



(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)



1-0035471

(51)⁷ H04R 3/00

(13) B

(21) 1-2018-01687

(22) 10/03/2017

(86) PCT/EP2017/055719 10/03/2017

(87) WO 2017/157803 21/09/2017

(30) 16160504.3 15/03/2016 EP

(45) 25/04/2023 421

(43) 27/08/2018 365A

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

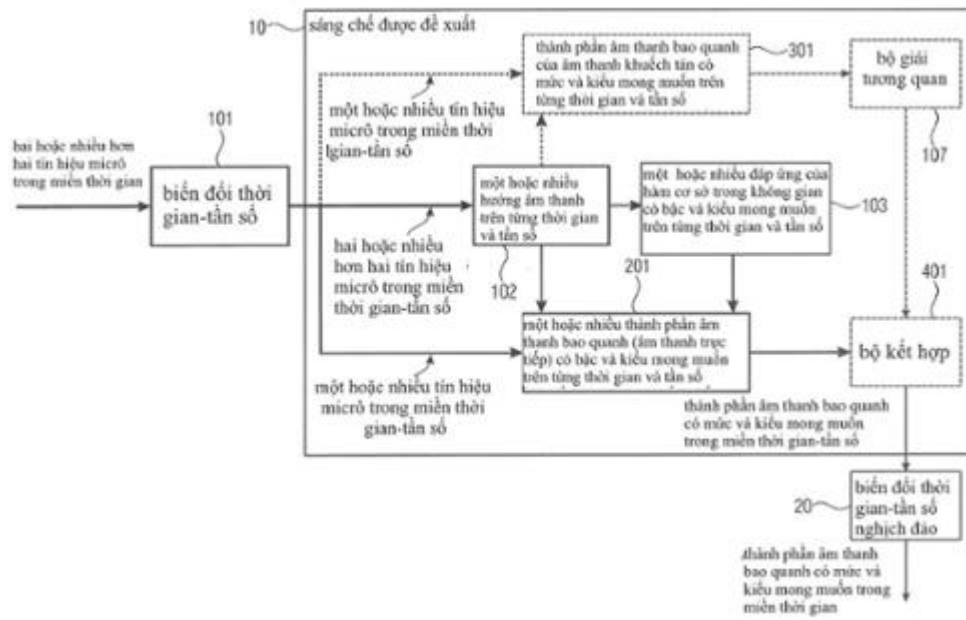
Hansastrasse 27c, 80686 Muenchen, Germany

(72) HABETS, Emanuel (NL); THIERGART, Oliver (DE); KUECH, Fabian (DE);
NIEDERLEITNER, Alexander (DE); KHAN, Affan-Hasan (PK); MAHNE, Dirk
(DE).

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

(54) THIẾT BỊ, PHƯƠNG PHÁP VÀ VẬT GHI CÓ THỂ ĐỌC ĐƯỢC BẰNG MÁY
TÍNH BAO GỒM CHƯƠNG TRÌNH MÁY TÍNH ĐỂ TẠO RA PHẦN MÔ TẢ
TRƯỜNG ÂM THANH

(57) Sáng chế đề cập đến thiết bị, phương pháp và vật ghi có thể đọc được bằng máy tính bao gồm chương trình máy tính để tạo ra phần mô tả trường âm thanh. Thiết bị tạo ra phần mô tả trường âm thanh có phép biểu diễn các thành trường âm thanh, bao gồm bộ xác định hướng (102) để xác định một hoặc nhiều hướng âm thanh cho từng ô thời gian-tần số thuộc nhiều ô thời gian-tần số của nhiều tín hiệu micro; bộ đánh giá hàm cơ sở trong không gian (103) để đánh giá, cho từng ô thời gian-tần số thuộc nhiều ô thời gian-tần số, một hoặc nhiều hàm cơ sở trong không gian sử dụng một hoặc nhiều hướng âm thanh; và bộ tính toán thành phần trường âm thanh (201) để tính toán, cho từng ô thời gian-tần số thuộc nhiều ô thời gian-tần số, một hoặc nhiều thành phần trường âm thanh tương ứng với một hoặc nhiều hàm cơ sở trong không gian được đánh giá sử dụng một hoặc nhiều hướng âm thanh và tín hiệu tham chiếu cho ô thời gian-tần số tương ứng, tín hiệu tham chiếu được suy ra từ một hoặc nhiều tín hiệu micro thuộc nhiều tín hiệu micro.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến thiết bị, phương pháp và vật ghi có thể đọc được bằng máy tính bao gồm chương trình máy tính để tạo ra phần mô tả trường âm thanh và còn tổng hợp các tín hiệu âm thanh bao quanh (bậc cao hơn) trong miền thời gian-tần số sử dụng thông tin hướng âm thanh.

Tính trạng kỹ thuật của sáng chế

Sáng chế thuộc lĩnh vực ghi âm và tái tạo âm thanh trong không gian. Việc ghi âm âm thanh trong không gian nhằm thu giữ trường âm thanh với nhiều micrô sao cho ở phía tái tạo, người nghe nhận biết được ảnh điện âm thanh tại vị trí ghi âm. Các cách tiếp cận thông thường đối với việc ghi âm âm thanh trong không gian thường sử dụng các micrô đẳng hướng được đặt cách nhau (ví dụ trong âm lập thể AB), hoặc các micrô cùng hướng (ví dụ trong âm lập thể có cường độ). Các tín hiệu được ghi có thể được tái tạo từ việc thiết lập loa phát thanh âm thanh nổi tiêu chuẩn để có được ảnh điện âm thanh của âm thanh nổi. Đối với việc tái tạo âm bao quanh, ví dụ, sử dụng thiết lập loa phát thanh 5.1, các kỹ thuật ghi tương tự có thể được sử dụng, ví dụ, năm micrô đường hình tim hướng thẳng về các vị trí của loa phát thanh trong mục tham khảo [ArrayDesign]. Gần đây, các hệ thống tái tạo âm thanh 3D đã xuất hiện, như thiết lập loa phát thanh 7.1+4, trong đó 4 loa đỉnh được sử dụng để tái tạo các âm cao. Các tín hiệu cho thiết lập loa phát thanh có thể được ghi ví dụ với các thiết lập loa phát thanh 3D được đặt cách quãng rất cụ thể trong mục tham khảo [MicSetup3D]. Tất cả các kỹ thuật ghi âm này có điểm chung là chúng được thiết kế cho một thiết lập loa cụ thể, điều này hạn chế khả năng ứng dụng thực tế, ví dụ, khi âm thanh đã ghi phải được tái tạo trên các cấu hình loa phát thanh khác nhau.

Sự linh hoạt hơn đạt được khi không ghi trực tiếp các tín hiệu cho thiết lập loa phát thanh cụ thể, mà thay vào đó ghi lại các tín hiệu có định dạng trung gian, từ đó

các tín hiệu của thiết lập loa phát thanh bất kỳ sau đó có thể được tạo ra ở phía tái tạo. Định dạng trung gian như vậy, mà được thiết lập tốt trong thực tiễn, được biểu diễn bởi âm thanh bao quanh (bậc cao hơn) trong mục tham khảo [Ambisonics]. Từ tín hiệu âm thanh bao quanh, có thể tạo ra các tín hiệu của từng thiết lập loa phát thanh mong muốn bao gồm các tín hiệu kép để tái tạo tai nghe. Việc này yêu cầu bộ kết xuất cụ thể được ứng dụng cho tín hiệu âm thanh bao quanh, như bộ kết xuất âm thanh bao quanh cổ điển trong mục tham khảo [Ambisonics], mã hóa âm thanh có hướng tính (Directional Audio Coding-DirAC) trong mục tham khảo [DirAC], hoặc HARPEX trong mục tham khảo [HARPEX].

Tín hiệu âm thanh bao quanh biểu diễn tín hiệu đa kênh trong đó từng kênh (được xem như thành phần âm thanh xung quanh) tương đương với cái gọi là hệ số của hàm cơ sở trong không gian. Với tổng được gán trọng số của các hàm cơ sở trong không gian này (với các trọng số tương ứng với các hệ số) có thể tạo ra trường âm thanh ban đầu ở vị trí ghi trong mục tham khảo [FourierAcoust]. Do đó, các hệ số của hàm cơ sở trong không gian (tức là các thành phần âm thanh bao quanh) biểu diễn phần mô tả thu gọn của trường âm thanh ở vị trí ghi. Có tồn tại các loại hàm cơ sở trong không gian khác nhau, ví dụ các hàm điều hòa cầu (spherical harmonics - SHs) trong mục tham khảo [FourierAcoust] hoặc hàm điều hòa trụ (cylindrical harmonics - CHs) trong mục tham khảo [FourierAcoust]. CHs có thể được sử dụng khi mô tả trường âm thanh trong không gian 2D (ví dụ đối với việc tái tạo âm thanh 2D) trong khi SHs có thể được sử dụng để mô tả trường âm thanh trong không gian 2D và 3D (ví dụ đối với việc tái tạo âm thanh 2D và 3D).

Các hàm cơ sở trong không gian tồn tại đối với các bậc l , và các kiểu m khác nhau trong các hàm cơ sở trong không gian 3D (như SHs). Trong trường hợp sau, có tồn tại các kiểu $m = 2l + 1$ cho từng bậc l , trong đó m và l là các số nguyên nằm trong khoảng $l \geq 0$ và $-l \leq m \leq l$. Ví dụ tương ứng về các hàm cơ sở trong không gian được thể hiện trên Fig.1a, mà thể hiện các hàm số của hàm điều hòa cầu cho các bậc l và các kiểu m khác nhau. Lưu ý là bậc l đôi khi được xem như các mức, và các kiểu m còn có thể được xem như các mức độ. Như có thể được thấy trên Fig.1a, hàm điều hòa cầu có

bậc không (mức thứ không) $l = 0$ biểu diễn áp suất âm thanh toàn hướng ở vị trí ghi, trong khi các hàm điều hòa cầu có bậc một (mức thứ nhất) $l = 1$ biểu diễn các thành phần lưỡng cực theo ba thứ nguyên của hệ tọa độ Đề-các. Điều này có nghĩa là, hàm cơ sở trong không gian có bậc (mức) cụ thể mô tả hướng tính của micrô có bậc l . Nói cách khác, hệ số của hàm cơ sở trong không gian tương ứng với tín hiệu của micrô có bậc (mức) l và kiểu m . Lưu ý là các hàm cơ sở trong không gian có các bậc và các kiểu khác nhau là trực giao với nhau. Điều này có nghĩa là ví dụ trong trường âm thanh khuếch tán hoàn toàn, các hệ số của tất cả các hàm cơ sở trong không gian không tương quan với nhau.

Như được giải thích ở trên, từng thành phần âm thanh bao quanh của tín hiệu âm thanh bao quanh tương ứng với hệ số của hàm cơ sở trong không gian có mức (và kiểu) cụ thể. Ví dụ, nếu trường âm thanh được mô tả đến mức $l = 1$ sử dụng SHs như là hàm cơ sở trong không gian, sau đó tín hiệu âm thanh bao quanh sẽ bao gồm bốn thành phần âm thanh bao quanh (vì chúng ta có một kiểu cho bậc $l = 0$ cộng với ba kiểu cho bậc $l = 1$). Các tín hiệu âm thanh bao quanh có bậc cao nhất $l = 1$ được xem như âm thanh bao quanh bậc một (first-order ambisonics - FOA) ở những phần sau, trong khi các tín hiệu âm thanh bao quanh có bậc cao nhất $l > 1$ được xem như âm thanh bao quanh bậc cao hơn (higher-order ambisonics - HOA). Khi sử dụng các bậc cao hơn l để mô tả trường âm thanh, độ phân giải trong không gian trở nên cao hơn, tức là, có thể mô tả hoặc khôi phục trường âm thanh với độ chính xác cao hơn. Do đó, có thể mô tả trường âm thanh với chỉ ít bậc hơn dẫn đến độ chính xác thấp (nhưng ít dữ liệu) hoặc có thể sử dụng các bậc cao hơn dẫn đến độ chính xác cao hơn (và dữ liệu nhiều hơn).

