại cho các loa phóng thanh được tính toán với VBAP được làm nhẵn bằng cách tích hợp tạm thời với các hằng số thời gian phụ thuộc vào tần số tương đương với khoảng 50 chu kỳ ở mỗi băng tần. Điều này có hiệu quả loại bỏ các thành phần lạ, tuy nhiên, các thay đổi theo hướng không được cảm nhận là chậm hơn so với không lấy trung bình trong hầu hết các trường hợp.

Mục đích của phép tổng hợp âm thanh khuếch tán là để tạo cảm nhận âm thanh bao quanh người nghe. Trong phiên bản tốc độ bit thấp, dòng khuếch tán được tái tạo bằng cách giải tương quan tín hiệu đầu vào và tái tạo nó từ mỗi loa phóng thanh. Trong phiên bản chất lượng cao, các tín hiệu micrô ảo của dòng khuếch tán đã không nhất quán ở một mức độ nào đó, và chúng cần được giải tương quan nhẹ. Việc tiếp cận này đưa ra chất lượng không gian tốt hơn cho âm vang xung quanh và âm thanh không gian xung quanh so với phiên bản tốc độ bit thấp.

Đối với phép tổng hợp DirAC với các ống nghe, DirAC được tạo ra với số lượng loa phóng thanh ảo nhất định xung quanh người nghe đối với dòng không khuếch tán và số lượng loa phóng thanh nhất định đối với dòng khuếch tán. Các loa phóng thanh ảo được triển khai dưới dạng tích chập của các tín hiệu đầu vào với các hàm chuyển đổi liên quan đến đầu đo được (head-related transfer function - HRTF).

Mặc dù một vài khía cạnh đã được mô tả trong ngữ cảnh của thiết bị, rõ ràng rằng các khía cạnh này cũng thể hiện mô tả của phương pháp tương ứng, trong đó khối hoặc thiết bị tương ứng với bước xử lý của phương pháp hoặc đặc điểm của bước xử lý của phương pháp. Tương tự, các khía cạnh được mô tả trong ngữ cảnh của bước xử lý

của phương pháp cũng biểu diễn sự mô tả của khối hoặc mục hoặc dấu hiệu tương ứng của thiết bị tương ứng.

Mô tả trường âm thanh nâng cao theo sáng chế có thể được lưu trữ trong vật ghi lưu trữ số hoặc vật ghi lưu trữ không tạm thời hoặc có thể được truyền trên phương tiện truyền dẫn chẳng hạn như phương tiện truyền không dây hoặc phương tiện truyền có dây chẳng hạn như Internet.

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

Một số phương án theo sáng chế bao gồm vật mang dữ liệu không tạm thời 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 để 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 mang có thể 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 khác của các phương pháp theo sáng chế là vật mang dữ liệu (hoặc vật ghi lưu trữ số, hoặc vật ghi có thể đọc được bằng máy tính) bao gồm chương

trình máy tính được ghi lại trên đó để thực hiện một trong số các phương pháp được mô tả ở đây.

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

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

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

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

YÊU CẦU BẢO HỘ

1. Thiết bị tạo mô tả trường âm thanh được biến đổi từ mô tả trường âm thanh và siêu dữ liệu liên quan đến thông tin không gian của mô tả trường âm thanh, thiết bị bao gồm:

giao diện đầu vào để nhận mô tả trường âm thanh, siêu dữ liệu, và thông tin tịnh tiến; và

bộ tính trường âm thanh để tính toán trường âm thanh được biến đổi sử dụng thông tin không gian, mô tả trường âm thanh và thông tin tịnh tiến, trong đó thông tin tịnh tiến biểu thị sự tịnh tiến từ vị trí tham chiếu tới vị trí tham chiếu khác, trong đó bộ tính trường âm thanh được tạo cấu hình:

để nhận mô tả trường âm thanh lớp thứ nhất và mô tả trường âm thanh lớp thứ hai làm mô tả trường âm thanh,

để thực hiện việc phân tách nguồn trên các mô tả trường âm thanh lớp thứ nhất và thứ hai, để trích xuất các nguồn của các mô tả trường âm thanh lớp thứ nhất và thứ hai và dữ liệu hướng đến (direction of arrival - DoA) cho các nguồn được trích xuất,

để tính toán dữ liệu DoA đã biến đổi liên quan đến vị trí khác nhau cho mỗi nguồn được trích xuất bằng cách sử dụng dữ liệu DoA và thông tin tịnh tiến, và để xử lý các nguồn được trích xuất và dữ liệu DoA đã biến đổi để thu được mô tả trường âm thanh được biến đổi.

2. Thiết bị theo điểm 1,

trong đó bộ tính trường âm thanh được tạo cấu hình

để thực hiện riêng việc phân tách nguồn cho mỗi mô tả trường âm thanh lớp và để trích xuất tín hiệu xung quanh/khuếch tán/dư cho mỗi lớp, hoặc

để thực hiện việc phân tách nguồn cho tất cả các lớp với nhau và để trích xuất tín hiệu xung quanh/khuếch tán/dư đơn cho ít nhất hai lớp.

3. Thiết bị theo điểm 1,

trong đó bộ tính trường âm thanh được tạo cấu hình để tính toán, cho mỗi nguồn trong số các nguồn, khoảng cách đã biến đổi và để tính toán mô tả trường âm thanh lớp đã biến đổi bằng cách sử dụng các khoảng cách đã biến đổi của các nguồn được trích xuất.

4. Thiết bị theo điểm 1,

trong đó bộ tính trường âm thanh được tạo cấu hình để thêm các mô tả trường âm thanh lớp cho mỗi thành phần tương ứng của mỗi mô tả trường âm thanh lớp để thu được mô tả trường âm thanh tổng thể.

5. Thiết bị theo điểm 1,

trong đó bộ tính trường âm thanh được tạo cấu hình để áp dụng phép xoay cho mô tả trường âm thanh được biến đổi, mô tả trường âm thanh được biến đổi là mô tả trường âm thanh lớp hoặc mô tả trường âm thanh tổng thể liên quan đến vị trí tham chiếu khác nhau.

6. Thiết bị theo điểm 1,

trong đó bộ tính trường âm thanh được tạo cấu hình, cho mỗi lớp,

để tính toán dữ liệu hướng được biến đổi cho các ngăn thời gian-tần số khác nhau cho mô tả DirAC làm mô tả trường âm thanh bằng cách sử dụng dữ liệu hướng, thông tin độ sâu và thông tin tịnh tiến, và

để kết xuất mô tả DirAC bằng cách sử dụng dữ liệu hướng được biến đổi thành mô tả âm thanh bao gồm nhiều kênh âm thanh hoặc để truyền hoặc lưu trữ mô tả DirAC bằng cách sử dụng dữ liệu hướng được biến đổi thay vì dữ liệu hướng cho các ngăn thời gian-tần số, và dữ liệu độ khuếch tán giống nhau tùy ý như được bao gồm trong mô tả DirAC.

7. Thiết bị theo điểm 1, trong đó bộ tính trường âm thanh được tạo cấu hình để xác định, đối với mỗi lớp, cho ngăn thời gian-tần số, để duy trì dữ liệu hướng hoặc để tính toán dữ liệu hướng được biến đổi dựa trên dữ liệu độ khuếch tán cho ngăn thời gian-tần số, trong đó dữ liệu hướng được biến đổi chỉ được tính cho dữ liệu độ khuếch tán cho thấy độ khuếch tán thấp hơn mức độ khuếch tán được xác định trước hoặc thích ứng.

8. Thiết bị theo điểm 1, còn bao gồm:

giao diện tịnh tiến để cung cấp thông tin tịnh tiến hoặc thông tin xoay cho biết sự xoay của người nghe dự định cho trường âm thanh được biến đổi;

bộ cung cấp siêu dữ liệu để cung cấp siêu dữ liệu cho bộ tính trường âm thanh;

bộ cung cấp trường âm thanh để cung cấp mô tả trường âm thanh cho bộ tính trường âm thanh; và

giao diện đầu ra để xuất ra trường âm thanh được biến đổi bao gồm mô tả trường âm thanh được biến đổi và siêu dữ liệu được biến đổi, siêu dữ liệu được biến đổi được suy ra từ siêu dữ liệu bằng cách sử dụng thông tin tịnh tiến, hoặc để xuất ra nhiều kênh loa phóng thanh, mỗi kênh loa phóng thanh có liên quan đến vị trí loa phóng thanh được xác định trước, hoặc để xuất ra phép biểu diễn âm lập thể của trường âm thanh được biến đổi.

9. Thiết bị theo điểm 1,

trong đó vector thứ nhất được tính bằng cách nhân vector đơn vị hướng đến với khoảng cách được bao gồm trong siêu dữ liệu, hoặc

trong đó vector thứ hai được tính bằng cách trừ vector tịnh tiến từ vector thứ nhất, hoặc trong đó giá trị biến đổi khoảng cách được tính bằng cách chia vector thứ hai cho mức chuẩn của vector thứ nhất.

10. Thiết bị theo điểm 1,

trong đó bộ tính trường âm thanh được tạo cấu hình để nhận, ngoài thông tin tịnh tiến, thông tin xoay, và

trong đó bộ tính toán trường âm thanh được tạo cấu hình để thực hiện phép biến đổi xoay để xoay dữ liệu hướng đến cho trường âm thanh bằng cách sử dụng thông tin xoay, trong đó dữ liệu hướng đến được suy ra từ dữ liệu hướng đến thu được bởi phép phân tích trường âm thanh của mô tả trường âm thanh và bằng cách sử dụng thông tin tịnh tiến.

11. Thiết bị theo điểm 1, trong đó bộ tính trường âm thanh được tạo cấu hình, đối với mỗi lớp:

để xác định các nguồn từ mô tả trường âm thanh và các hướng cho các nguồn bằng phép phân tích trường âm thanh;

để xác định, đối với mỗi nguồn, khoảng cách của nguồn từ vị trí tham chiếu bằng cách sử dụng siêu dữ liệu;

để xác định hướng mới của nguồn liên quan đến vị trí tham chiếu khác nhau bằng cách sử dụng hướng cho nguồn và thông tin tịnh tiến;

để xác định thông tin khoảng cách mới cho nguồn liên quan đến vị trí tham chiếu khác nhau; và

để tạo trường âm thanh được biến đổi bằng cách sử dụng hướng mới của nguồn, thông tin khoảng cách mới, và mô tả trường âm thanh hoặc các tín hiệu nguồn tương ứng với các nguồn được suy ra từ mô tả trường âm thanh.

12. Thiết bị theo điểm 1, trong đó bộ tính trường âm thanh được tạo cấu hình, đối với mỗi lớp:

để xác định các tín hiệu nguồn từ mô tả trường âm thanh và các hướng của các tín hiệu nguồn liên quan đến vị trí tham chiếu bằng phép phân tích âm thanh;

để tính toán các hướng mới của các tín hiệu nguồn liên quan đến vị trí tham chiếu khác nhau bằng cách sử dụng thông tin tịnh tiến;

để tính toán thông tin khoảng cách cho các nguồn âm thanh liên quan đến vị trí tham chiếu khác nhau; và

để tổng hợp trường âm thanh được biến đổi bằng cách sử dụng thông tin khoảng cách, các tín hiệu nguồn và các hướng mới.

13. Thiết bị theo điểm 12, trong đó bộ tính trường âm thanh được tạo cấu hình:

để tổng hợp trường âm thanh được biến đổi bằng cách quét tín hiệu nguồn âm thanh tới hướng được đưa ra bởi hướng mới liên quan đến thiết lập phát lại, và

bằng cách chia tỷ lệ tín hiệu nguồn âm thanh bằng cách sử dụng thông tin khoảng cách trước khi thực hiện việc quét hoặc sau khi thực hiện việc quét.

14. Thiết bị theo điểm 12,

trong đó bộ tính trường âm thanh được tạo cấu hình để thêm, cho mỗi lớp, hoặc chỉ cho một lớp duy nhất, tín hiệu khuếch tán đến phần có hướng của tín hiệu nguồn âm thanh, phần có hướng được biến đổi bởi thông tin khoảng cách trước khi được thêm vào tín hiệu khuếch tán.

15. Thiết bị theo điểm 12, trong đó bộ tính trường âm thanh được tạo cấu hình, đối với mỗi lớp,

để thực hiện phép chuyển đổi thời gian-tần số của mô tả trường âm thanh và để tính toán hướng đến cho nhiều ngăn tần số của khung thời gian;

để tính toán hướng mới cho mỗi ngăn tần số,

để tính toán thông tin khoảng cách cho mỗi ngăn tần số, và

để thực hiện phép tổng hợp hướng cho từng ngăn tần số bằng cách sử dụng tín hiệu âm thanh cho ngăn tần số, độ khuếch đại quét cho ngăn tần số được suy ra từ hướng mới cho ngăn tần số và giá trị tỷ lệ cho ngăn tần số được suy ra từ thông tin khoảng cách cho lớp tương ứng.

16. Thiết bị theo điểm 15, trong đó bộ tính trường âm thanh được tạo cấu hình, đối với mỗi lớp,

để thực hiện phép tổng hợp khuếch tán bằng cách sử dụng tín hiệu âm thanh khuếch tán phát ra từ tín hiệu âm thanh cho ngăn tần số và bằng cách sử dụng tham số độ khuếch tán có được từ việc phân tích âm thanh cho ngăn tần số và để kết hợp phân có hướng và phân khuếch tán để thu được tín hiệu âm thanh tổng hợp cho ngăn tần số; và

để thực hiện phép chuyển đổi tần số-thời gian bằng cách sử dụng các tín hiệu âm thanh cho các ngăn tần số cho khung thời gian để thu được tín hiệu âm thanh tổng hợp miền thời gian làm trường âm thanh được biến đổi.

17. Thiết bị theo điểm 12, trong đó bộ tính trường âm thanh được tạo cấu hình để tổng hợp, cho mỗi nguồn âm thanh, trường âm thanh liên quan đến vị trí tham chiếu khác nhau, phép tổng hợp bao gồm:

đối với mỗi nguồn, xử lý tín hiệu nguồn sử dụng hướng mới cho tín hiệu nguồn để thu được mô tả trường âm thanh của tín hiệu nguồn liên quan đến vị trí tham chiếu khác nhau;

biến đổi tín hiệu nguồn trước khi xử lý tín hiệu nguồn hoặc biến đổi mô tả trường âm thanh bằng cách sử dụng thông tin hướng; và

thêm các mô tả trường âm thanh đối với các nguồn để thu được trường âm thanh được biến đổi liên quan đến vị trí tham chiếu khác.

18. Thiết bị theo điểm 12,

trong đó bộ phân tích âm thanh được tạo cấu hình để xác định các tín hiệu nguồn bằng thuật toán phân tách nguồn và để trừ ít nhất một số tín hiệu nguồn từ mô tả trường âm thanh để thu được tín hiệu khuếch tán.

19. Thiết bị theo điểm 1,

trong đó bộ tính trường âm thanh được tạo cấu hình để nhận, đối với mỗi mô tả trường âm thanh lớp, khoảng cách biểu diễn cho mô tả trường âm thanh lớp từ siêu dữ liệu,

trong đó bộ tính trường âm thanh được tạo cấu hình để xác định giá trị tỷ lệ cho nguồn được xác định từ mô tả trường âm thanh lớp bằng cách sử dụng việc xác định, dù nguồn đó có nằm trong cùng một lớp đối với vị trí tham chiếu mới được xác định bởi thông tin tĩnh tiến hay không,

trong đó giá trị tỷ lệ được xác định là lớn hơn một, nếu nguồn ở lớp thấp hơn đối với vị trí tham chiếu khác nhau, hoặc

trong đó giá trị tỷ lệ được xác định là nhỏ hơn một, nếu nguồn ở lớp cao hơn đối với vị trí tham chiếu khác nhau.

20. Thiết bị theo điểm 1,

trong đó bộ tính trường âm thanh được tạo cấu hình để phân tích từng mô tả trường âm thanh lớp để thu được thông tin hướng đến cho từng nguồn được bao gồm trong mô tả trường âm thanh lớp,

để xác định khoảng cách của nguồn của mô tả trường âm thanh lớp đến vị trí tham chiếu mới bằng cách sử dụng thông tin hướng đến và siêu dữ liệu cho mô tả trường âm thanh lớp, và

để xác định hệ số tỷ lệ bằng cách sử dụng khoảng cách của nguồn của mô tả trường âm thanh lớp đến vị trí tham chiếu mới và siêu dữ liệu cho mô tả trường âm thanh mà từ đó nguồn được trích xuất.

21. Thiết bị theo điểm 20,

trong đó bộ tính trường âm thanh được tạo cấu hình để xác định khoảng cách từ vị trí tham chiếu mới đến vị trí của nguồn bằng cách sử dụng vị trí của nguồn và thông tin tịnh tiến, và

để so sánh khoảng cách đến vị trí tham chiếu mới với khoảng cách biểu diễn của lớp xung quanh vị trí tham chiếu khác nhau, sao cho nguồn được kết xuất thành lớp xung quanh vị trí tham chiếu khác nhau theo sự so sánh.

22. Thiết bị theo điểm 1,

trong đó bộ tính trường âm thanh được tạo cấu hình để tạo, cho nguồn, vector thứ nhất từ vị trí tham chiếu đến nguồn bằng cách sử dụng thông tin hướng đến và thông tin hình học, trong đó thông tin hình học giống nhau được sử dụng cho mỗi nguồn trong mô tả trường âm thanh lớp,

trong đó bộ tính trường âm thanh được tạo cấu hình để tạo vector thứ hai tạo thành vector thứ nhất và thông tin tịnh tiến cho từng nguồn trong mô tả trường âm thanh lớp, và

trong đó bộ tính trường âm thanh được tạo cấu hình để tính giá trị tỷ lệ cho mỗi nguồn bằng cách chia độ dài của vector thứ hai cho độ dài của vector thứ nhất.

23. Thiết bị tạo mô tả trường âm thanh được biến đổi từ mô tả trường âm thanh và siêu dữ liệu liên quan đến thông tin không gian của mô tả trường âm thanh, thiết bị bao gồm:

giao diện đầu vào để nhận mô tả trường âm thanh, siêu dữ liệu, và thông tin tịnh tiến; và

bộ tính trường âm thanh để tính toán trường âm thanh được biến đổi sử dụng thông tin không gian, mô tả trường âm thanh và thông tin tịnh tiến, trong đó thông tin tịnh tiến biểu thị sự tịnh tiến từ vị trí tham chiếu tới vị trí tham chiếu khác,

trong đó mô tả trường âm thanh được tạo cấu hình để tạo mô tả trường âm thanh cho mỗi nguồn trong số các nguồn được trích xuất bằng cách sử dụng dữ liệu DoA đã biến đổi,

chọn các mô tả trường âm thanh của các nguồn âm thanh dựa trên khoảng cách giữa nguồn được trích xuất tương ứng và vị trí tham chiếu khác nhau cho lớp cụ thể và thêm mô tả trường âm thanh của mỗi nguồn trong số các nguồn được chọn trong lớp để thu được mô tả trường âm thanh lớp được biến đổi, hoặc

thêm mô tả trường âm thanh của từng nguồn trong cảnh âm thanh để thu được mô tả trường âm thanh tổng thể đã biến đổi.

24. Thiết bị theo điểm 23,

trong đó bộ tính trường âm thanh được tạo cấu hình để thêm tín hiệu xung quanh/khuếch tán/dur cho mỗi lớp trong việc tính toán mô tả trường âm thanh lớp được biến đổi hoặc để tính toán tín hiệu xung quanh/khuếch tán/dur duy nhất được thêm vào mô tả trường âm thanh tổng thể.

25. Thiết bị tạo mô tả trường âm thanh được biến đổi từ mô tả trường âm thanh và siêu dữ liệu liên quan đến thông tin không gian của mô tả trường âm thanh, thiết bị bao gồm:

giao diện đầu vào để nhận mô tả trường âm thanh, siêu dữ liệu, và thông tin tịnh tiến; và

bộ tính trường âm thanh để tính toán trường âm thanh được biến đổi sử dụng thông tin không gian, mô tả trường âm thanh và thông tin tịnh tiến, trong đó thông tin tịnh tiến biểu thị sự tịnh tiến từ vị trí tham chiếu tới vị trí tham chiếu khác,

trong đó mô tả trường âm thanh bao gồm, đối với mỗi lớp, nhiều thành phần trường âm thanh, nhiều thành phần trường âm thanh bao gồm thành phần đa hướng và ít nhất một thành phần định hướng, và

trong đó bộ tính trường âm thanh được tạo cấu hình, cho mỗi lớp:

để phân tích các thành phần trường âm thanh để suy ra, đối với các ngăn tần số khác nhau, thông tin hướng đến;

để tính toán thông tin hướng đến được biến đổi trên mỗi ngăn tần số bằng cách sử dụng thông tin hướng và siêu dữ liệu, siêu dữ liệu bao gồm bản đồ độ sâu liên kết thông tin khoảng cách với lớp; và

để tính toán trường âm thanh được biến đổi bằng cách sử dụng thông tin bù khoảng cách tùy thuộc vào khoảng cách được cung cấp bởi bản đồ độ sâu cho lớp, và khoảng cách mới được liên kết với ngăn tần số có liên quan đến thông tin hướng đến được biến đổi.

26. Thiết bị tạo mô tả trường âm thanh được biến đổi từ mô tả trường âm thanh và siêu dữ liệu liên quan đến thông tin không gian của mô tả trường âm thanh, thiết bị bao gồm:

giao diện đầu vào để nhận mô tả trường âm thanh, siêu dữ liệu, và thông tin tịnh tiến; và

bộ tính trường âm thanh để tính toán trường âm thanh được biến đổi sử dụng thông tin không gian, mô tả trường âm thanh và thông tin tịnh tiến, trong đó thông tin tịnh tiến biểu thị sự tịnh tiến từ vị trí tham chiếu tới vị trí tham chiếu khác, trong đó bộ tính trường âm thanh được tạo cấu hình, đối với mỗi lớp,

để tính toán vector thứ nhất chỉ từ vị trí tham chiếu đến nguồn âm thanh thu được từ phép phân tích của trường âm thanh;

để tính toán vector thứ hai chỉ từ vị trí tham chiếu khác đến nguồn âm thanh bằng cách sử dụng vector thứ nhất và thông tin tịnh tiến, thông tin tịnh tiến xác định vector tịnh tiến từ vị trí tham chiếu đến vị trí tham chiếu khác; và

để tính toán giá trị biến đổi khoảng cách bằng cách sử dụng vị trí tham chiếu khác, vị trí của nguồn âm thanh, và vector thứ hai, hoặc bằng cách sử dụng khoảng cách từ vị trí tham chiếu khác đến vị trí của nguồn âm thanh và vector thứ hai.

27. Phương pháp tạo mô tả trường âm thanh được biến đổi từ mô tả trường âm thanh và siêu dữ liệu liên quan đến thông tin không gian của mô tả trường âm thanh, phương pháp bao gồm bước:

tính toán trường âm thanh được biến đổi bằng cách sử dụng thông tin không gian, mô tả trường âm thanh và thông tin tịnh tiến biểu thị sự tịnh tiến từ vị trí tham chiếu tới vị trí tham chiếu khác,

trong đó bước tính toán bao gồm:

nhận mô tả trường âm thanh lớp thứ nhất và mô tả trường âm thanh lớp thứ hai làm mô tả trường âm thanh,

thực hiện việc phân tách nguồn trên các mô tả trường âm thanh lớp thứ nhất và thứ hai, để trích xuất các nguồn của các mô tả trường âm thanh lớp thứ nhất và thứ hai và dữ liệu hướng đến (DoA) cho các nguồn được trích xuất,

tính toán dữ liệu DoA đã biến đổi liên quan đến vị trí khác nhau cho mỗi nguồn được trích xuất bằng cách sử dụng dữ liệu DoA và thông tin tịnh tiến, và

xử lý các nguồn được trích xuất và dữ liệu DoA đã biến đổi để thu được mô tả trường âm thanh được biến đổi, hoặc

trong đó bước tính toán bao gồm:

tạo mô tả trường âm thanh cho mỗi nguồn trong số các nguồn được trích xuất bằng cách sử dụng dữ liệu DoA đã biến đổi,

chọn các mô tả trường âm thanh của các nguồn âm thanh dựa trên khoảng cách giữa nguồn được trích xuất tương ứng và vị trí tham chiếu khác nhau cho lớp cụ thể và thêm mô tả trường âm thanh của mỗi nguồn trong số các nguồn được chọn trong lớp để thu được mô tả trường âm thanh lớp được biến đổi, hoặc

thêm mô tả trường âm thanh của từng nguồn trong cảnh âm thanh để thu được mô tả trường âm thanh tổng thể đã biến đổi.