Có những định nghĩa toán học khác nhau nhưng liên quan chặt chẽ cho các hàm cơ sở trong không gian khác nhau. Ví dụ, có thể tính toán các hàm điều hòa cầu được định giá trị phức cũng như các hàm điều hòa cầu được định giá trị thực. Hơn nữa, các hàm điều hòa cầu có thể được tính toán với các thuật ngữ chuẩn hóa khác nhau như chuẩn hóa SN3D, N3D, hoặc N2D. Những định nghĩa khác nhau có thể được tìm thấy

ví dụ trong mục tham khảo [Ambix]. Một số ví dụ cụ thể sẽ được thể hiện sau cùng với phần mô tả của sáng chế và các phương án.

Tín hiệu âm thanh bao quanh được mong muốn có thể được xác định từ những hoạt động ghi với nhiều micrô. Cách đơn giản để thu được các tín hiệu âm thanh bao quanh là việc tính toán trực tiếp các thành phần âm thanh bao quanh (các hệ số của hàm cơ sở trong không gian) từ các tín hiệu của micrô. Cách tiếp cận này yêu cầu đo áp suất âm thanh tại các vị trí rất cụ thể, ví dụ trên đường tròn hoặc trên bề mặt của hình cầu. Sau đó, các hệ số của hàm cơ sở trong không gian có thể được tính toán bằng cách lấy tích phân đối với áp suất âm thanh đã đo, như được mô tả ví dụ trong mục tham khảo [FourierAcoust, p. 218]. Cách tiếp cận trực tiếp này yêu cầu thiết lập micrô cụ thể, ví dụ, dãy hình tròn hoặc dãy hình cầu của các micrô đẳng hướng. Hai ví dụ điển hình của các thiết lập micrô sẵn có trên thị trường là micrô SoundField ST350 hoặc EigenMike® trong mục tham khảo [EigenMike]. Không may là, các yêu cầu về hình học micrô cụ thể làm hạn chế nhiều đến khả năng ứng dụng trong thực tiễn, ví dụ khi các micrô cần được tích hợp thành thiết bị nhỏ hoặc nếu dãy micrô cần được kết hợp với máy quay video. Hơn nữa, việc xác định các hệ số trong không gian của bậc cao hơn với cách tiếp cận trực tiếp này yêu cầu số lượng micrô tương đối lớn để đảm bảo đủ mạnh chống lại nhiễu âm. Do đó, cách tiếp cận trực tiếp của việc thu tín hiệu âm thanh bao quanh thường rất đắt đỏ.

Tài liệu tham khảo

[Ambisonics] R. K. Furness, "Âm thanh bao quanh - An overview," in AES 8th International Conference, April 1990, pp. 181–189.

[Ambix] C. Nachbar, F. Zotter, E. Deleflie, and A. Sontacchi, "AMBIX - A Suggested Âm thanh bao quanh Format", Proceedings of the Âm thanh bao quanh Symposium 2011.

[ArrayDesign] M. Williams and G. Le Du, "Multichannel Microphone Array Design," in Audio Engineering Society Convention 108, 2008.

[CovRender] J. Vilkamo and V. Pulkki, "Minimization of Decorrelator Artifacts in Directional Audio Coding by Covariance Domain Rendering ", J. Audio Eng. Soc,

vol. 61, no. 9, 2013.

[DiffuseBF] O. Thiergart and E. A. P. Habets, "Extracting Reverberant Sound Using a Linearly Constrained Minimum Variance Spatial Filter," *IEEE Signal Processing Letters*, vol. 21, no. 5, May 2014.

[DirAC] V. Pulkki, "Directional audio coding in spatial sound reproduction and stereo upmixing," in *Proceedings of The AES 28th International Conference*, pp. 251-258, June, 2006.

[EigenMike] J. Meyer and T. Agnello, "Spherical microphone array for spatial sound recording," in *Audio Engineering Society Convention 115*, October 2003

[ERBsmooth] A. Favrot and C. Faller, "Perceptually Motivated Gain Filter Smoothing for Noise Suppression", *Audio Engineering Society Convention 123*, 2007.

[ESPRIT] R. Roy, A. Paulraj, and T. Kailath, "Direction-of-arrival estimation by subspace rotation methods – ESPRIT," in *IEEE International Conference on Acoustics, Speech, and Signal Processing(ICASSP)*, Stanford, CA, USA, April, 1986.

[FourierAcoust] E. G. Williams, "Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography," *Academic Press*, 1999.

[HARPEX] S. Berge and N. Barrett, "High Angular Resolution Planewave Expansion," in *2nd International Symposium on Âm thanh bao quanh and Spherical Acoustics*, May, 2010.

[InformedSF] O. Thiergart, M. Taseska, and E. A. P. Habets, "An Informed Parametric Spatial Filter Based on Instantaneous Direction-of-Arrival Estimates," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 22, no. 12, December 2014.

[MicSetup3D] H. Lee and C. Gribben, "On the optimum microphone array configuration for height channels," in *134 AES Convention*, Rome, 2013.

[MUSIC] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 276–280, 1986.

[OptArrayPr] B. D. Van Veen and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering", IEEE ASSP Magazine, vol. 5, no. 2, 1988.

[RootMUSIC1] B. Rao and K. Hari, "Performance analysis of root-MUSIC," in Signals, Systems and Computers, 1988. Twenty-Second Asilomar Conference on, vol. 2, 1988, pp. 578–582.

[RootMUSIC2] A. Mhamdi and A. Samet, "Direction of arrival estimation for nonuniform linear antenna," in Communications, Computing and Control Applications (CCCA), 2011 International Conference on, March 2011, pp. 1–5.

[RootMUSIC3] M. Zoltowski and C. P. Mathews, "Direction finding with uniform circular arrays via phase mode excitation and beamspace root-MUSIC," in Acoustics, Speech, and Signal Processing, 1992. ICASSP-92., 1992 IEEE International Conference on, vol. 5, 1992, pp. 245–248.

[SDRestim] O. Thiergart, G. Del Galdo, and E. A. P. Habets, "On the spatial coherence in mixed sound fields and its application to signal-to-diffuse ratio estimation", The Journal of the Acoustical Society of America, vol. 132, no. 4, 2012.

[SourceNum] J.-S. Jiang and M.-A. Ingram, "Robust detection of number of sources using the transformed rotational matrix," in Wireless Communications and Networking Conference, 2004. WCNC. 2004 IEEE, vol. 1, March, 2004.

[SpCoherence] D. P. Jarrett, O. Thiergart, E. A. P. Habets, and P. A. Naylor, "Coherence-Based Diffuseness Estimation in the Spherical Harmonic Domain," IEEE 27th Convention of Electrical and Electronics Engineers in Israel (IEEEI), 2012.

[SphHarm] F. Zotter, "Analysis and Synthesis of Sound-Radiation with Spherical Arrays", PhD thesis, University of Music and Performing Arts Graz, 2009.

[VirtualMic] O. Thiergart, G. Del Galdo, M. Taseska, and E. A. P. Habets, "Geometry-based Spatial Sound Acquisition Using Distributed Microphone Arrays," IEEE Transactions on Audio, Speech, and Language Processing, vol. 21, no. 12, Dec 2013.

Bản chất kỹ thuật của sáng chế

Mục đích của sáng chế là đề xuất khái niệm được cải thiện để tạo ra phần mô tả trường âm thanh có phép biểu diễn của các thành phần của trường âm thanh.

Mục đích của sáng chế đạt được bởi thiết bị theo điểm 1, phương pháp theo điểm 23 hoặc vật ghi có thể đọc được bằng máy tính bao gồm chương trình máy tính theo điểm 24.

Sáng chế đề cập đến thiết bị hoặc phương pháp hoặc vật ghi có thể đọc được bằng máy tính bao gồm chương trình máy tính để tạo ra phần mô tả trường âm thanh có phép biểu diễn của các thành phần trường âm thanh. Trong bộ xác định hướng, một hoặc nhiều hướng âm thanh đối với mỗi ô thời gian-tần số của nhiều ô thời gian-tần số của nhiều tín hiệu micrô được xác định. Bộ đánh giá hàm cơ sở trong không gian đánh giá, cho từng ô thời gian-tần số của nhiều ô thời gian-tần số, một hoặc nhiều hàm cơ sở trong không gian sử dụng một hoặc nhiều hướng âm thanh. Hơn nữa, bộ tính toán thành phần trường âm thanh tính toán, cho từng ô thời gian-tần số trong số nhiều ô thời gian-tần số, một hoặc nhiều thành phần trường âm thanh tương ứng với một hoặc nhiều hàm cơ sở trong không gian được đánh giá sử dụng một hoặc nhiều hướng âm thanh và sử dụng tín hiệu tham chiếu cho ô thời gian tần số tương ứng, trong đó tín hiệu tham chiếu được suy ra từ một hoặc nhiều tín hiệu micrô thuộc nhiều tín hiệu micrô.

Sáng chế dựa trên việc tìm ra rằng phần mô tả trường âm thanh mô tả trường âm thanh phức tạp tùy ý có thể được suy ra theo cách thức hiệu quả từ nhiều tín hiệu micrô trong phép biểu diễn thời gian-tần số chứa các ô thời gian-tần số. Các ô thời gian-tần số này một mặt đề cập đến nhiều tín hiệu micrô, và mặt khác được sử dụng để xác định các hướng âm thanh. Do đó, việc xác định hướng âm thanh diễn ra bên trong miền phổ sử dụng các ô thời gian-tần số của phép biểu diễn thời gian-tần số. Sau đó, phần lớn việc xử lý tiếp theo tốt hơn là được thực hiện trong phép biểu diễn thời gian-tần số. Để đạt được mục đích này, việc đánh giá các hàm cơ sở trong không gian được thực hiện sử dụng một hoặc nhiều hướng được xác định cho từng ô thời gian-tần số. Các hàm cơ sở trong không gian phụ thuộc vào các hướng âm thanh nhưng không phụ thuộc vào tần số. Do đó, việc đánh giá các hàm cơ sở trong không gian với các tín hiệu

miền tần số, tức là tín hiệu trong các ô thời gian-tần số được áp dụng. Trong cùng phép biểu diễn thời gian-tần số, một hoặc nhiều thành phần trường âm thanh tương ứng với một hoặc nhiều hàm cơ sở trong không gian mà đã được đánh giá sử dụng một hoặc nhiều hướng âm thanh được tính toán cùng với tín hiệu tham chiếu cũng có sẵn trong phép biểu diễn thời gian-tần số giống nhau.

Một hoặc nhiều thành phần trường âm thanh này cho từng khối và từng ngăn tần số của tín hiệu, tức là, đối với từng ô thời gian-tần số có thể là kết quả cuối cùng hoặc, một cách thay thế, sự chuyển đổi lại thành miền thời gian có thể được thực hiện để thu được một hoặc nhiều thành phần trường âm thanh trong miền thời gian tương ứng với một hoặc nhiều hàm cơ sở trong không gian. Phụ thuộc vào cách thực hiện, một hoặc nhiều thành phần trường âm thanh có thể là các thành phần trường âm thanh trực tiếp được xác định trong phép biểu diễn thời gian-tần số sử dụng các ô thời gian-tần số hoặc có thể các thành phần trường âm thanh khuếch tán thường sẽ được xác định bên cạnh các thành phần trường âm thanh trực tiếp. Các thành phần trường âm thanh cuối cùng có phần trực tiếp và phần khuếch tán sau đó có thể được thu bằng cách kết hợp các thành phần trường âm thanh trực tiếp với các thành phần trường âm thanh khuếch tán, trong đó sự kết hợp này có thể được thực hiện trong miền thời gian hoặc trong miền tần số phụ thuộc vào cách thực tiễn thực tế.

Các tiến trình khác nhau có thể được thực hiện để suy ra tín hiệu tham chiếu từ một hoặc nhiều tín hiệu micro. Các tiến trình này có thể bao gồm việc lựa chọn đơn giản tín hiệu micro nhất định từ nhiều tín hiệu micro hoặc lựa chọn được đề xuất dựa trên một hoặc nhiều hướng âm thanh. Việc xác định tín hiệu tham chiếu được đề xuất lựa chọn tín hiệu micro cụ thể từ nhiều tín hiệu micro tức là từ micro được đặt gần nhất với hướng âm thanh trong số các micro mà từ đó các tín hiệu micro đã được suy ra. Sự thay thế nữa là áp dụng bộ lọc đa kênh cho hai hoặc nhiều hơn hai tín hiệu micro để cùng lọc các tín hiệu micro đó để tín hiệu tham chiếu thông thường cho tất cả các ô tần số của khối thời gian được thu. Một cách thay thế, các tín hiệu tham chiếu khác nhau cho các ô tần số khác nhau trong khối thời gian có thể được suy ra. Dĩ nhiên là, các tín hiệu tham chiếu khác nhau cho các khối thời gian khác nhau mà không cho các tần số giống nhau trong các khối thời gian khác nhau cũng có thể được tạo ra. Do

đó, phụ thuộc vào cách thực hiện, tín hiệu tham chiếu cho ô thời gian-tần số có thể được lựa chọn một cách tự do hoặc được suy ra từ nhiều tín hiệu micrô.

Trong bối cảnh này, cần nhấn mạnh rằng micrô có thể được đặt ở các vị trí tùy ý. Các micrô cũng có thể có các đặc điểm hướng khác nhau. Hơn nữa, nhiều tín hiệu micrô không nhất thiết phải là tín hiệu đã được ghi lại bởi micrô vật lý thực. Thay vào đó, tín hiệu micrô có thể là tín hiệu micrô được tạo nhân tạo từ một trường âm thanh nhất định bằng cách sử dụng các thao tác xử lý dữ liệu nhất định bắt chước micrô vật lý thực.

Với mục đích xác định các thành phần trường âm thanh khuếch tán theo các phương án nhất định, các quy trình khác nhau có thể thực hiện được và hữu ích cho một số triển khai nhất định. Thông thường, một phần khuếch tán có nguồn gốc từ nhiều tín hiệu micrô làm tín hiệu tham chiếu và tín hiệu tham chiếu (khuếch tán) này sau đó được xử lý cùng với phản hồi trung bình của hàm cơ sở trong không gian có bậc nhất định (hoặc mức và/hoặc kiểu) để thu được thành phần âm thanh khuếch tán cho bậc hoặc mức hoặc kiểu này. Do đó, thành phần âm thanh trực tiếp được tính toán sử dụng sự đánh giá về hàm cơ sở trong không gian nhất định với hướng đến nhất định và thành phần âm thanh khuếch tán, vốn dĩ, không được tính toán sử dụng hướng đến nhất định nhưng được tính toán bằng cách sử dụng tín hiệu tham chiếu khuếch tán và bằng cách kết hợp tín hiệu tham chiếu khuếch tán và đáp ứng trung bình của hàm cơ sở trong không gian có bậc hoặc mức hoặc kiểu nhất định bởi hàm nhất định. Việc kết hợp hàm số này có thể, ví dụ, là phép nhân vì còn có thể được thực hiện trong phép tính của thành phần âm thanh trực tiếp hoặc kết hợp này có thể là phép nhân được gán trọng số hoặc phép cộng hoặc phép trừ, ví dụ khi các phép tính trong miền lôgarit được thực hiện. Các kết hợp khác khác với phép nhân hoặc phép cộng/phép trừ được thực hiện sử dụng hàm tuyến tính hoặc phi tuyến tính nữa, trong đó các hàm phi tuyến tính được ưu tiên hơn. Nối tiếp với việc tạo thành thành phần trường âm thanh trực tiếp và thành phần trường âm thanh khuếch tán có bậc nhất định, việc kết hợp có thể được thực hiện bằng cách kết hợp thành phần trường âm thanh trực tiếp với thành phần trường âm thanh khuếch đại bên trong miền phổ cho từng ô thời gian/tần số riêng rẽ. Một cách thay thế, các thành phần trường âm thanh khuếch tán và các thành phần

trường âm thanh trực tiếp cho bậc nhất định có thể được biến đổi từ miền tần số thành miền thời gian và sau đó việc kết hợp miền thời gian của thành phần miền thời gian trực tiếp và thành phần miền thời gian khuếch tán có bậc nhất định cũng có thể được thực hiện.

Tùy thuộc vào trạng thái, các bộ giải tương quan khác có thể được sử dụng để giải tương quan các thành phần trường âm thanh khuếch tán. Một cách thay thế, các thành phần trường âm thanh khuếch tán được giải tương quan có thể được tạo ra bằng cách sử dụng các tín hiệu micrô khác nhau hoặc các ngăn thời gian/tần số khác nhau cho các thành phần trường âm thanh khuếch tán khác nhau có các bậc khác nhau hoặc bằng cách sử dụng tín hiệu micrô khác nhau cho việc tính toán thành phần trường âm thanh trực tiếp và tín hiệu micrô khác nhau nữa cho việc tính toán thành phần trường âm thanh khuếch tán.

Trong phương án được ưu tiên, các hàm cơ sở trong không gian là các hàm cơ sở trong không gian được kết hợp với các mức (các bậc) và các kiểu nhất định của phần mô tả trường âm thanh của âm thanh bao quanh đã biết. Thành phần trường âm thanh có bậc nhất định và kiểu nhất định sẽ tương ứng với thành phần trường âm thanh của âm thanh bao quanh được kết hợp với mức nhất định và kiểu nhất định. Thông thường, thành phần trường âm thanh thứ nhất sẽ là thành phần trường âm thanh được kết hợp với hàm cơ sở trong không gian đẳng hướng như được biểu thị trên Fig.1 đối với bậc $l=0$ và kiểu $m=0$.

Thành phần trường âm thanh thứ hai ví dụ có thể được kết hợp với hàm cơ sở trong không gian có hướng tính cực đại trong hướng x tương ứng với bậc $l=1$ và kiểu $m=-1$ đối với Fig.1a. Thành phần trường âm thanh thứ ba ví dụ có thể là hàm cơ sở trong không gian được định hướng theo hướng y mà sẽ tương ứng với kiểu $m=0$ và bậc $l=1$ thuộc Fig.1a và thành phần trường âm thanh thứ tư ví dụ có thể là hàm cơ sở trong không gian định hướng theo hướng z tương ứng với kiểu $m=1$ và bậc $l=1$ thuộc Fig.1a.

Tuy nhiên, những phần mô tả trường âm thanh khác ngoài âm thanh bao quanh, tất nhiên, đã được biết đến bởi người có hiểu biết trong lĩnh vực kỹ thuật tương ứng và các thành phần trường âm thanh khác dựa trên các hàm cơ sở trong không gian khác nhau từ các hàm cơ sở trong không gian của âm thanh bao quanh còn có thể được tính

toán thuận lợi bên trong phép biến đổi miền thời gian-tần số như được thảo luận trước.

Các phương án thuộc sáng chế mô tả phương thức thực tiễn để thu được các tín hiệu âm thanh bao quanh. Ngược lại với các cách tiếp cận thuộc tình trạng kỹ thuật đã được nói ở trên, cách tiếp cận của sáng chế có thể được áp dụng cho các thiết lập micro tùy ý mà xử lý hai hoặc nhiều hơn hai micro. Hơn nữa, các thành phần âm thanh bao quanh có các bậc cao hơn có thể được tính toán sử dụng chỉ vài micro tương ứng. Do đó, cách tiếp cận của sáng chế là tương đối rẻ và thiết thực. Trong phương án được đề xuất, các thành phần âm thanh bao quanh không được tính toán trực tiếp từ thông tin áp suất âm thanh theo bề mặt cụ thể, như đối với các cách tiếp cận trong tình trạng kỹ thuật được giải thích ở trên, mà chúng được tổng hợp dựa trên cách tiếp cận theo tham số. Với mục đích này, mô hình trường âm thanh khá đơn giản được giả định, tương tự như mô hình được sử dụng ví dụ trong DirAC trong mục tham khảo [DirAC]. Chính xác hơn, giả sử trường âm thanh ở vị trí ghi chứa một hoặc vài âm thanh trực tiếp đến từ các hướng âm thanh cụ thể cộng với âm thanh khuếch tán đến từ tất cả các hướng. Dựa trên mô hình này, và bằng cách sử dụng thông tin theo tham số trên trường âm thanh như hướng âm thanh của các âm thanh trực tiếp, có khả năng tổng hợp các thành phần âm thanh bao quanh hoặc các thành phần trường âm thanh khác bất kỳ chỉ từ vài phép đo áp suất âm thanh. Cách tiếp cận của sáng chế được giải thích chi tiết trong các phần sau đây.

Mô tả vắn tắt các hình vẽ

Các phương án được ưu tiên của sáng chế lần lượt được giải thích với các hình vẽ kèm theo, trong đó:

Fig.1a thể hiện các hàm sóng hài hình cầu cho các bậc và các kiểu khác nhau;

Fig.1b thể hiện ví dụ về cách lựa chọn micro tham chiếu dựa trên thông tin hướng đến;

Fig.1c thể hiện cách thực hiện được ưu tiên của thiết bị hoặc phương pháp để tạo ra phần mô tả trường âm thanh;

Fig.1d minh họa sự chuyển đổi thời gian-tần số của tín hiệu micro làm mẫu trong đó các ô thời gian-tần số cụ thể (10, 1) cho ngăn tần số 10 và khối thời gian 1 và ô (5, 2) cho ngăn tần số 5 và khối thời gian 2 được biểu thị một cách cụ thể;

Fig.1e minh họa sự đánh giá của bốn hàm cơ sở trong không gian làm mẫu sử dụng các hướng âm thanh cho các ngăn tần số được biểu thị (10, 1) và (5, 2);

Fig.1f minh họa việc tính toán các thành phần trường âm thanh cho hai ngăn (10, 1) and (5, 2) và việc chuyển đổi tần số-thời gian tiếp theo và việc xử lý nhiễu xuyên kênh/chồng lấp-bổ sung;

Fig.1g minh họa phép biểu diễn trong miền thời gian của bốn thành phần trường âm thanh làm mẫu b_1 đến b_4 như được thu bởi việc xử lý thuộc Fig.1f;

Fig.2a thể hiện sơ đồ khối chung của sáng chế;

Fig.2b thể hiện sơ đồ khối chung của sáng chế trong đó phép biến đổi thời gian-tần số nghịch đảo được áp dụng trước bộ kết hợp;

Fig.3a thể hiện phương án của sáng chế trong đó thành phần âm thanh bao quanh có mức và kiểu mong muốn được tính toán từ tín hiệu micro tham chiếu và thông tin hướng âm thanh;

Fig.3b thể hiện phương án của sáng chế trong đó micro tham chiếu được lựa chọn dựa trên thông tin hướng đến;

Fig.4 thể hiện phương án của sáng chế trong đó thành phần âm thanh bao quanh của âm thanh trực tiếp và thành phần âm thanh bao quanh của âm thanh khuếch tán được tính toán;

Fig.5 thể hiện phương án của sáng chế trong đó thành phần âm thanh bao quanh của âm thanh khuếch tán được giải tương quan;

Fig.6 thể hiện phương án của sáng chế trong đó âm thanh trực tiếp và âm thanh khuếch tán được trích xuất từ nhiều micro và thông tin hướng âm thanh;

Fig.7 thể hiện phương án của sáng chế trong đó âm thanh khuếch tán được trích xuất từ nhiều micro và trong đó thành phần âm thanh bao quanh của âm thanh khuếch tán được giải tương quan; và

Fig.8 thể hiện phương án của sáng chế trong đó việc làm nhấn độ khuếch đại được áp dụng cho đáp ứng của hàm cơ sở trong không gian.