28. Phương tiện lưu trữ số không tạm thời lưu trữ trên đó chương trình máy tính để thực hiện, khi chương trình máy tính này được chạy bởi máy tính, phương pháp tạo mô

tả trường âm thanh được biến đổi từ mô tả trường âm thanh và siêu dữ liệu liên quan đến thông tin không gian của mô tả trường âm thanh, phương pháp bao gồm bước:

tính toán trường âm thanh được biến đổi bằng cách sử dụng thông tin không gian, mô tả trường âm thanh và thông tin tịnh tiến biểu thị sự tịnh tiến từ vị trí tham chiếu tới vị trí tham chiếu khác,

trong đó bước tính toán bao gồm:

nhận mô tả trường âm thanh lớp thứ nhất và mô tả trường âm thanh lớp thứ hai làm mô tả trường âm thanh,

thực hiện việc phân tách nguồn trên các mô tả trường âm thanh lớp thứ nhất và thứ hai, để trích xuất các nguồn của các mô tả trường âm thanh lớp thứ nhất và thứ hai và dữ liệu hướng đến (DoA) cho các nguồn được trích xuất,

tính toán dữ liệu DoA đã biến đổi liên quan đến vị trí khác nhau cho mỗi nguồn được trích xuất bằng cách sử dụng dữ liệu DoA và thông tin tịnh tiến, và

xử lý các nguồn được trích xuất và dữ liệu DoA đã biến đổi để thu được mô tả trường âm thanh được biến đổi, hoặc

trong đó bước tính toán bao gồm:

tạo mô tả trường âm thanh cho mỗi nguồn trong số các nguồn được trích xuất bằng cách sử dụng dữ liệu DoA đã biến đổi,

chọn các mô tả trường âm thanh của các nguồn âm thanh dựa trên khoảng cách giữa nguồn được trích xuất tương ứng và vị trí tham chiếu khác nhau cho lớp cụ thể và thêm mô tả trường âm thanh của mỗi nguồn trong số các nguồn được chọn trong lớp để thu được mô tả trường âm thanh lớp được biến đổi, hoặc

thêm mô tả trường âm thanh của từng nguồn trong cảnh âm thanh để thu được mô tả trường âm thanh tổng thể đã biến đổi.

29. Phương pháp tạo mô tả trường âm thanh được biến đổi từ mô tả trường âm thanh và siêu dữ liệu liên quan đến thông tin không gian của mô tả trường âm thanh, phương pháp bao gồm bước:

tính toán trường âm thanh được biến đổi bằng cách sử dụng thông tin không gian, mô tả trường âm thanh và thông tin tịnh tiến biểu thị sự tịnh tiến từ vị trí tham chiếu tới vị trí tham chiếu khác, trong đó mô tả trường âm thanh bao gồm, đối với mỗi lớp, nhiều thành phần trường âm thanh, nhiều thành phần trường âm thanh bao gồm thành phần đẳng hướng và ít nhất một thành phần định hướng,

trong đó bước tính toán trường âm thanh bao gồm, cho mỗi lớp:

phân tích các thành phần trường âm thanh để suy ra thông tin hướng đến cho các ngăn tần số khác nhau;

tính toán thông tin hướng đến được biến đổi trên mỗi ngăn tần số bằng cách sử dụng thông tin hướng và siêu dữ liệu, siêu dữ liệu bao gồm bản đồ độ sâu liên kết thông tin khoảng cách với lớp; và

tính toán trường âm thanh được biến đổi bằng cách sử dụng thông tin bù khoảng cách tùy thuộc vào khoảng cách được cung cấp bởi bản đồ độ sâu cho lớp, và khoảng cách mới liên quan đến ngăn tần số có liên quan đến thông tin hướng đến được biến đổi, hoặc

trong đó bước tính toán bao gồm, cho mỗi lớp:

tính toán vector thứ nhất chỉ từ vị trí tham chiếu đến nguồn âm thanh thu được từ phép phân tích của trường âm thanh;

tính toán vector thứ hai chỉ từ vị trí tham chiếu khác đến nguồn âm thanh bằng cách sử dụng vector thứ nhất và thông tin tịnh tiến, thông tin tịnh tiến xác định vector tịnh tiến từ vị trí tham chiếu đến vị trí tham chiếu khác; và

tính toán giá trị biến đổi khoảng cách bằng cách sử dụng vị trí tham chiếu khác nhau, vị trí của nguồn âm thanh, và vector thứ hai, hoặc bằng cách sử dụng khoảng cách từ vị trí tham chiếu khác nhau đến vị trí của nguồn âm thanh và vector thứ hai.

30. Phương tiện lưu trữ số không tạm thời lưu trữ trên đó chương trình máy tính để thực hiện, khi chương trình máy tính này được chạy bởi máy tính, phương pháp tạo mô tả trường âm thanh được biến đổi từ mô tả trường âm thanh và siêu dữ liệu liên quan đến thông tin không gian của mô tả trường âm thanh, phương pháp bao gồm bước:

tính toán trường âm thanh được biến đổi bằng cách sử dụng thông tin không gian, mô tả trường âm thanh và thông tin tịnh tiến biểu thị sự tịnh tiến từ vị trí tham chiếu tới vị trí tham chiếu khác,

trong đó bước tính toán trường âm thanh bao gồm, cho mỗi lớp:

phân tích các thành phần trường âm thanh để suy ra thông tin hướng đến cho các ngăn tần số khác nhau;

tính toán thông tin hướng đến được biến đổi trên mỗi ngăn tần số bằng cách sử dụng thông tin hướng và siêu dữ liệu, siêu dữ liệu bao gồm bản đồ độ sâu liên kết thông tin khoảng cách với lớp; và

tính toán trường âm thanh được biến đổi bằng cách sử dụng thông tin bù khoảng cách tùy thuộc vào khoảng cách được cung cấp bởi bản đồ độ sâu cho lớp, và khoảng cách mới liên quan đến ngăn tần số có liên quan đến thông tin hướng đến được biến đổi, hoặc

trong đó bước tính toán bao gồm, cho mỗi lớp:

tính toán vector thứ nhất chỉ từ vị trí tham chiếu đến nguồn âm thanh thu được từ phép phân tích của trường âm thanh;

tính toán vector thứ hai chỉ từ vị trí tham chiếu khác đến nguồn âm thanh bằng cách sử dụng vector thứ nhất và thông tin tịnh tiến, thông tin tịnh tiến xác định vector tịnh tiến từ vị trí tham chiếu đến vị trí tham chiếu khác; và

tính toán giá trị biến đổi khoảng cách bằng cách sử dụng vị trí tham chiếu khác nhau, vị trí của nguồn âm thanh, và vector thứ hai, hoặc bằng cách sử dụng khoảng cách từ vị trí tham chiếu khác nhau đến vị trí của nguồn âm thanh và vector thứ hai.