Mô tả chi tiết sáng chế

Phương án ưu tiên được minh họa trên Fig.1c. Fig.1c minh họa phương án của thiết bị hoặc phương pháp để tạo ra phần mô tả trường âm thanh 130 có phép biểu diễn của các thành phần trường âm thanh như phép biểu diễn trong miền thời gian của các thành phần trường âm thanh hoặc phép biểu diễn trong miền tần số của các thành phần trường âm thanh, phép biểu diễn được mã hóa hoặc được giải mã hoặc phép biểu diễn trung gian.

Để đạt được mục đích này, bộ xác định hướng 102 xác định một hoặc nhiều hướng âm thanh 131 cho từng ô thời gian-tần số thuộc nhiều ô thời gian-tần số của nhiều tín hiệu micrô.

Do đó, bộ xác định hướng nhận, tại đầu vào của nó 132, ít nhất hai tín hiệu micrô khác nhau và, đối với từng tín hiệu trong số hai tín hiệu khác nhau đó, phép biểu diễn thời gian-tần số thường chứa các khối ngăn phổ liên tiếp sẵn có trong đó khối gồm các ngăn phổ đã được kết hợp chỉ số thời gian nhất định n , trong đó chỉ số tần số là k . Khối gồm các ngăn tần số đối với chỉ số thời gian biểu diễn phổ của tín hiệu miền thời gian cho khối có các mẫu miền thời gian được tạo ra bởi hoạt động tạo cửa sổ nhất định.

Các hướng âm thanh 131 được sử dụng bởi bộ đánh giá hàm cơ sở trong không gian 103 để đánh giá một hoặc nhiều hàm cơ sở trong không gian cho từng ô thời gian-tần số thuộc nhiều ô thời gian-tần số. Do đó, kết quả của việc xử lý trong khối 103 là một hoặc nhiều hàm cơ sở trong không gian được đánh giá cho từng ô thời gian-tần số. Tốt hơn là, hai hoặc thậm chí nhiều hơn hai hàm cơ sở trong không gian khác nhau được sử dụng như bốn hàm cơ sở trong không gian như được mô tả đối với các hình vẽ Fig.1e và Fig.1f. Do đó, tại đầu ra 133 của khối 103, các hàm cơ sở trong không gian được đánh giá có các bậc và các kiểu khác nhau cho các ô thời gian-tần số khác nhau của phép biểu diễn thời gian- phổ là có thể sử dụng được và được nhập vào trong bộ tính toán thành phần trường âm thanh 201. Bộ tính toán thành phần trường âm thanh 201 sử dụng thêm tín hiệu tham chiếu 134 được tạo ra bởi bộ tính toán tín hiệu tham chiếu (không được thể hiện trên Fig.1c). Tín hiệu tham chiếu 134 được suy ra từ một hoặc nhiều tín hiệu micrô của nhiều tín hiệu micrô và được sử dụng bởi bộ tính toán thành phần trường âm thanh bên trong phép biểu diễn thời gian/tần số giống nhau.

Do đó, bộ tính toán thành phần trường âm thanh 201 được tạo cấu hình để tính toán, cho từng ô thời gian-tần số thuộc nhiều ô thời gian-tần số, một hoặc nhiều thành phần trường âm thanh tương ứng với một hoặc nhiều hàm cơ sở trong không gian được đánh giá sử dụng một hoặc nhiều hướng âm thanh với sự giúp đỡ của một hoặc nhiều tín hiệu tham chiếu cho ô thời gian-tần số tương ứng.

Phụ thuộc vào cách thực hiện, bộ đánh giá hàm cơ sở không gian 103 được tạo cấu hình để sử dụng, đối với hàm cơ sở trong không gian, phép biểu diễn được tham số hóa, trong đó tham số của phép biểu diễn được tham số hóa là hướng âm thanh, hướng âm thanh là một chiều trong trạng thái không gian hai chiều hoặc hai chiều trong trạng thái không gian ba chiều, và đề chèn tham số tương ứng với hướng âm thanh vào trong phép biểu diễn được tham số hóa để thu được kết quả đánh giá cho từng hàm cơ sở trong không gian.

Một cách thay thế, bộ đánh giá hàm cơ sở không gian được tạo cấu hình để sử dụng bảng tìm kiếm cho từng hàm cơ sở trong không gian có sự nhận biết hàm cơ sở trong không gian và hướng âm thanh làm đầu vào và có kết quả đánh giá làm đầu ra. Trong trạng thái này, bộ đánh giá hàm cơ sở trong không gian được tạo cấu hình để xác định, cho một hoặc nhiều hướng âm thanh được xác định bởi bộ xác định hướng 102, hướng âm thanh tương ứng của đầu vào bảng tìm kiếm. Diễn hình là, các đầu vào của hướng khác nhau được lượng tử hóa theo cách để, ví dụ, số lượng nhất định của các đầu vào của bảng có mười hướng âm thanh khác nhau.

Bộ đánh giá hàm cơ sở trong không gian 103 được tạo cấu hình để xác định, đối với âm thanh cụ thể nhất định không trùng khớp ngay với đầu vào hướng âm thanh cho bảng tìm kiếm, đầu vào bảng tìm kiếm tương ứng. Ví dụ, nó có thể được thực hiện bằng cách sử dụng, cho hướng âm thanh được xác định nhất định, đầu vào hướng âm thanh cao hơn kế tiếp hoặc đầu vào hướng âm thanh thấp hơn kế tiếp vào trong bảng tìm kiếm. Một cách thay thế, bảng được sử dụng theo cách mà giá trị trung bình được gán trọng số giữa hai đầu vào của bảng tìm kiếm liên kế được tính toán. Do đó, tiến trình sẽ là đầu ra của bảng cho đầu vào có hướng thấp hơn kế tiếp được xác định. Hơn nữa, đầu ra của bảng tìm kiếm cho đầu vào cao hơn kế tiếp được xác định và sau đó mức trung bình các giá trị của chúng được tính toán.

Mức trung bình này có thể là mức trung bình đơn giản được thu bằng cách thêm hai đầu vào và chia các kết quả cho hai hoặc có thể là mức trung bình được gán trọng số phụ thuộc vào vị trí của hướng âm thanh được xác định đối với đầu ra của bảng cao hơn kế tiếp hoặc đầu ra của bảng thấp hơn kế tiếp. Do đó, một cách điển hình, thừa số được gán trọng số sẽ phụ thuộc vào sự chênh lệch giữa hướng âm thanh được xác định và đầu vào cao hơn kế tiếp/thấp hơn kế tiếp vào trong bảng tìm kiếm. Ví dụ, khi hướng được đo gần với đầu vào thấp hơn kế tiếp sau đó kết quả của bảng tìm kiếm cho đầu vào thấp hơn kế tiếp được nhân với thừa số gán trọng số cao hơn khi so sánh với thừa số gán trọng số, với đầu ra của bảng tìm kiếm cho đầu vào cao hơn kế tiếp được gán trọng số. Do đó, đối với chênh lệch nhỏ giữa hướng được xác định với đầu vào thấp hơn kế tiếp, đầu ra của bảng tìm kiếm đối với đầu vào thấp hơn kế tiếp sẽ được gán trọng số với thừa số gán trọng số cao hơn khi so sánh với thừa số gán trọng số được sử dụng cho việc gán trọng số đầu ra của bảng tìm kiếm tương ứng với đầu vào của bảng tìm kiếm cao hơn kế tiếp cho hướng âm thanh.

Tiếp theo, các hình vẽ từ Fig.1d đến Fig.1g được thảo luận để thể hiện các ví dụ về việc tính toán cụ thể các khối khác nhau một cách chi tiết hơn.

Phần minh họa ở trên của Fig.1d thể hiện tín hiệu micro dưới dạng giản đồ. Tuy nhiên, biên độ thực tế của tín hiệu micro không được minh họa. Thay vào đó, các cửa sổ được minh họa, và cụ thể là các cửa sổ 151 và 152. Cửa sổ 151 xác định khối thứ nhất 1 và cửa sổ 142 nhận diện và xác định khối thứ hai 2. Do đó, tín hiệu micro được xử lý tốt hơn là với các khối chồng lấp trong đó sự chồng lấp bằng 50%. Tuy nhiên, sự chồng lấp cao hơn hoặc thấp hơn cũng có thể được sử dụng, và thậm chí không có sự chồng lấp nào sẽ là khả thi. Tuy nhiên, việc xử lý chồng lấp được thực hiện để tránh các thành phần lạ tạo khối.

Từng khối có các giá trị tạo mẫu của tín hiệu micro được chuyển đổi thành phép biểu diễn phổ. Phép biểu diễn phổ hoặc phổ cho khối với chỉ số thời gian $n=1$, tức là, đối với khối 151, được minh họa trong phép biểu diễn ở giữa trên Fig.1d, và phép biểu diễn phổ của khối thứ hai 2 tương ứng với số tham chiếu 152 được minh họa trong ảnh phía dưới trên Fig.1d. Hơn nữa, đối với các lý do làm mẫu, từng phổ được thể hiện có mười ngàn tần số, tức là chỉ số phổ k mở rộng đến giữa 1 và 10, là ví dụ.

Do đó, ô thời gian-tần số (k, n) là ô thời gian-tần số $(10, 1)$ tại 153 và, ví dụ nữa thể hiện ô thời gian-tần số khác $(5, 2)$ tại 154. Việc xử lý nữa được thực hiện bởi thiết bị để tạo ra phần mô tả trường âm thanh, ví dụ, được minh họa trên Fig.1d, được minh họa làm mẫu bằng cách sử dụng các ô thời gian-tần số này được biểu thị bởi các số tham chiếu 153 và 154.

Hơn nữa, giả sử bộ xác định hướng 102 xác định hướng âm thanh hoặc hướng đến "DOA" được biểu thị làm mẫu bởi vector đơn vị chuẩn n . Các biểu thị hướng thay thế bao gồm góc phương vị, góc ngẩng hoặc cả hai góc này. Để đạt được mục đích này, tất cả các tín hiệu micro thuộc nhiều tín hiệu micro, trong đó từng tín hiệu micro được biểu diễn bởi các khối tiếp theo của các ngăn tần số như được minh họa trên Fig.1d, được sử dụng bởi bộ xác định hướng 102, và bộ xác định hướng 102 của Fig.1c ví dụ sau đó xác định hướng âm thanh hoặc DOA. Do đó, một cách điển hình, ô thời gian-tần số $(10, 1)$ có hướng âm thanh $n(10, 1)$ và ô thời gian-tần số $(5, 2)$ có hướng âm thanh $n(5, 2)$ như được minh họa trong phần phía trên của Fig.1e. Trong trường hợp không gian ba chiều, hướng âm thanh là vector ba chiều có thành phần x , y hoặc z . Đương nhiên là, các hệ tọa độ khác như các tọa độ hình cầu cũng có thể được sử dụng mà dựa trên hai góc và bán kính. Ngoài ra, các góc có thể là, ví dụ, góc phương vị và góc ngẩng. Sau đó không cần bán kính. Tương tự, có hai thành phần hướng âm thanh trong trường hợp không gian hai chiều như các tọa độ Đề-các, tức là hướng x và y nhưng, một cách thay thế, tọa độ hình tròn có bán kính và góc hoặc góc ngẩng và góc nâng cũng có thể được sử dụng.