1/19

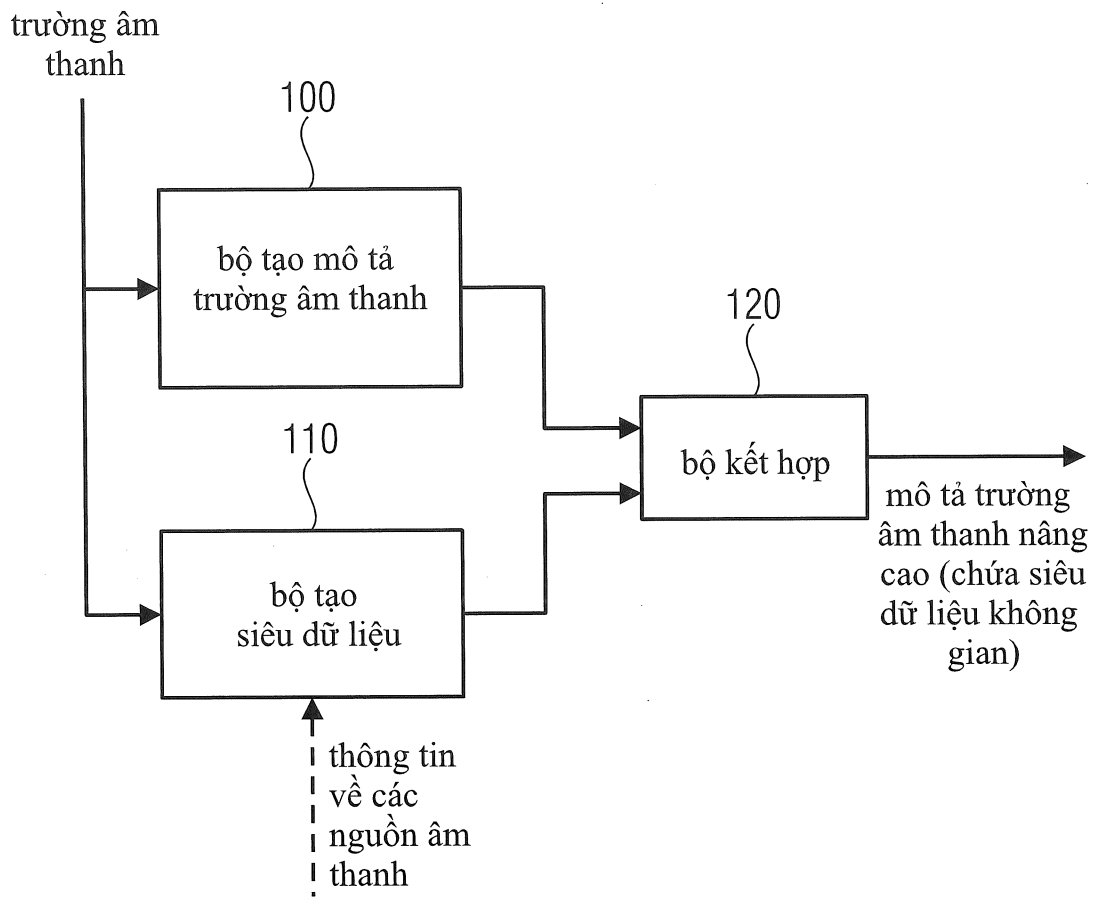


Fig. 1a

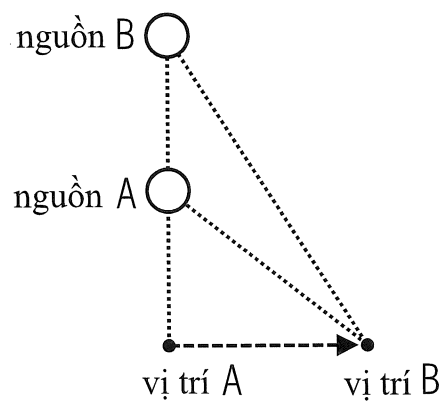


Fig. 1b

2/19

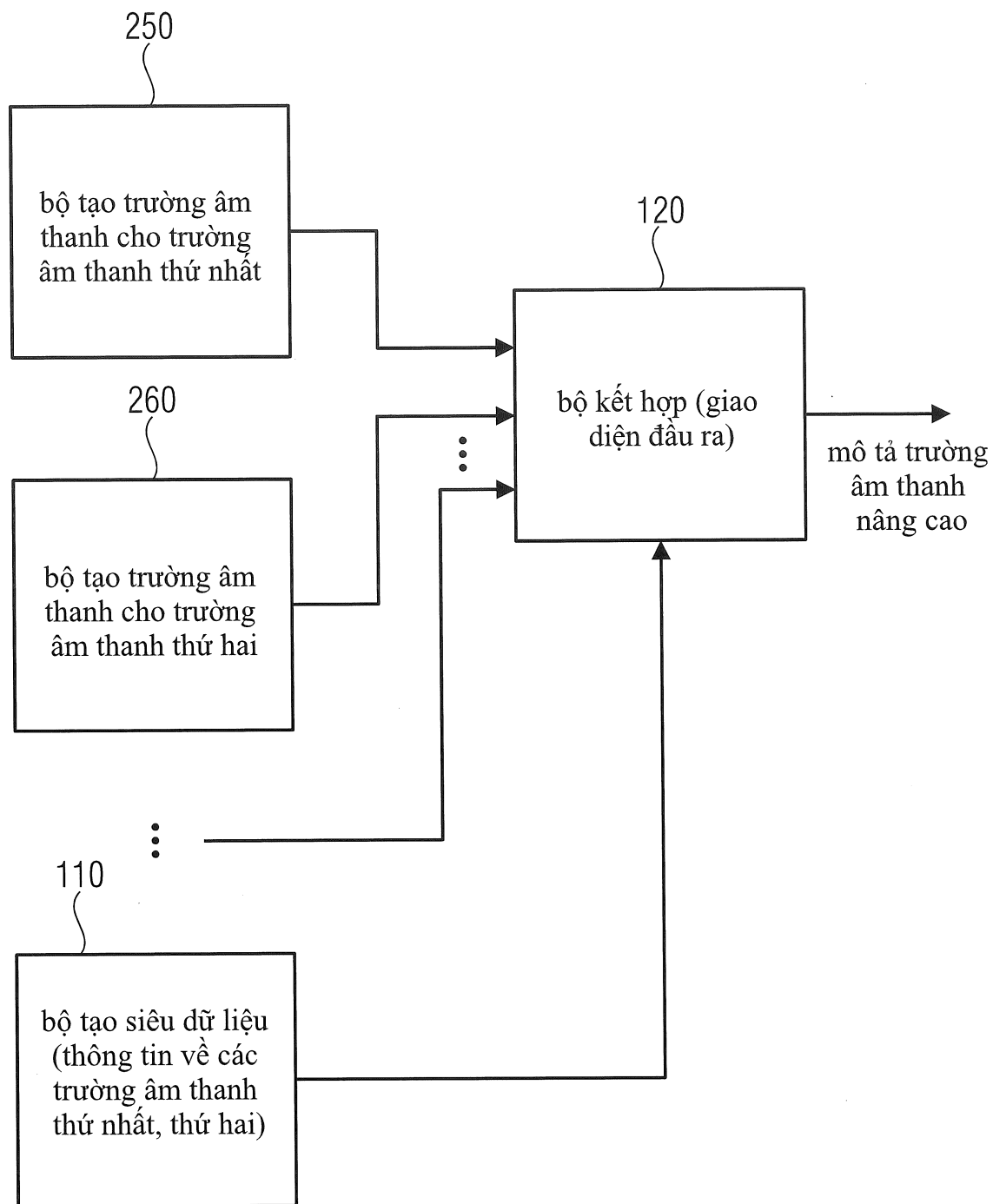


Fig. 2

3/19

	330 mô tả trường âm thanh thứ nhất	340 mô tả trường âm thanh thứ hai	350 thông tin về các trường âm thanh thứ nhất và thứ hai
--	---------------------------------------	--------------------------------------	---

Fig. 3a

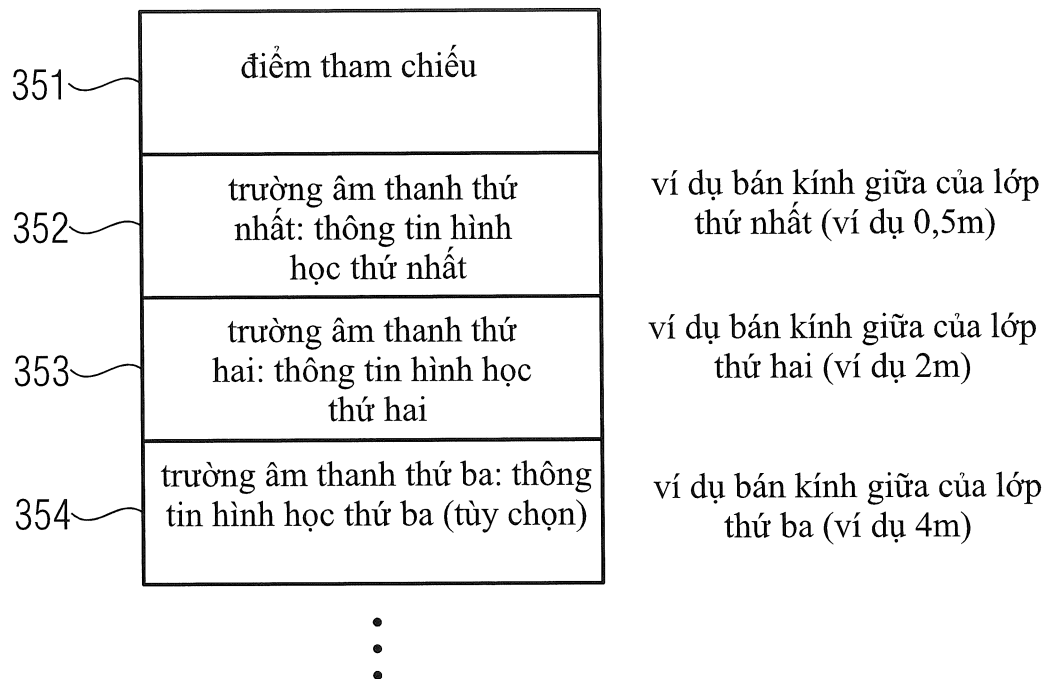


Fig. 3b

4/19

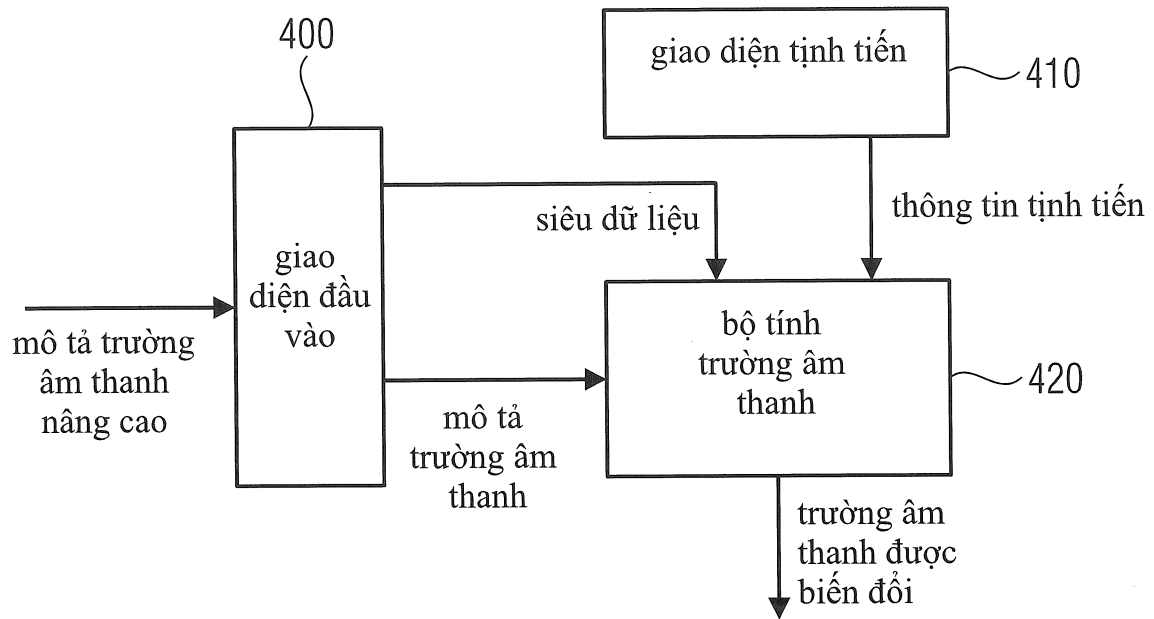


Fig. 4a

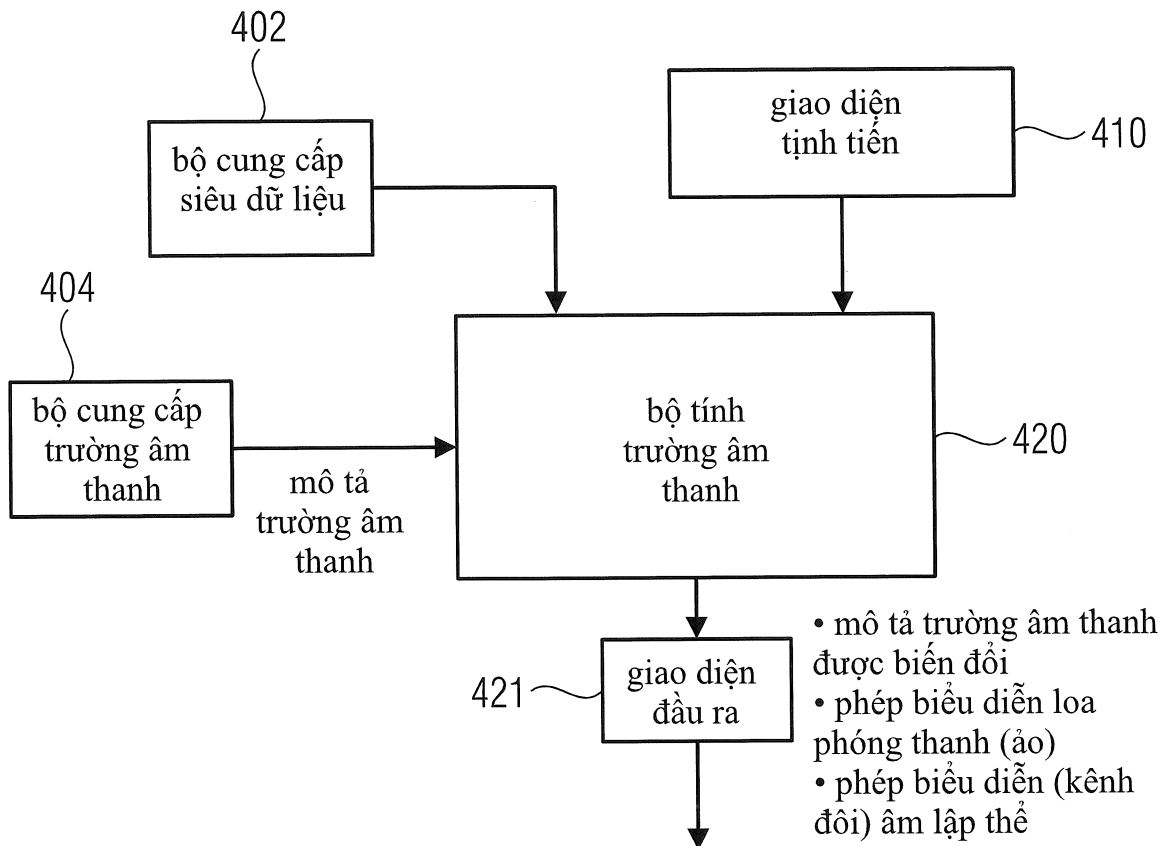


Fig. 4b

6/19

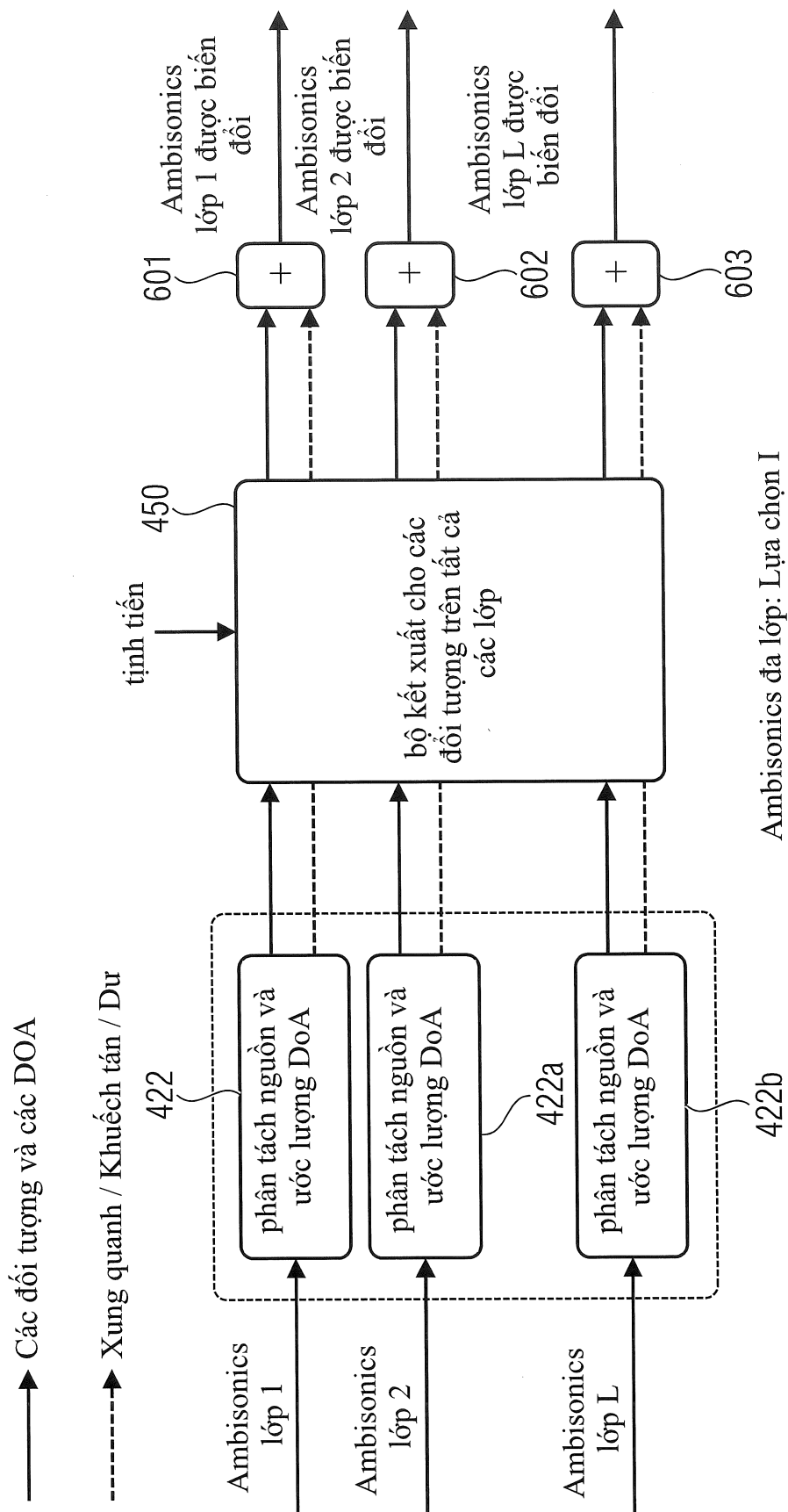


Fig. 4d

7/19

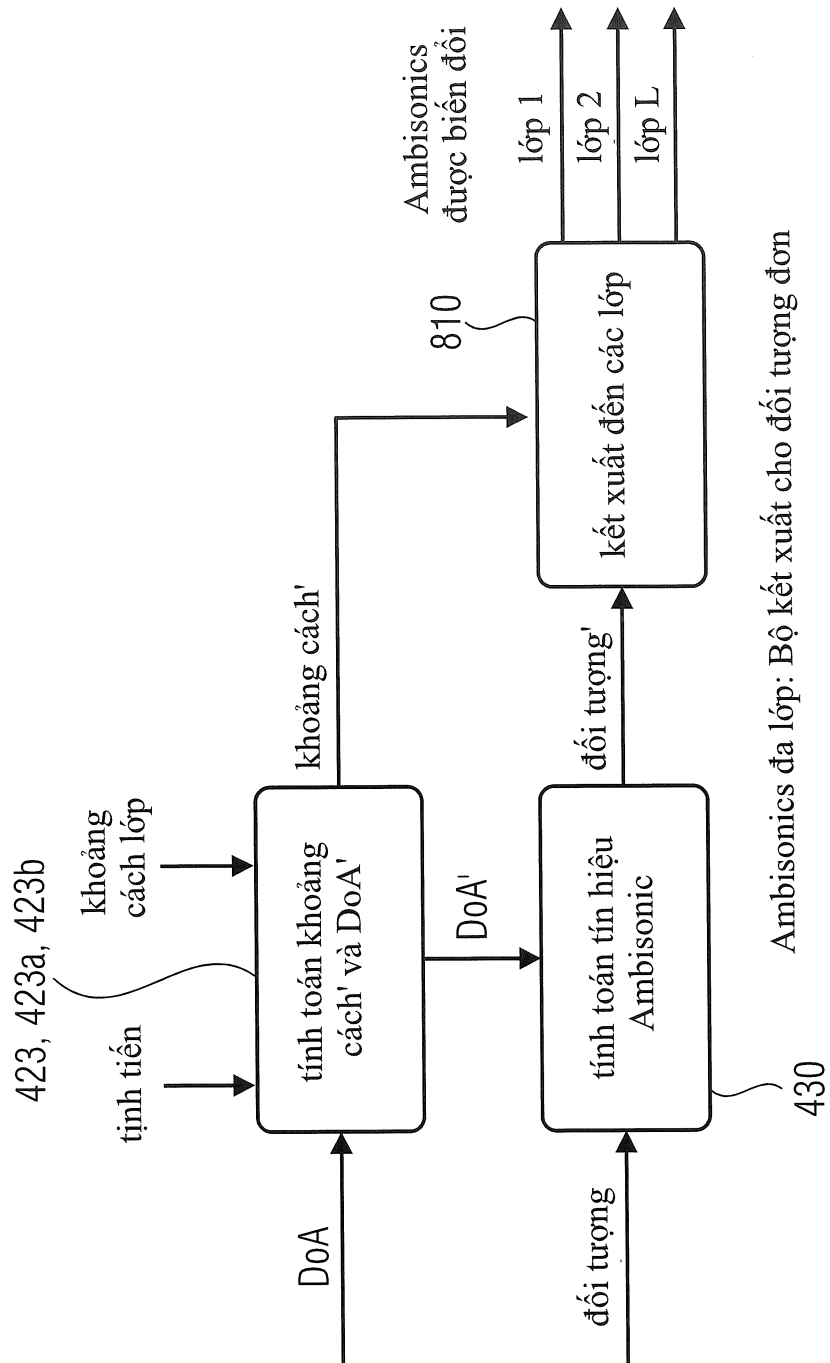


Fig. 4e

8/19

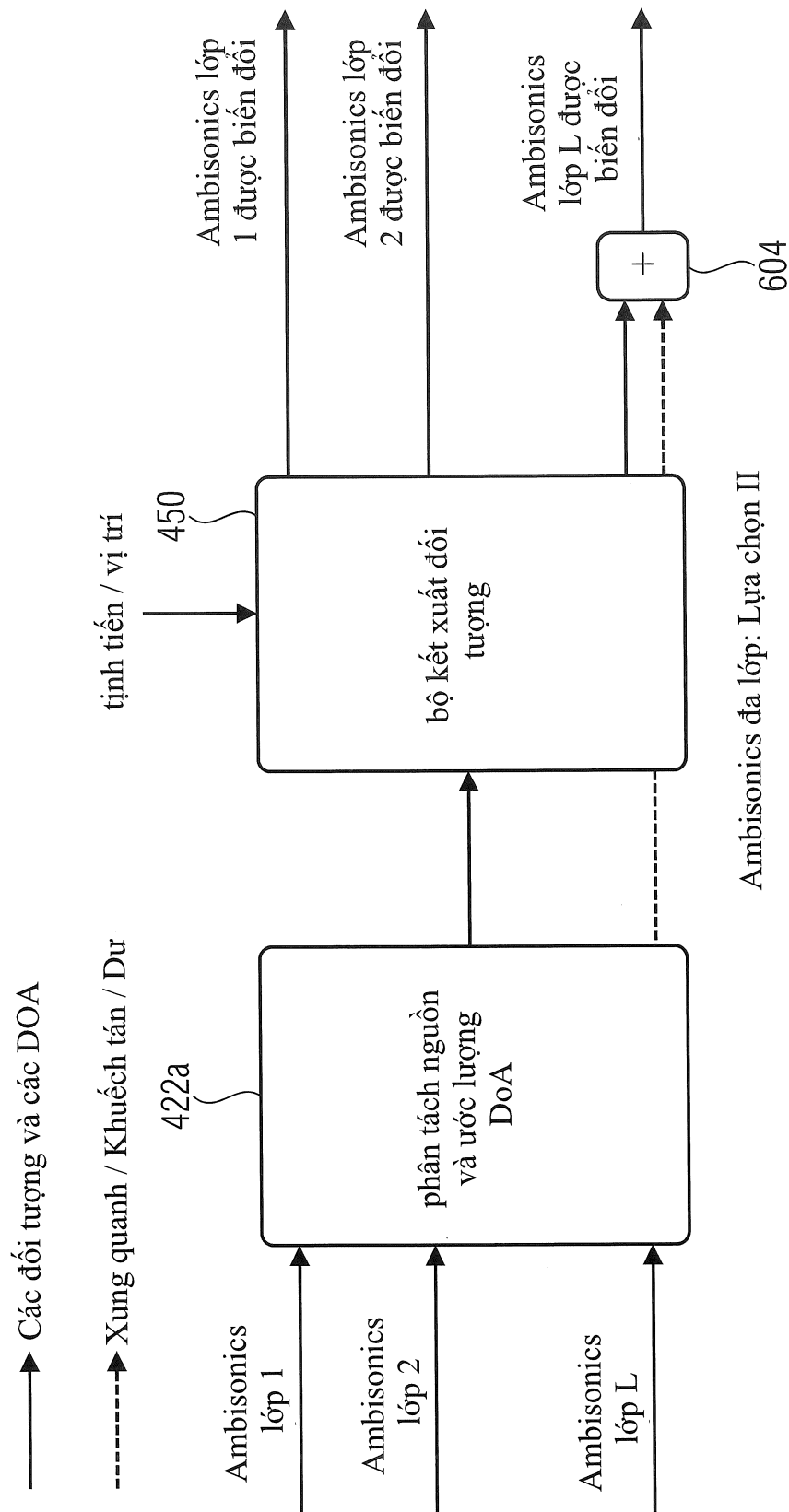
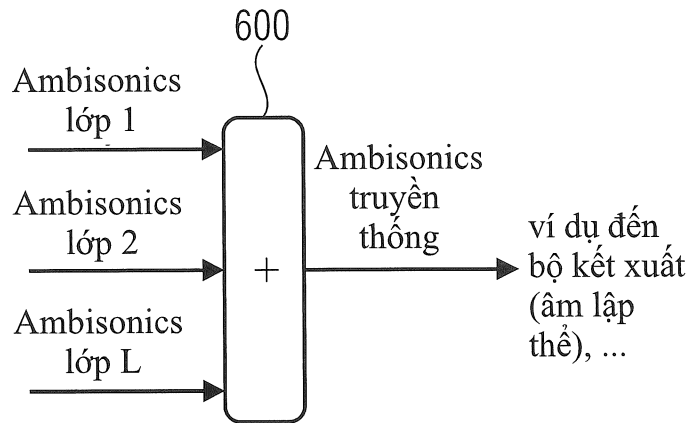