Tiến trình này không chỉ được thực hiện cho các ô thời gian-tần số $(10, 1)$ và $(5, 2)$, mà còn cho tất cả các ô thời gian-tần số, bởi các tín hiệu micro được biểu diễn.

Sau đó, một hoặc nhiều hàm cơ sở trong không gian cần được xác định. Cụ thể, được xác định là số thành phần trường âm thanh hoặc, thông thường, phép biểu diễn của các thành phần trường âm thanh sẽ được tạo ra. Số hàm cơ sở trong không gian mà bây giờ được sử dụng bởi bộ đánh giá hàm cơ sở trong không gian 103 thuộc Fig.1c cuối cùng xác định số thành phần trường âm thanh cho từng ô thời gian-tần số trong phép biểu diễn phổ hoặc số lượng thành phần trường âm thanh trong miền thời gian.

Đối với phương án nữa, giả sử số lượng bốn thành phần trường âm thanh được xác định trong đó, một cách điển hình, bốn thành phần trường âm thanh này có thể là thành phần trường âm thanh đẳng hướng (tương ứng với bậc bằng 0) và các thành phần trường âm thanh ba chiều mà định hướng theo các hướng tọa độ tương ứng của hệ tọa độ Đề-các.

Phần minh họa phía dưới trên Fig.1e minh họa các hàm cơ sở trong không gian được đánh giá G_i cho các ô thời gian-tần số khác nhau. Do đó, sẽ trở nên rõ ràng là, trong ví dụ này bốn hàm cơ sở trong không gian được đánh giá cho từng ô thời gian-tần số được xác định. Khi được giả định minh họa là từng khối có mười ngàn tần số, sau đó số lượng của 40 hàm cơ sở trong không gian được đánh giá G_i được xác định cho từng khối như khối $n=1$ hoặc cho khối $n=2$ như được minh họa trên Fig.1e. Do đó, khi chỉ hai khối được xem xét và từng khối có mười ngàn tần số, sau đó các kết quả của tiến trình trong 80 hàm cơ sở trong không gian được đánh giá, vì có hai mươi ô thời gian-tần số trong hai khối và từng ô thời gian-tần số có bốn hàm cơ sở trong không gian được đánh giá.

Fig.1f minh họa các cách hoạt động được ưu tiên của bộ tính toán thành phần trường âm thanh 201 của Fig.1c. Fig.1f minh họa trong hai phần minh họa ở trên hai khối có các ngàn tần số cho đầu vào tín hiệu tham chiếu được xác định vào trong khối 201 trên Fig.1c thông qua đường 134. Cụ thể, tín hiệu tham chiếu có thể là tín hiệu micrô cụ thể hoặc kết hợp các tín hiệu micrô cụ thể đã được xử lý với cùng cách thức như đã được thảo luận đối với Fig.1d. Do đó, một cách làm mẫu, tín hiệu tham chiếu được biểu diễn bởi phổ tham chiếu cho khối $n=1$ và phổ tín hiệu tham chiếu cho khối $n=2$. Do đó, tín hiệu tham chiếu được phân tách thành mô hình thời gian-tần số giống nhau như đã được sử dụng cho việc tính toán các hàm cơ sở trong không gian được đánh giá cho đầu ra của các ô thời gian-tần số thông qua đường 133 từ khối 103 đến khối 201.

Sau đó, việc tính toán thực tế các thành phần của trường âm thanh được thực hiện qua việc kết hợp theo hàm số giữa ô thời gian-tần số tương ứng cho tín hiệu tham chiếu P và hàm cơ sở trong không gian được đánh giá đã được kết hợp G, như được biểu thị tại 155. Tốt hơn là, kết hợp theo hàm số được biểu diễn bởi $f(\dots)$ là phép nhân

được minh họa tại 115 trên các hình vẽ Fig.3a và Fig.3b được thảo luận sau đây. Tuy nhiên, các kết hợp theo hàm số khác cũng có thể được sử dụng, như được thảo luận ở trên. Bằng phương thức kết hợp theo hàm số trong khối 155, một hoặc nhiều thành phần trường âm thanh B_i được tính toán cho từng ô thời gian-tần số để thu được phép biểu diễn trong miền tần số (phổ) của các thành phần trường âm thanh B_i như được minh họa tại 156 cho khối $n=1$ và tại 157 cho khối $n=2$.

Do đó, một cách làm mẫu, phép biểu diễn trong miền tần số của các thành phần trường âm thanh B_i một mặt được minh họa cho ô thời gian-tần số (10, 1) và mặt khác còn được minh họa cho ô thời gian-tần số (5, 2) cho khối thứ hai. Tuy nhiên, rõ ràng là số lượng của các thành phần trường âm thanh B_i được minh họa trên Fig.1f tại 156 và 157 là giống với số hàm cơ sở trong không gian được đánh giá được minh họa ở phần dưới cùng thuộc Fig.1e.

Khi chỉ cần các thành phần trường âm thanh của miền tần số, việc tính toán được hoàn thành với đầu ra của các khối 156 và 157. Tuy nhiên, trong các phương án khác, cần phép biểu diễn trong miền thời gian của các thành phần trường âm thanh để thu được phép biểu diễn trong miền thời gian cho thành phần trường âm thanh thứ nhất B_1 , phép biểu diễn trong miền thời gian nữa cho thành phần trường âm thanh thứ hai B_2 và v.v.

Để đạt được mục đích này, các thành phần trường âm thanh B_1 từ ngăn tần số 1 đến ngăn tần số 10 trong khối thứ nhất 156 được chèn vào trong khối chuyển tần số-thời gian 159 để thu được phép biểu diễn trong miền thời gian cho khối thứ nhất và thành phần thứ nhất.

Tương tự, để xác định và tính toán thành phần thứ nhất trong miền thời gian, tức là $b_1(t)$, các thành phần trường âm thanh phổ B_1 cho khối thứ hai chạy từ ngăn tần số 1 đến ngăn tần số 10 được chuyển đổi thành phép biểu diễn trong miền thời gian bằng phép biến đổi tần số-thời gian nữa 160.

Thực tế là các cửa sổ chồng lấp được sử dụng như được minh họa trong phần phía trên của Fig.1d, hoạt động nhiễu xuyên kênh hoặc chồng lấp bổ sung 161 được minh họa ở phía dưới cùng trên Fig.1f có thể được sử dụng để tính toán các mẫu miền

thời gian được xuất ra thuộc phép biểu diễn phổ thứ nhất b_1 (d) trong khoảng chồng lấp giữa khối 1 và khối 2 được minh họa tại 162 trên Fig.1g.

Cùng một tiến trình được thực hiện để tính toán thành phần trường âm thanh trong miền thời gian thứ hai b_2 (t) bên trong khoảng chồng lấp 163 ở giữa khối thứ nhất và khối thứ hai. Hơn nữa, để tính toán thành phần trường âm thanh thứ ba b_3 (t) trong miền thời gian và, cụ thể, để tính toán các mẫu trong khoảng chồng lấp 164, các thành phần D_3 từ khối thứ nhất và các thành phần D_3 từ khối thứ hai được chuyển đổi tương ứng thành phép biểu diễn trong miền thời gian bởi các tiến trình 159, 160 và các giá trị thu được sau đó được tạo nhiễu xuyên kênh/chồng lấp bổ sung trong khối 161.

Cuối cùng, cùng một tiến trình được thực hiện cho các thành phần thứ tư B_4 cho khối thứ nhất và B_4 cho khối thứ hai để thu được các mẫu cuối cùng của thành phần trường âm thanh thuộc phép biểu diễn trong miền thời gian thứ tư b_4 (t) trong khoảng chồng lấp 165 như được minh họa trên Fig.1g.

Cần lưu ý là không cần sự nhiễu xuyên kênh/chồng lấp bổ sung bất kỳ như được minh họa trong khối 161 khi xử lý để thu được các ô thời gian-tần số, không được thực hiện với các khối chồng lấp nhưng lại được thực hiện với các khối không chồng lấp.

Hơn nữa, trong trường hợp chồng lấp cao hơn khi nhiều hơn hai khối chồng lấp nhau, cần số lớn hơn tương ứng của các khối 159, 160 và sự nhiễu xuyên kênh/chồng lấp bổ sung của khối 161 được tính toán không chỉ với hai đầu ra mà thậm chí với ba đầu ra để cuối cùng thu được các mẫu của các phép biểu diễn trong miền thời gian được minh họa trên Fig.1g.

Hơn nữa, cần lưu ý là các mẫu cho các phép biểu diễn trong miền thời gian, ví dụ, cho khoảng chồng lấp OL_{23} được thu bằng cách áp dụng các tiến trình trong khối 159, 160 cho khối thứ hai và khối thứ ba. Một cách tương ứng, các mẫu đối với khoảng chồng lấp $OL_{0,1}$ được tính toán bằng cách thực hiện các tiến trình 159, 160 cho các thành phần trường âm thanh phổ tương ứng B_i cho số i nhất định đối với khối 0 và khối 1.

Hơn nữa, như đã được nêu vắn tắt, phép biểu diễn của các thành phần trường âm thanh có thể là phép biểu diễn trong miền tần số như được minh họa tại Fig.1f cho 156

và 157. Một cách thay thế, phép biểu diễn của các thành phần trường âm thanh có thể là phép biểu diễn trong miền thời gian như được minh họa trên Fig.1g, trong đó bốn thành phần trường âm thanh biểu diễn các tín hiệu âm thanh đơn giản có dãy mẫu được kết hợp với tốc độ lấy mẫu nhất định. Hơn nữa, phép biểu diễn trong miền tần số hoặc phép biểu diễn trong miền thời gian của các thành phần trường âm thanh có thể được mã hóa. Việc mã hóa này có thể được thực hiện một cách riêng rẽ để từng thành phần trường âm thanh được mã hóa như tín hiệu đơn, hoặc việc mã hóa có thể được thực hiện cùng nhau để, ví dụ, bốn thành phần trường âm thanh B_1 đến B_4 được xem như tín hiệu đa kênh có bốn kênh. Do đó, phép biểu diễn được mã hóa miền tần số hoặc phép biểu diễn trong miền thời gian được mã hóa với thuật toán mã hóa hữu dụng bất kỳ cũng là phép biểu diễn của các thành phần trường âm thanh.

Hơn nữa, thậm chí phép biểu diễn trong miền thời gian trước khi việc nhiều xuyên kênh/chồng lấp bổ sung được thực hiện bởi khối 161 có thể là phép biểu diễn hữu dụng cho các thành phần trường âm thanh đối với việc thực hiện nhất định. Hơn nữa, dạng lượng tử hóa vector đối với các khối n cho thành phần nhất định như thành phần 1 còn có thể được thực hiện để bao gồm phép biểu diễn trong miền tần số của thành phần trường âm thanh để truyền dẫn hoặc lưu trữ các công việc xử lý khác.

Các phương án ưu tiên

Fig.2a thể hiện cách tiếp cận tiên tiến của sáng chế được đưa ra bởi khối (10), mà cho phép tổng hợp thành phần âm thanh bao quanh có bậc (mức) và kiểu mong muốn từ các tín hiệu có nhiều (hai hoặc nhiều hơn hai) micrô. Không giống như các cách tiếp cận của tình trạng kỹ thuật có liên quan, không có sự hạn chế nào được đưa ra đối với thiết lập micrô. Theo phương thức này, nhiều micrô có thể được sắp xếp theo dạng hình học tùy ý, ví dụ, như thiết lập trùng khớp, dãy tuyến tính, dãy hai chiều, hoặc dãy trong không gian ba chiều. Hơn nữa, từng micrô có thể xử lý hướng tính đồng hướng hoặc theo hướng tùy ý. Các hướng tính của các micrô khác nhau có thể khác nhau.

Để có được thành phần âm thanh bao quanh mong muốn, các tín hiệu của nhiều micrô trước tiên được biến đổi trước thành phép biểu diễn thời gian-tần số sử dụng khối (101). Với mục đích này, ví dụ có thể sử dụng giàn lọc hoặc biến đổi Fourier thời gian ngắn (short-time Fourier transform - STFT). Đầu ra của khối (101) là nhiều tín

hiệu micro trong miền thời gian-tần số. Lưu ý là việc xử lý sau đây được tiến hành một cách riêng rẽ cho các ô thời gian-tần số.

Sau khi biến đổi nhiều tín hiệu micro trong miền thời gian-tần số, ta xác định một hoặc nhiều hướng âm thanh (đối với ô thời gian-tần số) trong khối (102) từ hai hoặc nhiều hơn hai tín hiệu micro. Hướng âm thanh mô tả hướng mà âm thanh nổi bật cho ô thời gian-tần số đang đi đến dãy micro. Hướng này thường được xem như hướng đến (direction-of-arrival - DOA) của âm thanh. Ngoài DOA, cũng có thể xem xét hướng lan truyền của âm thanh, mà là hướng ngược lại của DOA, hoặc bất kỳ phép đo nào mà mô tả hướng âm thanh. Một hoặc nhiều hướng âm thanh hoặc các DOA được ước lượng trong khối (102) bằng cách sử dụng ví dụ các bộ ước lượng DOA bằng hợp trong tình trạng kỹ thuật, mà sẵn có cho hầu hết thiết lập micro bất kỳ. Các bộ ước lượng DOA làm mẫu thích hợp được liệt kê trong phương án 1. Số lượng hướng âm thanh hoặc số lượng các DOA (một hoặc nhiều), mà được tính toán trong khối (102), ví dụ phụ thuộc vào sự phức tạp tính toán cho phép nhưng cũng phụ thuộc vào các khả năng sử dụng của bộ ước lượng DOA được sử dụng hoặc dạng hình học micro. Hướng âm thanh có thể được ước lượng ví dụ trong không gian 2D (ví dụ được biểu diễn ở dạng góc phương vị) hoặc trong không gian 3D (ví dụ được biểu diễn ở dạng góc phương vị và góc ngẩng). Tiếp theo, hầu như các phần mô tả dựa trên trường hợp 3D thông thường hơn, ngay cả khi áp dụng thẳng tất cả các bước xử lý cho trường hợp 2D. Trong nhiều trường hợp, người dùng chỉ rõ bao nhiêu hướng âm thanh hoặc bao nhiêu DOA (ví dụ 1, 2 hoặc 3) được ước lượng trên từng ô thời gian-tần số. Một cách thay thế, số lượng âm thanh lan truyền có thể được ước lượng bằng cách sử dụng các cách tiếp cận của tình trạng kỹ thuật, ví dụ các cách tiếp cận được giải thích trong mục tham khảo [SourceNum].

Một hoặc nhiều hướng âm thanh, mà được ước lượng trong khối (102) cho ô thời gian-tần số, được sử dụng trong khối (103) để tính toán cho các đáp ứng của một hoặc nhiều ô thời gian-tần số của hàm cơ sở trong không gian có bậc (mức) và kiểu mong muốn. Một đáp ứng được tính toán cho từng hướng âm thanh được ước lượng. Như được giải thích trong phần trên, hàm cơ sở trong không gian có thể biểu diễn ví dụ sóng hài hình cầu (ví dụ nếu việc xử lý được thực hiện trong không gian 3D) hoặc

sóng hài hình trụ (ví dụ nếu việc xử lý được thực hiện trong không gian 2D). Đáp ứng của hàm cơ sở trong không gian là hàm cơ sở trong không gian được đánh giá tại hướng âm thanh được ước lượng tương ứng, như được giải thích chi tiết hơn trong phương án thứ nhất.

Một hoặc nhiều hướng âm thanh, mà được ước lượng cho ô thời gian-tần số, còn được sử dụng trong khối (201), cụ thể để tính toán cho ô thời gian-tần số một hoặc nhiều thành phần âm thanh bao quanh có bậc (mức) hoặc kiểu mong muốn. Thành phần âm thanh bao quanh tổng hợp thành phần âm thanh bao quanh cho âm thanh có hướng đến từ hướng âm thanh được ước lượng. Đầu vào bổ sung cho khối (201) là một hoặc nhiều đáp ứng của hàm cơ sở trong không gian mà được tính toán cho ô thời gian-tần số trong khối (103), cũng như một hoặc nhiều tín hiệu micro cho ô thời gian-tần số đã định. Trong khối (201) một thành phần âm thanh bao quanh có bậc (mức) và kiểu mong muốn được tính toán cho từng hướng âm thanh được ước lượng và đáp ứng tương ứng của hàm cơ sở trong không gian. Các bước xử lý của khối (201) được thảo luận thêm trong các phương án dưới đây.

Sáng chế (10) chứa khối tùy ý (301) mà có thể tính toán cho ô thời gian-tần số thành phần âm thanh bao quanh của âm thanh khuếch tán có bậc (mức) và kiểu mong muốn. Thành phần này tổng hợp thành phần âm thanh bao quanh ví dụ cho trường âm thanh khuếch tán hoàn toàn hoặc cho âm thanh của môi trường xung quanh. Đầu vào đến khối (301) là một hoặc nhiều hướng âm thanh, mà được ước lượng trong khối (102), cũng như một hoặc nhiều tín hiệu micro. Các bước xử lý khối (301) được thảo luận thêm trong các phương án sau.

Các thành phần âm thanh bao quanh của âm thanh khuếch tán, mà được tính toán trong khối tùy ý (301), có thể còn được giải tương quan trong khối tùy ý (107). Với mục đích này, các phép giải tương quan trong tình trạng kỹ thuật có thể được sử dụng. Một số ví dụ được liệt kê trong phương án 4. Thông thường, có thể áp dụng các bộ giải tương quan khác nhau hoặc các cách thực hiện khác nhau của bộ giải tương quan có các bậc (các mức) hoặc các kiểu khác nhau. Để làm được điều này, các thành phần âm thanh bao quanh âm thanh khuếch tán được giải tương quan có các bậc (các mức) và các kiểu khác nhau sẽ không tương quan với nhau. Việc này mô phỏng cách xử lý theo

quy luật tự nhiên được kỳ vọng, cụ thể là các thành phần âm thanh bao quanh có các bậc (các mức) và các kiểu khác nhau là không tương quan với nhau đối với các âm thanh khuếch tán hoặc âm thanh của môi trường xung quanh, như được giải thích ví dụ trong mục tham khảo [SpCoherence].

Một hoặc nhiều thành phần âm thanh bao quanh (âm thanh trực tiếp) có bậc (mức) và kiểu mong muốn, mà được tính toán cho ô thời gian-tần số trong khối (201), và thành phần âm thanh bao quanh của âm thanh khuếch tán tương ứng mà được tính toán trong khối (301), được kết hợp trong khối (401). Như được thảo luận trong các phương án sau, việc kết hợp có thể được thấy rõ ví dụ như tổng số (được gán trọng số). Đầu ra của khối (401) là thành phần âm thanh bao quanh được tổng hợp cuối cùng có bậc (mức) và kiểu mong muốn cho ô thời gian-tần số đã định. Một cách rõ ràng, nếu chỉ thành phần âm thanh bao quanh (âm thanh trực tiếp) đơn lẻ có bậc (mức) và kiểu mong muốn được tính toán trong khối (201) cho ô thời gian-tần số (và không cho thành phần âm thanh bao quanh của âm thanh khuếch tán nào), sau đó bộ kết hợp (401) là không cần thiết.

Sau khi tính toán thành phần âm thanh bao quanh cuối cùng có bậc (mức) và kiểu mong muốn cho tất cả các ô thời gian-tần số, thành phần âm thanh bao quanh có thể được biến đổi lại thành miền thời gian với phép biến đổi thời gian-tần số nghịch đảo (20), mà có thể được thấy rõ ví dụ như giàn lọc nghịch đảo hoặc STFT nghịch đảo. Lưu ý là phép biến đổi thời gian-tần số nghịch đảo không bị đòi hỏi trong mọi ứng dụng, và do đó, nó không là phần nói đến của sáng chế. Trong thực tế, có thể tính toán các thành phần âm thanh bao quanh cho tất cả các bậc và các kiểu mong muốn để thu được tín hiệu âm thanh bao quanh mong muốn có bậc (mức) cực đại mong muốn.

Fig.2b thể hiện việc thực hiện được thay đổi một chút của sáng chế. Trên hình vẽ này, phép biến đổi thời gian-tần số nghịch đảo (20) được áp dụng trước bộ kết hợp (401). Có thể phép biến đổi thời gian-tần số nghịch đảo thường là phép biến đổi tuyến tính. Bằng cách áp dụng phép biến đổi thời gian-tần số nghịch đảo trước bộ kết hợp (401), ví dụ nó có thể thực hiện việc giải tương quan trong miền thời gian (thay vì miền thời gian-tần số như trên Fig.2a). Có thể có những ưu điểm thực tiễn cho một số ứng dụng khi thực hiện sáng chế.

Cần lưu ý là giàn lọc nghịch đảo còn có thể ở đâu đó khác. Thông thường, bộ kết hợp và bộ giải tương quan sẽ (và sau đó là thường) được áp dụng trong miền thời gian. Nhưng cả hai hoặc chỉ một khối cũng có thể được áp dụng trong miền tần số.

Do đó, các phương án được ưu tiên bao gồm bộ tính toán thành phần khuếch tán 301 để tính toán, cho từng ô thời gian-tần số thuộc nhiều ô thời gian-tần số, một hoặc nhiều thành phần âm thanh khuếch tán. Hơn nữa, các phương án này bao gồm bộ kết hợp 401 để kết hợp thông tin âm thanh khuếch tán và thông tin trường âm thanh trực tiếp để thu được phép biểu diễn trong miền tần số hoặc phép biểu diễn trong miền thời gian của các thành phần trường âm thanh. Hơn nữa, phụ thuộc vào cách thực hiện, bộ tính toán thành phần khuếch tán còn bao gồm bộ giải tương quan 107 để giải tương quan thông tin âm thanh khuếch tán, trong đó bộ giải tương quan có thể được thực hiện bên trong miền tần số để việc hiệu chỉnh được thực hiện với phép biểu diễn ô thời gian-tần số của thành phần âm thanh khuếch tán. Một cách thay thế, bộ giải tương quan được tạo cấu hình trong miền thời gian như được minh họa trên Fig.2b để việc giải tương quan trong miền thời gian của phép biểu diễn thời gian của thành phần âm thanh khuếch tán nhất định có bậc nhất định được thực hiện.

Các phương án nữa liên quan đến sáng chế bao gồm bộ chuyển đổi thời gian-tần số như bộ chuyển đổi thời gian-tần số 101 để chuyển đổi từng tín hiệu trong số nhiều tín hiệu micro của miền thời gian thành phép biểu diễn tần số có nhiều ô thời gian-tần số. Các phương án nữa bao gồm các bộ chuyển đổi tần số-thời gian như khối 20 của Fig.2a hoặc Fig.2b để chuyển đổi một hoặc nhiều thành phần trường âm thanh hoặc sự kết hợp của một hoặc nhiều thành phần trường âm thanh, tức là các thành phần trường âm thanh trực tiếp và các thành phần âm thanh khuếch tán thành phép biểu diễn trong miền thời gian của thành phần trường âm thanh.