Fig. 4f

9/19



Chuyển đổi từ Ambisonics đa lớp thành Ambisonics truyền thống

Fig. 4g

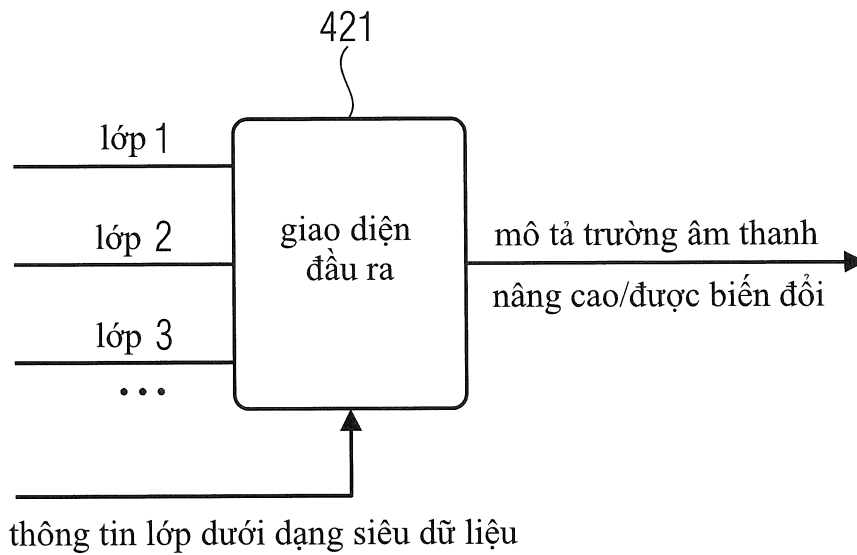
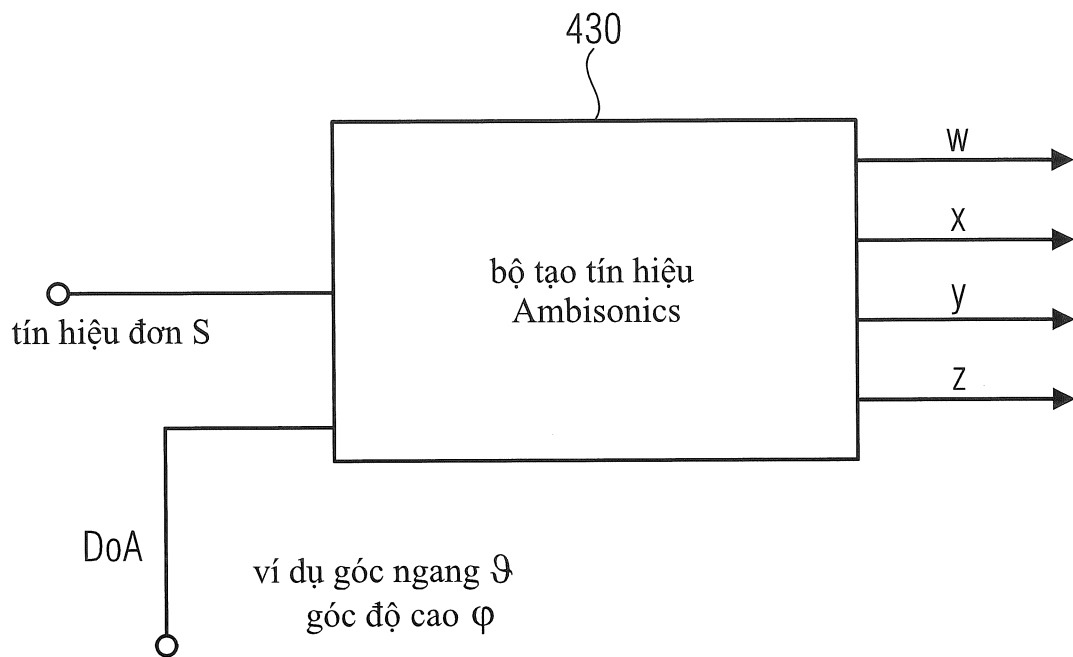


Fig. 4h

10/19



thành phần đẳng hướng $W = S \cdot \sqrt{\frac{1}{2}}$

các thành phần định hướng

$$\begin{cases} X = S \cdot \cos \vartheta \cos \varphi \\ Y = S \cdot \sin \vartheta \cos \varphi \\ Z = S \cdot \sin \varphi \end{cases}$$