Cụ thể, bộ chuyển đổi tần số-thời gian 20 được tạo cấu hình để xử lý một hoặc nhiều thành phần trường âm thanh để thu được nhiều thành phần trường âm thanh miền thời gian trong đó các thành phần trường âm thanh trong miền thời gian này là các thành phần trường âm thanh trực tiếp. Hơn nữa, bộ chuyển đổi tần số-thời gian 20 được tạo cấu hình để xử lý các thành phần (trường) âm thanh khuếch tán để thu được nhiều thành phần (trường âm thanh) khuếch tán của miền thời gian và bộ kết hợp được

tạo cấu hình để thực hiện việc kết hợp các thành phần trường âm thanh (trực tiếp) của miền thời gian với (các thành phần trường âm thanh) khuếch tán của miền thời gian trong miền thời như được minh họa, ví dụ, trên Fig.2b. Một cách thay thế, bộ kết hợp 401 được tạo cấu hình để kết hợp một hoặc nhiều thành phần trường âm thanh (trực tiếp) cho ô thời gian-tần số và các thành phần (trường) âm thanh khuếch tán cho ô thời gian-tần số tương ứng trong miền tần số, và bộ chuyển đổi tần số-thời gian 20 sau đó được tạo cấu hình để xử lý kết quả của bộ kết hợp 401 để thu được các thành phần trường âm thanh trong miền thời gian, tức là, phép biểu diễn các thành phần trường âm thanh trong miền thời gian như, ví dụ, được minh họa trên Fig.2a.

Các phương án sau đây mô tả chi tiết hơn sự thực hiện khác nhau của sáng chế. Lưu ý là các phương án từ 1 đến 7 xem xét một hướng âm thanh trên từng ô thời gian-tần số (và do đó, duy nhất một đáp ứng của hàm cơ sở trong không gian và duy nhất một thành phần âm thanh bao quanh của âm thanh trực tiếp trên từng mức và kiểu và thời gian và tần số). Phương án 8 mô tả ví dụ mà nhiều hơn một hướng âm thanh được xem xét trên từng ô thời gian-tần số. Khía cạnh của phương án này có thể được áp dụng theo cách thức đơn giản cho tất cả các phương án khác.

Phương án 1

Fig.3a thể hiện phương án của sáng chế mà cho phép tổng hợp thành phần âm thanh bao quanh có bậc (mức) l và kiểu m mong muốn từ các tín hiệu của nhiều (hai hoặc nhiều hơn hai) micrô.

Đầu vào là các tín hiệu của nhiều (hai hoặc nhiều hơn hai) micrô. Các micrô có thể được sắp xếp ở dạng hình học tùy ý, ví dụ, như thiết lập trùng nhau, dãy tịnh tiến, dãy dạng mặt phẳng hai chiều, dãy dạng không gian ba chiều. Hơn nữa, từng micrô có thể xử lý hướng tính đồng hướng hoặc theo hướng tùy ý. Các hướng tính của các micrô khác nhau có thể khác nhau.

Nhiều tín hiệu micrô được biến đổi thành miền thời gian-tần số trong khối (101) ví dụ bằng cách sử dụng giàn lọc hoặc phép biến đổi Fourier thời gian ngắn (STFT). Đầu ra của phép biến đổi thời gian-tần số (101) là nhiều tín hiệu micrô trong miền thời gian-tần số, mà được biểu thị bởi $P_{1-M}(k, n)$, trong đó k là chỉ số tần số, n là chỉ số thời

gian, và M là số micrô. Lưu ý là quy trình sau được thực hiện một cách riêng rẽ cho các ô thời gian-tần số (k, n) .

Sau khi biến đổi các tín hiệu micrô thành miền thời gian-tần số, việc ước lượng hướng âm thanh được thực hiện trong khối (102) trên từng thời gian và từng tần số sử dụng hai hoặc nhiều hơn hai tín hiệu micrô $P_{1..M}(k, n)$. Trong phương án này, hướng âm thanh đơn lẻ được xác định trên từng thời gian và từng tần số. Đối với việc ước lượng hướng âm thanh trong bộ ước lượng hướng đến (DOA) bằng hợp thuộc tình trạng kỹ thuật (102) có thể được sử dụng, mà có sẵn trong tài liệu tham khảo cho các dạng hình học dãy micrô khác nhau. Ví dụ, thuật toán MUSIC trong mục tham khảo [MUSIC] có thể được sử dụng mà có khả năng áp dụng cho các thiết lập micrô tùy ý. Trong trường hợp các dãy tịnh tiến đều, các dãy tịnh tiến không đều với các điểm lưới cách đều, hoặc các dãy tròn của các micrô đẳng hướng, thuật toán Root MUSIC trong mục tham khảo [RootMUSIC1, RootMUSIC2, RootMUSIC3] có thể được áp dụng mà có hiệu quả hơn về mặt tính toán so với thuật toán MUSIC. Bộ ước lượng DOA bằng hợp đã được biết đến khác, mà có thể được áp dụng cho các dãy tịnh tiến hoặc các dãy trong mặt phẳng hai chiều với cấu trúc dãy phụ bất biến xoay tròn là ESPRIT trong mục tham khảo [ESPRIT].

Trong phương án này, đầu ra của bộ ước lượng hướng âm thanh (102) là hướng âm thanh cho thời điểm n và chỉ số tần số k . Hướng âm thanh có thể được biểu diễn ví dụ trong các giới hạn của vectơ đơn vị định mức $\mathbf{n}(k, n)$ hoặc trong giới hạn của góc phương vị $\varphi(k, n)$ và/hoặc góc ngẩng $\vartheta(k, n)$, mà có liên quan ví dụ.

$$\mathbf{n}(k, n) = \begin{bmatrix} \cos \varphi(k, n) \cos \vartheta(k, n) \\ \sin \varphi(k, n) \cos \vartheta(k, n) \\ \sin \vartheta(k, n) \end{bmatrix}.$$

Nếu không góc ngẩng $\vartheta(k, n)$ nào được ước lượng (trường hợp 2D), ta có thể giả định độ ngẩng bằng không, tức là, $\vartheta(k, n) = 0$, trong các bước tiếp theo. Trong trường hợp này, vectơ đơn vị định mức $\mathbf{n}(k, n)$ có thể được viết như sau

$$\mathbf{n}(k, n) = \begin{bmatrix} \cos \varphi(k, n) \\ \sin \varphi(k, n) \end{bmatrix}.$$

Sau khi ước lượng hướng âm thanh trong khối (102), đáp ứng của hàm cơ sở trong không gian có bậc (mức) l và kiểu m mong muốn được xác định trong khối (103) một cách riêng rẽ trên từng thời gian và tần số sử dụng thông tin hướng âm thanh được ước lượng. Đáp ứng của hàm cơ sở trong không gian có bậc (mức) l và kiểu m được biểu thị bởi $G_l^m(k, n)$ và được tính toán như sau

$$G_l^m(k, n) = Y_l^m(\varphi, \vartheta).$$

Ở đây, $Y_l^m(\varphi, \vartheta)$ là hàm cơ sở trong không gian có bậc (mức) l và kiểu m mà phụ thuộc vào hướng được nhận dạng bởi vector $\mathbf{n}(k, n)$ hoặc góc phương vị $\varphi(k, n)$ và/hoặc góc ngẩng $\vartheta(k, n)$. Do đó, đáp ứng $G_l^m(k, n)$ mô tả đáp ứng của hàm cơ sở trong không gian $Y_l^m(\varphi, \vartheta)$ cho âm thanh đến từ hướng được biểu thị bởi vector $\mathbf{n}(k, n)$ hoặc góc phương vị $\varphi(k, n)$ và/hoặc góc ngẩng $\vartheta(k, n)$. Ví dụ, khi xem xét sóng hài hình cầu được định giá trị thực với chuẩn hóa N3D như hàm cơ sở trong không gian, $Y_l^m(\varphi, \vartheta)$ có thể được tính toán như trong mục tham khảo [SphHarm, Ambix, FourierAcoust]

$$Y_l^m(\varphi, \vartheta) = \begin{cases} \sqrt{2} K_l^m \cos(m\varphi) L_l^m(\cos\vartheta) & \text{if } m > 0 \\ K_l^m L_l^m(\cos\vartheta) & \text{if } m = 0 \\ \sqrt{2} K_l^m \sin(-m\varphi) L_l^{-m}(\cos\vartheta) & \text{if } m < 0 \end{cases}$$

trong đó

$$K_l^m = \sqrt{\frac{(2l+1)(l-|m|)!}{4\pi(l+|m|)!}}$$

là các hằng số chuẩn hóa N3D và $L_l^m(\cos\vartheta)$ là đa thức Legendre được kết hợp có bậc (mức) l và kiểu m phụ thuộc vào góc ngẩng, mà được xác định ví dụ trong mục tham khảo [FourierAcoust]. Lưu ý là đáp ứng của hàm cơ sở trong không gian $Y_l^m(k, n)$ có bậc (mức) l và kiểu m mong muốn còn có thể được tính toán trước cho từng góc ngẩng và/hoặc góc phương vị và được lưu trữ trong bảng tìm kiếm và sau đó được lựa chọn phụ thuộc vào hướng âm thanh được ước lượng.

Trong phương án này, không có sự suy giảm nào về tính tổng quát, tín hiệu micrô thứ nhất được tham chiếu đến như tín hiệu micrô tham chiếu $P_{\text{ref}}(k, n)$, tức là

$$P_{\text{ref}}(k, n) = P_1(k, n).$$

Trong phương án này, tín hiệu micrô tham chiếu $P_{\text{ref}}(k, n)$ được kết hợp như được nhân lên 115 cho ô thời gian-tần số (k, n) với đáp ứng $G_l^m(k, n)$ của hàm cơ sở trong không gian được xác định trong khối (103), tức là

$$B_l^m(k, n) = P_{\text{ref}}(k, n)G_l^m(k, n),$$

dẫn đến thành phần âm thanh bao quanh mong muốn $B_l^m(k, n)$ có bậc (mức) l và kiểu m cho ô thời gian-tần số (k, n) . Các thành phần âm thanh bao quanh thu được $B_l^m(k, n)$ suy cho cùng có thể được biến đổi ngược trở lại thành miền thời gian sử dụng giãn lọc nghịch đảo hoặc STFT nghịch đảo, được truyền, hoặc được sử dụng ví dụ cho các ứng dụng tái tạo âm thanh trong không gian. Trong thực tế, có thể tính toán các thành phần âm thanh bao quanh cho tất cả các bậc và các kiểu mong muốn để thu được tín hiệu âm thanh bao quanh mong muốn có bậc (mức) cực đại mong muốn.

Phương án 2

Fig.3b thể hiện phương án khác của sáng chế mà cho phép tổng hợp thành phần âm thanh bao quanh có bậc (mức) l và kiểu m mong muốn từ các tín hiệu của nhiều (hai hoặc nhiều hơn hai) micrô. Phương án tương tự với phương án 1 nhưng chứa thêm khối (104) để xác định tín hiệu micrô tham chiếu từ nhiều tín hiệu micrô.

Như trong phương án 1, đầu vào là các tín hiệu của nhiều (hai hoặc nhiều hơn hai) micrô. Các micrô có thể được sắp xếp ở dạng hình học tùy ý, ví dụ, như thiết lập trùng nhau, dãy tịnh tiến, dãy dạng mặt phẳng hai chiều, dãy dạng không gian ba chiều. Hơn nữa, từng micrô có thể xử lý hướng tính đồng hướng hoặc theo hướng tùy ý. Các hướng tính của các micrô khác nhau có thể khác nhau.

Như trong phương án 1, nhiều tín hiệu micrô được biến đổi thành miền thời gian-tần số trong khối (101) sử dụng ví dụ giãn lọc hoặc phép biến đổi Fourier thời gian ngắn (STFT). Đầu ra của phép biến đổi thời gian-tần số (101) là các tín hiệu micrô

trong miền thời gian-tần số, mà được biểu thị bởi $P_{1..M}(k, n)$. Việc xử lý sau được thực hiện một cách riêng rẽ cho các ô thời gian-tần số (k, n) .