Fig. 4i

11/19

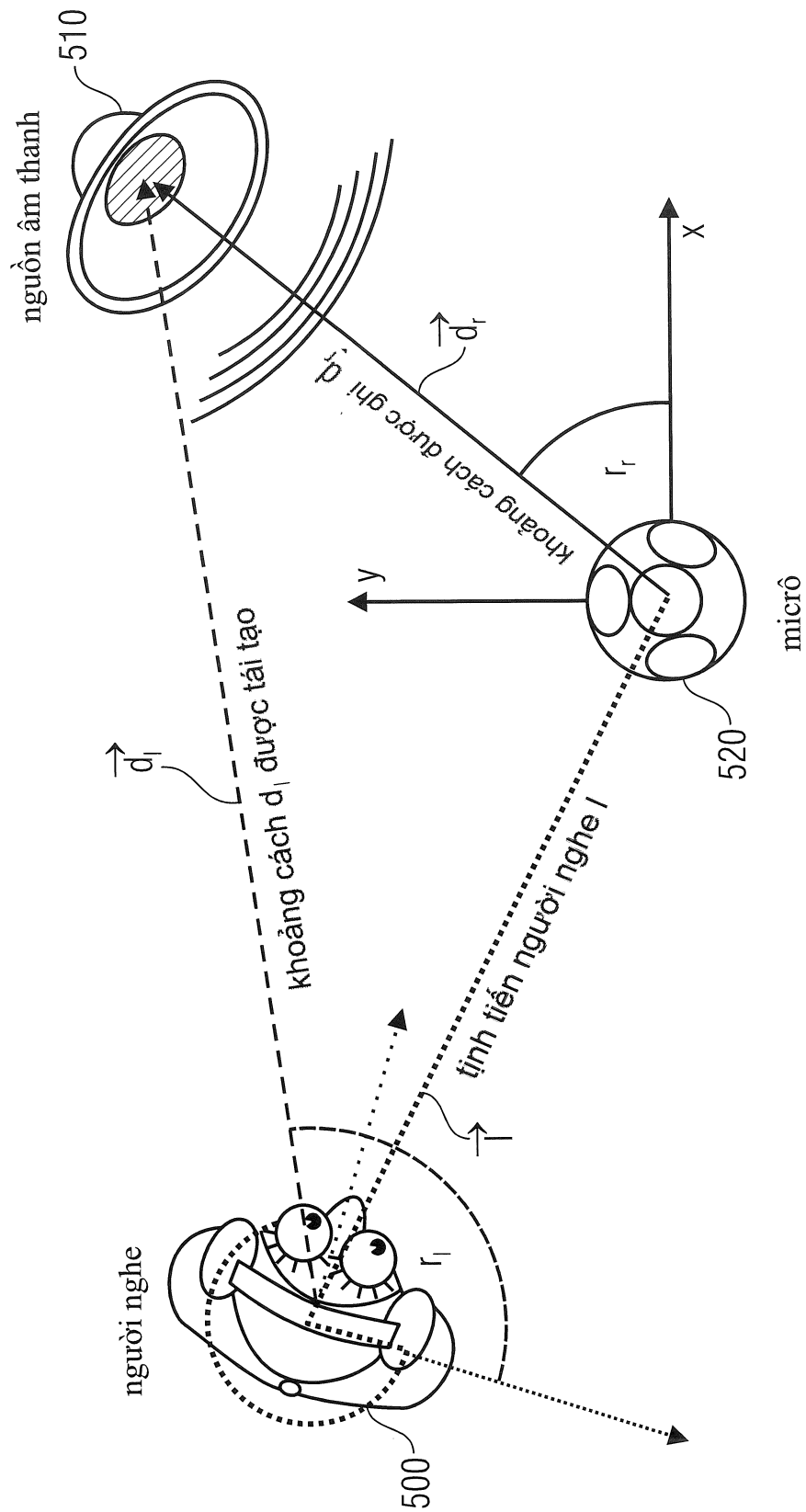


Fig. 5

12/19

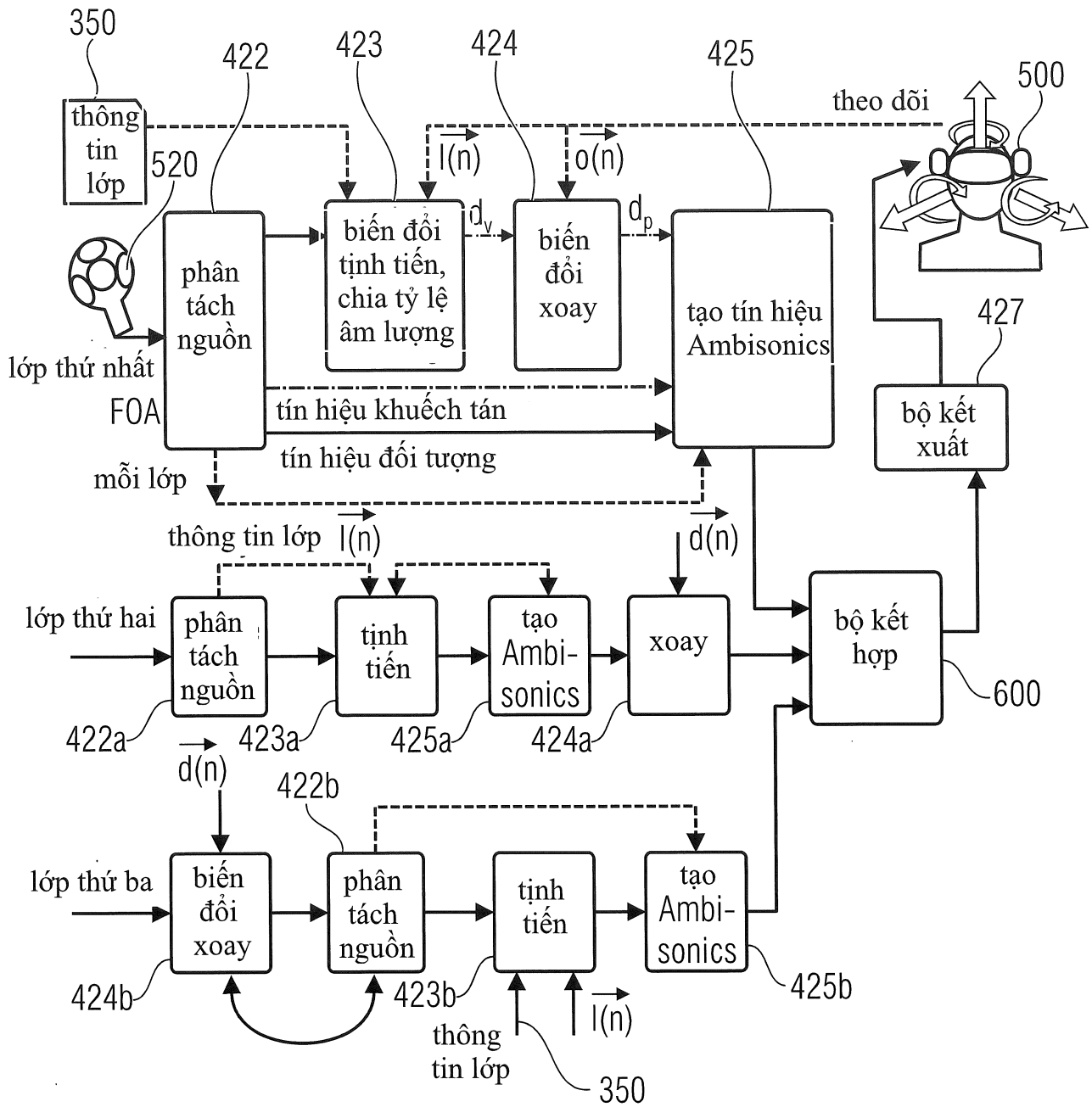


Fig. 6

13/19

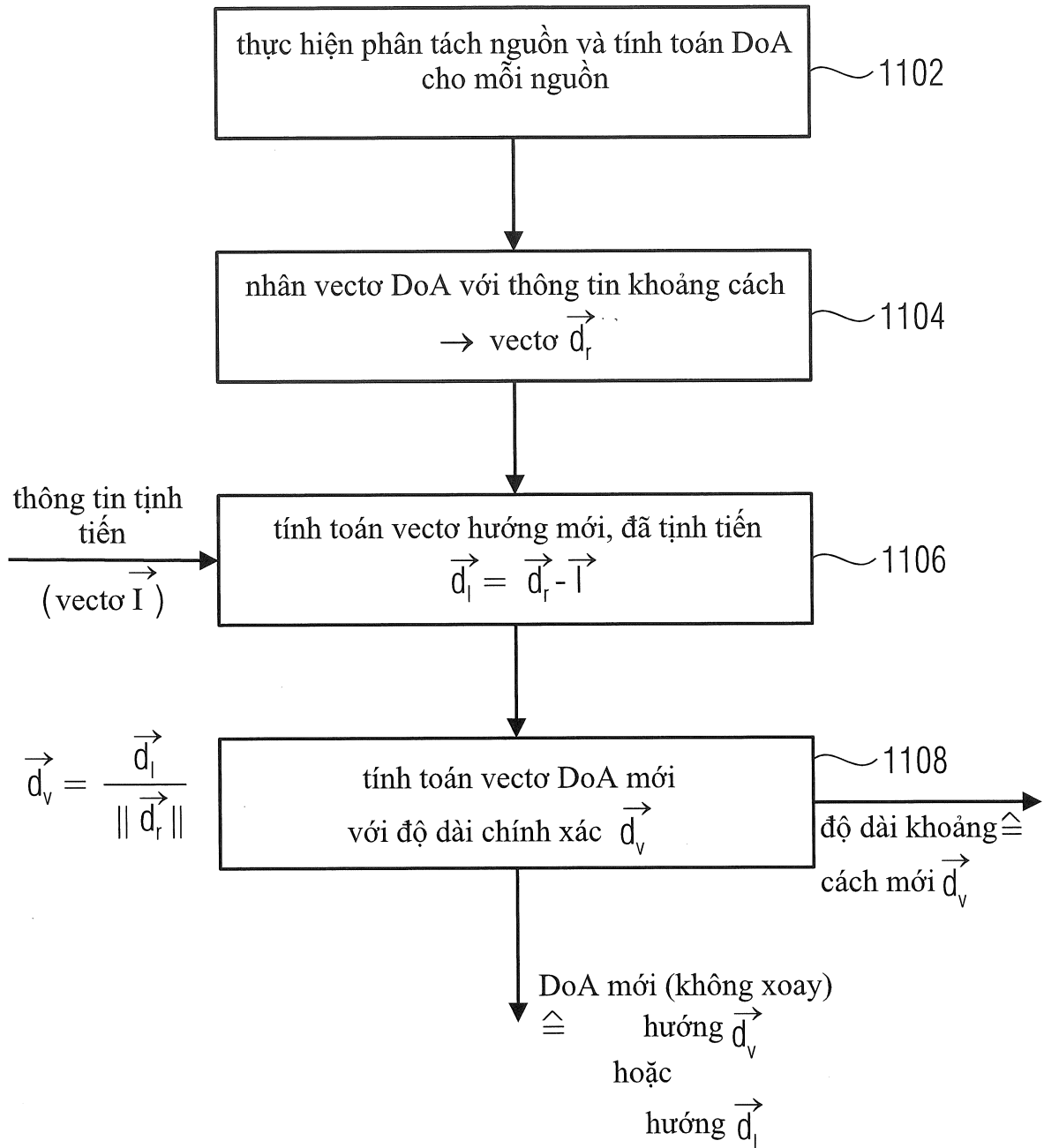


Fig. 7

14/19

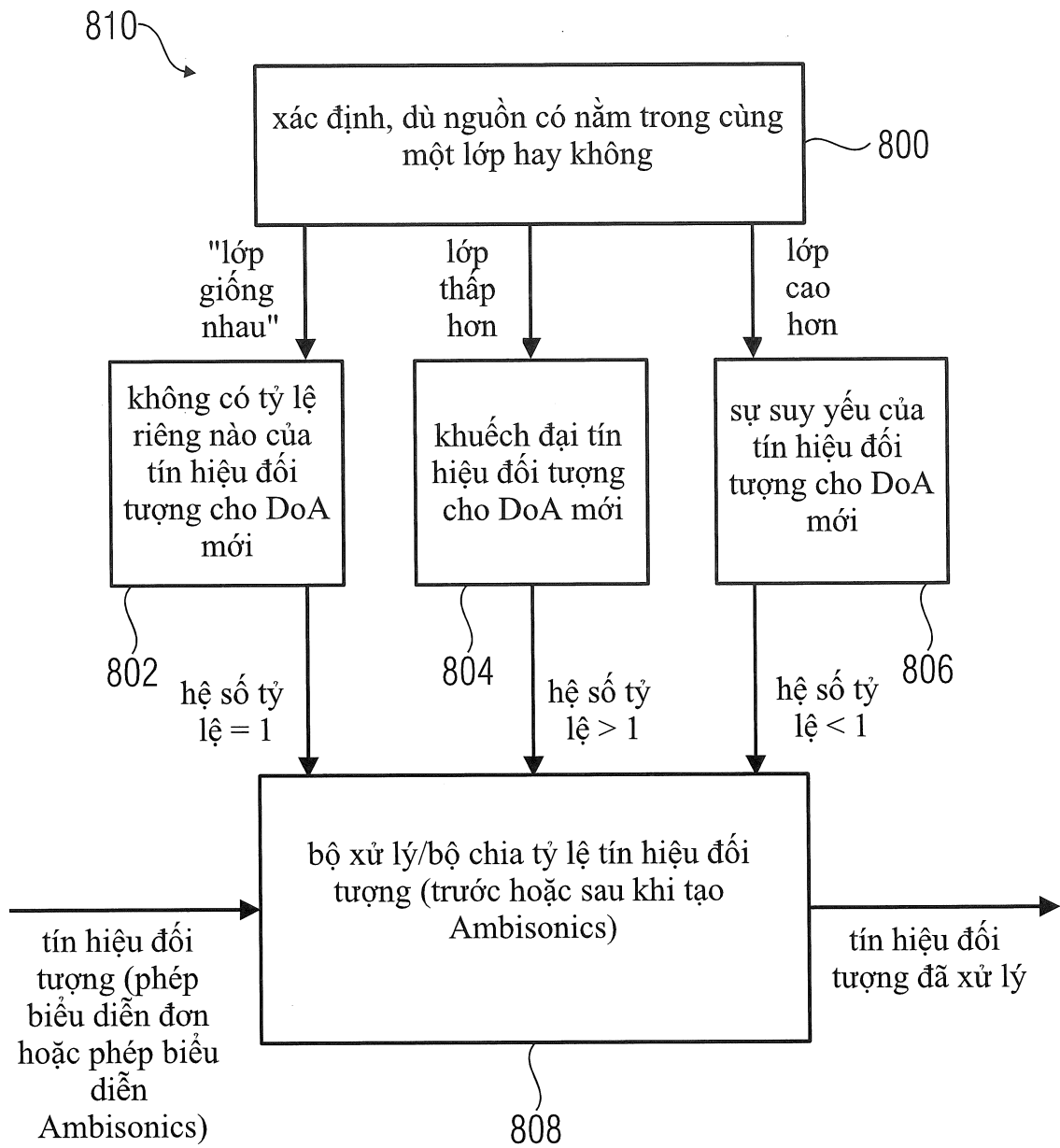


Fig. 8

15/19

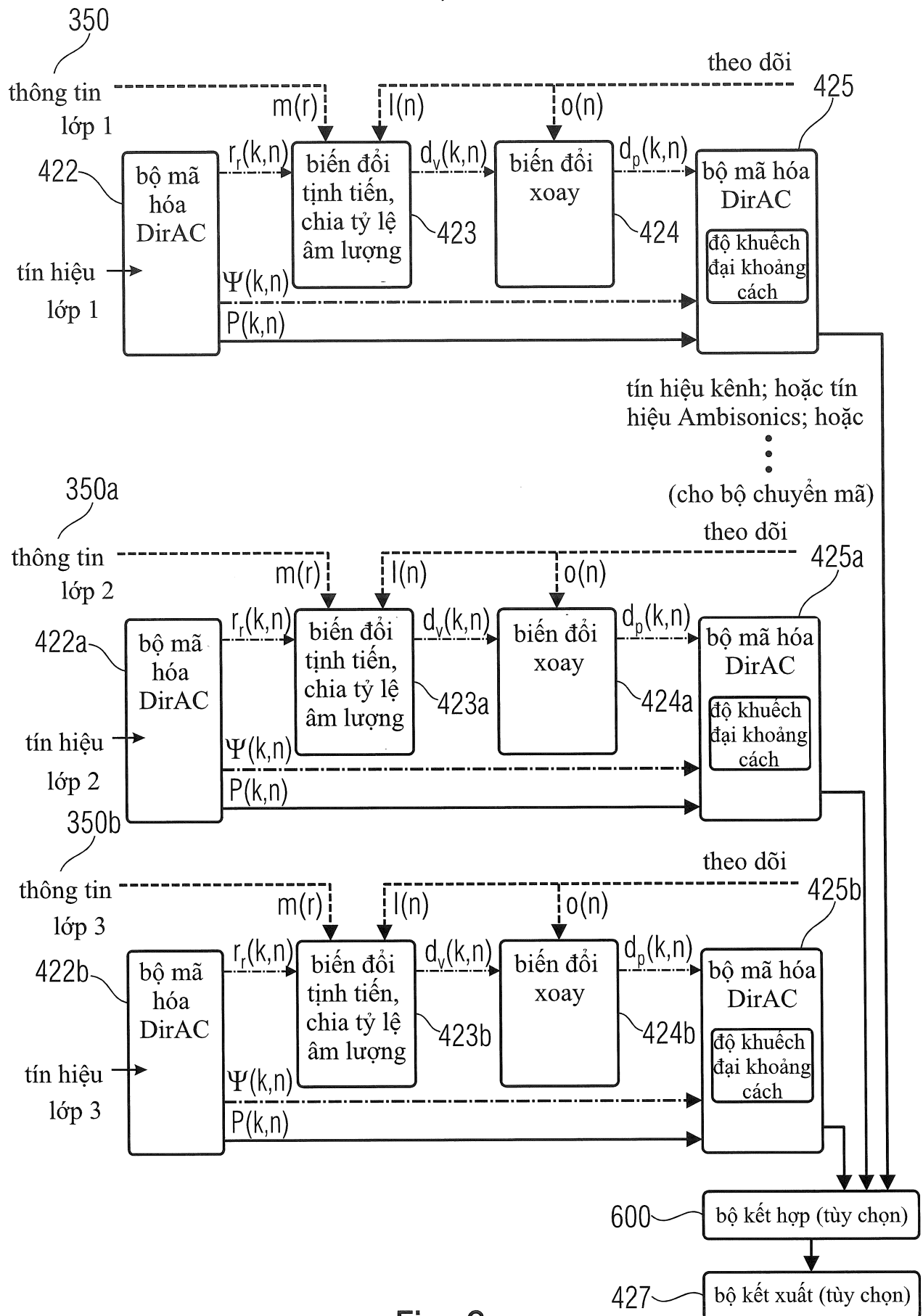


Fig. 9

16/19

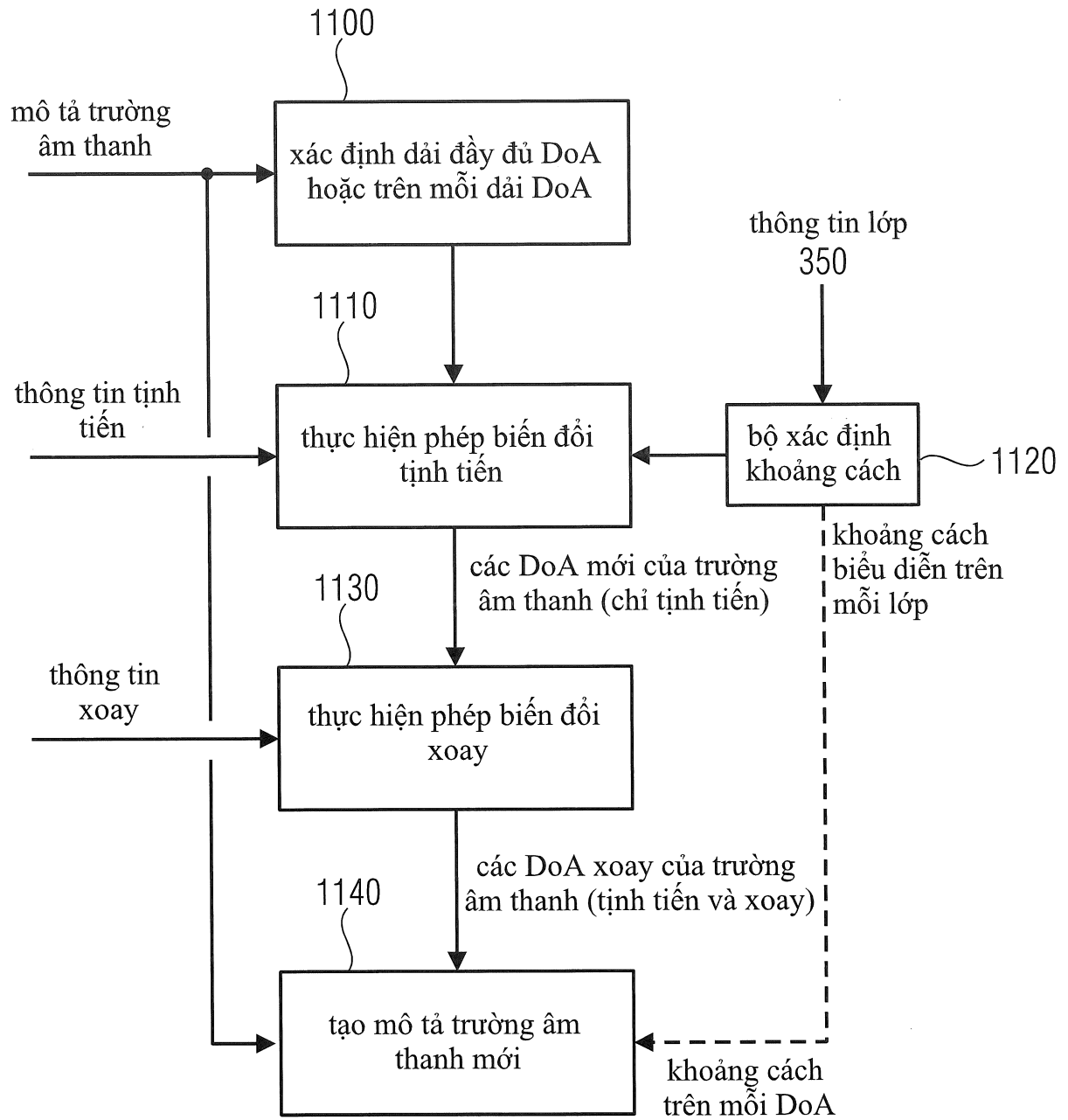


Fig. 10

17/19

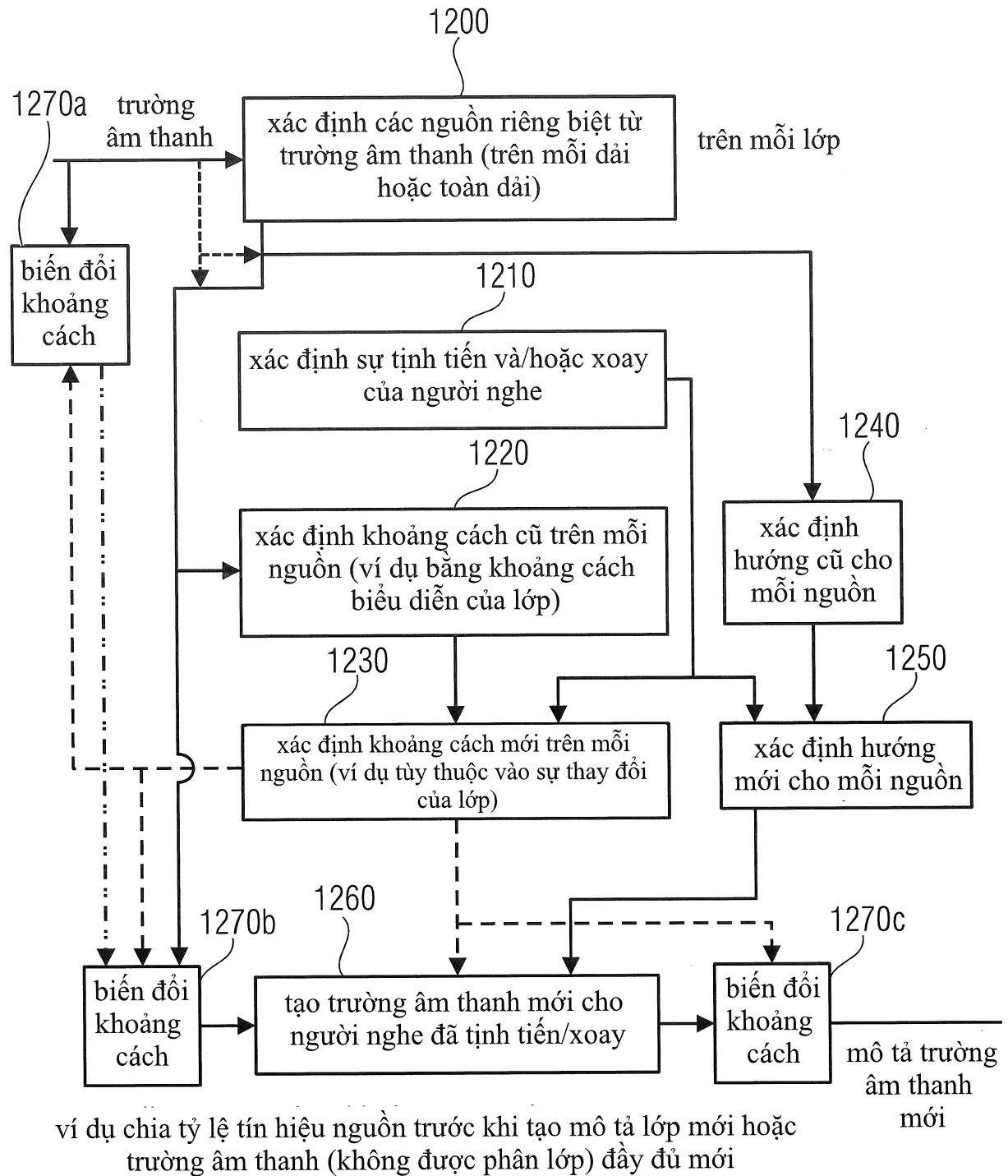
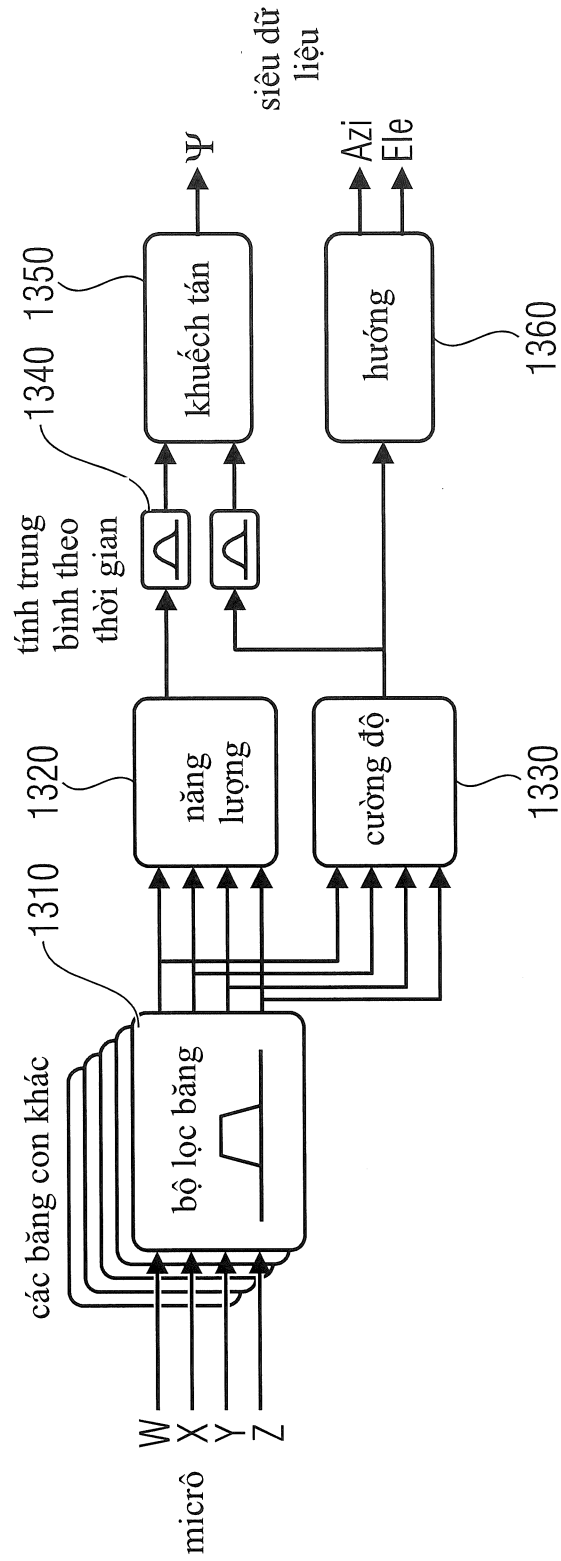
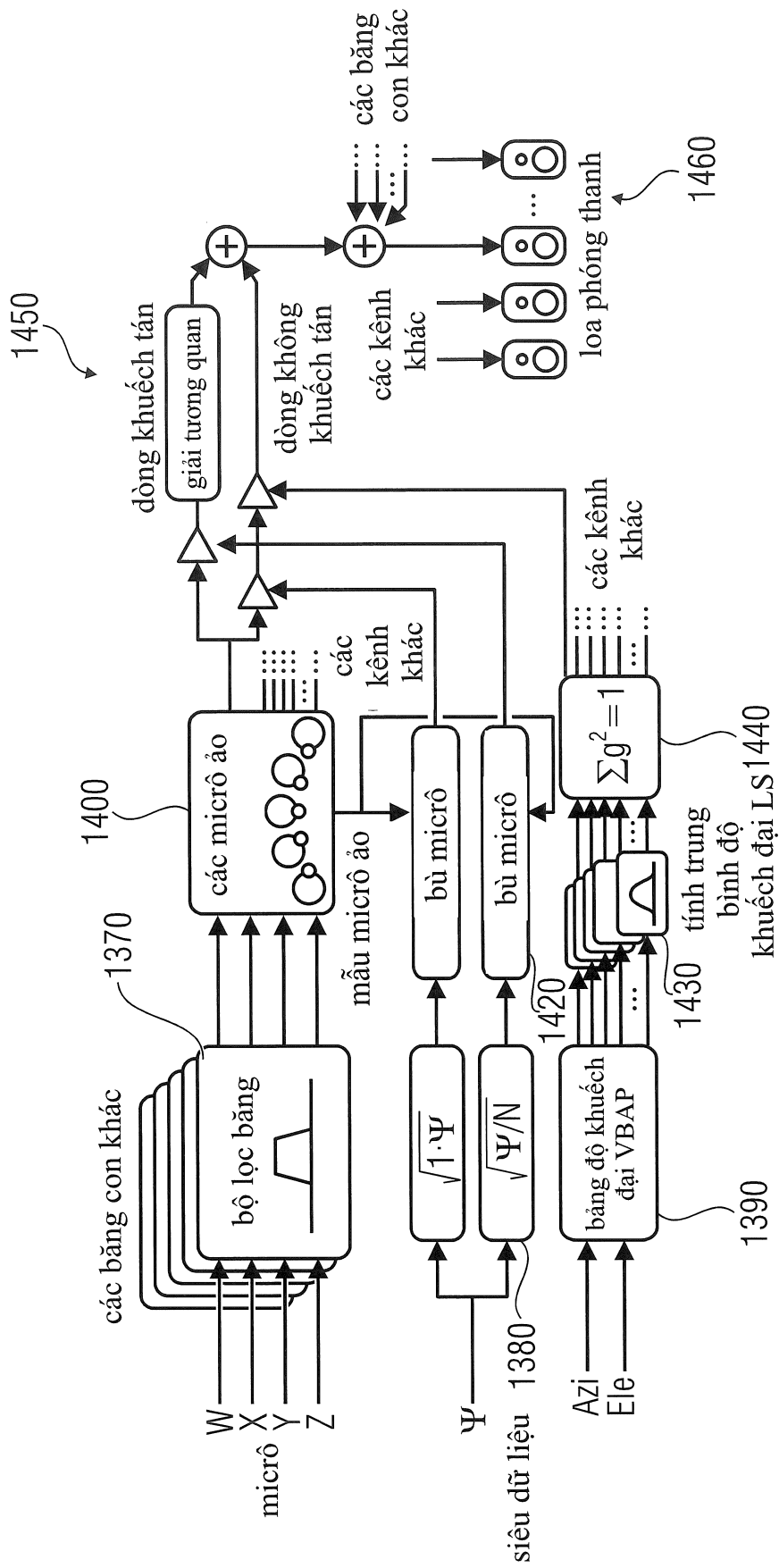


Fig. 11

18/19

**bộ phân tích DirAC****Fig. 12a**

19/19



bộ tổng hợp DirAC

Fig. 12b