Như trong phương án 1, việc ước lượng hướng âm thanh được thực hiện trong khối (102) cho từng thời gian và tần số sử dụng hai hoặc nhiều hơn hai tín hiệu micro $P_{1..M}(k, n)$. Các bộ ước lượng tương ứng được thảo luận trong phương án 1. Đầu ra của bộ ước lượng hướng âm thanh (102) là hướng âm thanh trên từng khoảng thời gian n và chỉ số tần số k . Hướng âm thanh có thể được biểu diễn ví dụ trong giới hạn vector đơn vị định mức $\mathbf{n}(k, n)$ hoặc trong giới hạn góc phương vị $\varphi(k, n)$ và/hoặc góc ngẩng $\vartheta(k, n)$, mà có liên quan như được giải thích trong phương án 1.

Như trong phương án 1, đáp ứng của hàm cơ sở trong không gian có bậc (mức) l và kiểu m mong muốn được xác định trong khối (103) trên từng thời gian và tần số sử dụng thông tin hướng âm thanh được ước lượng. Đáp ứng của hàm cơ sở trong không gian được biểu thị bởi $G_l^m(k, n)$. Ví dụ, ta có thể xem xét các sóng hài hình cầu được định giá trị thực với chuẩn hóa N3D như hàm cơ sở trong không gian $G_l^m(k, n)$ có thể được xác định như được giải thích trong phương án 1.

Trong phương án này, tín hiệu micro tham chiếu $P_{\text{ref}}(k, n)$ được xác định từ nhiều tín hiệu micro $P_{1..M}(k, n)$ trong khối (104). Với mục đích này, khối (104) sử dụng thông tin hướng âm thanh mà được ước lượng trong khối (102). Các tín hiệu micro tham chiếu khác nhau có thể được xác định cho các ô thời gian-tần số khác nhau. Những khả năng khác nhau tồn tại để xác định tín hiệu micro tham chiếu $P_{\text{ref}}(k, n)$ từ nhiều tín hiệu micro $P_{1..M}(k, n)$ dựa trên thông tin hướng âm thanh. Ví dụ, có thể lựa chọn trên thời gian và tần số micro từ nhiều micro mà gần nhất với hướng âm thanh được ước lượng. Cách tiếp cận này được hiển thị hóa trên Fig.1b. Ví dụ, giả định là các vị trí micro được đưa ra bởi các vector định vị $\mathbf{d}_{1..M}$, chỉ số $i(k, n)$ của các micro gần nhất có thể được tìm ra bằng cách giải quyết vấn đề

$$i(k, n) = \arg \min_{j \in [1..M]} \|\mathbf{d}_j - \mathbf{n}(k, n)\|$$

sao cho tín hiệu micro tham chiếu cho thời gian và tần số được xem xét được định trước bởi phương trình

$$P_{\text{ref}}(k, n) = P_{i(k, n)}(k, n).$$

Trong ví trên Fig.1b, micro tham chiếu cho ô thời gian-tần số (k, n) sẽ là số micro bằng 3, tức là $i(k, n) = 3$, như $\mathbf{d}_3$ là gần với $\mathbf{n}(k, n)$. Cách tiếp cận bổ sung để xác định tín hiệu micro tham chiếu $P_{\text{ref}}(k, n)$ là để áp dụng bộ lọc đa kênh cho các tín hiệu micro, tức là

$$P_{\text{ref}}(k, n) = \mathbf{w}^H(\mathbf{n})\mathbf{p}(k, n),$$

trong đó $\mathbf{w}(\mathbf{n})$ là bộ lọc đa kênh mà phụ thuộc vào hướng âm thanh được ước lượng và vector $\mathbf{p}(k, n) = [P_1(k, n), \dots, P_M(k, n)]^T$ chứa nhiều tín hiệu micro. Có nhiều bộ lọc đa kênh tối ưu khác nhau $\mathbf{w}(\mathbf{n})$ trong tài liệu tham khảo mà có thể được sử dụng $P_{\text{ref}}(k, n)$, ví dụ bộ lọc trễ & tổng hoặc bộ lọc LCMV, mà được suy ra ví dụ trong mục tham khảo [OptArrayPr]. Việc sử dụng các bộ lọc đa kênh cung cấp các ưu điểm và bất lợi khác nhau mà được giải thích trong mục tham khảo [OptArrayPr], ví dụ, chúng cho phép làm giảm chính nhiễu âm của micro.

Như trong phương án 1, tín hiệu micro tham chiếu $P_{\text{ref}}(k, n)$ cuối cùng được kết hợp như được nhân 115 trên khung và tần số với đáp ứng $G_i^m(k, n)$ của hàm cơ sở trong không gian được xác định trong khối (103) dẫn đến thành phần âm thanh bao quanh mong muốn $B_i^m(k, n)$ có bậc (mức) l và kiểu m cho ô thời gian-tần số (k, n) . Các thành phần âm thanh bao quanh thu được $B_i^m(k, n)$ cuối cùng có thể được biến đổi lại thành miền thời gian sử dụng giàn lọc nghịch đảo hoặc STFT nghịch đảo, được truyền, hoặc được sử dụng ví dụ cho phép tái tạo âm thanh trong không gian. Trong thực tế, có thể tính toán các thành phần âm thanh bao quanh có tất cả các bậc và các kiểu mong muốn để thu được tín hiệu âm thanh bao quanh mong muốn có bậc cực đại mong muốn.

Phương án 3

Fig.4 thể hiện phương án khác của sáng chế cho phép tổng hợp thành phần âm thanh bao quanh có bậc l và kiểu m từ các tín hiệu của nhiều (hai hoặc nhiều hơn hai) micrô. Phương án này tương tự như phương án 1 nhưng tính toán các thành phần âm thanh bao quanh cho tín hiệu âm thanh trực tiếp và tín hiệu âm thanh khuếch tán.

Như trong phương án 1, đầu vào là các tín hiệu của nhiều (hai hoặc nhiều hơn hai) micrô. Các micrô có thể được sắp xếp ở dạng hình học tùy ý, ví dụ, như thiết lập trùng nhau, dãy tịnh tiến, dãy dạng mặt phẳng hai chiều, dãy dạng không gian ba chiều. Hơn nữa, từng micrô có thể xử lý hướng tính đồng hướng hoặc theo hướng tùy ý. Các hướng tính của các micrô khác nhau có thể khác nhau.

Như trong phương án 1, nhiều tín hiệu micrô được biến đổi thành miền tần số trong khối (101) sử dụng ví dụ giàn lọc hoặc phép biến đổi Fourier thời gian ngắn (STFT). Đầu ra của phép biến đổi thời gian-tần số (101) là các tín hiệu micrô trong miền thời gian-tần số, mà được biểu thị bởi $P_{l,m}(k,n)$. Việc xử lý sau được thực hiện một cách riêng rẽ cho các ô thời gian-tần số (k,n) .

Như trong phương án 1, việc ước lượng hướng âm thanh được thực hiện trong khối (102) cho từng thời gian và tần số sử dụng hai hoặc nhiều hơn hai tín hiệu micrô $P_{l,m}(k,n)$. Các bộ ước lượng tương ứng được thảo luận trong phương án 1. Đầu ra của bộ ước lượng hướng âm thanh (102) là hướng âm thanh trên từng khoảng thời gian n và chỉ số tần số k . Hướng âm thanh có thể được biểu diễn ví dụ trong giới hạn vectơ đơn vị định mức $\mathbf{n}(k,n)$ hoặc trong giới hạn góc phương vị $\varphi(k,n)$ và/hoặc góc ngẩng $\theta(k,n)$, mà có liên quan như được giải thích trong phương án 1.

Như trong phương án 1, đáp ứng của hàm cơ sở trong không gian có bậc (mức) mong muốn l và kiểu m được xác định trong khối (103) trên từng thời gian và tần số sử dụng thông tin hướng âm thanh được ước lượng. Đáp ứng của hàm cơ sở trong không gian được biểu thị bởi $G_l^m(k,n)$. Ví dụ, ta có thể xem xét các sóng hài hình cầu được định giá trị thực với chuẩn hóa N3D như hàm cơ sở trong không gian $G_l^m(k,n)$ có thể được xác định như được giải thích trong phương án 1.

Trong phương án này, đáp ứng trung bình của hàm cơ sở trong không gian có bậc (mức) l và kiểu m mong muốn, mà độc lập với chỉ số thời gian n , thu được từ khối (106). Đáp ứng trung bình này được biểu thị bởi $D_l^m(k)$ và mô tả đáp ứng của hàm cơ sở trong không gian cho âm thanh đến từ mọi hướng có thể (như các âm thanh khuếch tán hoặc âm thanh trong môi trường xung quanh). Một ví dụ để xác định đáp ứng trung bình $D_l^m(k)$ là để xem xét tích phân của cường độ bình phương của hàm cơ sở trong không gian $Y_l^m(\varphi, \vartheta)$ đối với tất cả các góc có thể φ và/hoặc ϑ . Ví dụ, khi tích hợp tất cả các góc trên mặt cầu, chúng ta thu được

$$D_l^m(k) = \int_0^{2\pi} \int_0^\pi |Y_l^m(\varphi, \vartheta)|^2 \sin \vartheta \, d\vartheta d\varphi.$$

Như việc xác định đáp ứng trung bình $D_l^m(k)$ có thể được hiểu như sau: Như được giải thích trong phương án 1, hàm cơ sở trong không gian $Y_l^m(\varphi, \vartheta)$ có thể được hiểu như hướng tính của micrô có bậc l . Để làm tăng các bậc này, vì micrô sẽ trở thành có hướng tính ngày càng nhiều, và do đó, năng lượng âm thanh khuếch tán ít hoặc năng lượng âm thanh của môi trường xung quanh ít sẽ được thu trong trường âm thanh thực tế đến micrô đẳng hướng (micrô có bậc $l = 0$). Với việc xác định $D_l^m(k)$ đã định ở trên, đáp ứng trung bình $D_l^m(k)$ sẽ dẫn đến hệ số được định giá trị thực mà mô tả bao nhiêu năng lượng âm thanh khuếch tán hoặc năng lượng âm thanh của môi trường xung quanh được làm giảm trong tín hiệu của micrô có bậc l khi so sánh với micrô đẳng hướng. Một cách rõ ràng, bên cạnh việc lấy tích phân cường độ bình phương của hàm cơ sở trong không gian $Y_l^m(\varphi, \vartheta)$ với các hướng của mặt cầu, những thay thế khác nhau tồn tại để xác định đáp ứng trung bình $D_l^m(k)$, ví dụ, lấy tích phân cường độ bình phương của $Y_l^m(\varphi, \vartheta)$ với các hướng trên hình tròn, lấy tích phân cường độ bình phương của $Y_l^m(\varphi, \vartheta)$ với tập hợp bất kỳ của các hướng mong muốn (φ, ϑ) , lấy trung bình cường độ bình phương của $Y_l^m(\varphi, \vartheta)$ với tập hợp bất kỳ của các hướng mong muốn (φ, ϑ) , lấy tích phân hoặc lấy trung bình cường độ của $Y_l^m(\varphi, \vartheta)$ thay vì cường độ bình phương, xem xét tổng được gán trọng số $Y_l^m(\varphi, \vartheta)$ với tập hợp bất kỳ của các

hướng mong muốn (φ, ϑ) , hoặc chỉ rõ số được định giá trị thực mong muốn bất kỳ cho $D_l^m(k)$ mà tương ứng với độ nhảy mong muốn của micrô được tưởng tượng được đề cập ở trên có bậc l đối với các âm thanh khuếch tán hoặc các âm thanh trong môi trường xung quanh.

Đáp ứng của hàm cơ sở trong không gian trung bình còn có thể được tính toán trước và được lưu trữ trong bảng tìm kiếm và việc xác định các giá trị đáp ứng được thực hiện bằng cách truy cập bảng tìm kiếm và khôi phục lại giá trị tương ứng.

Như trong phương án 1, không mất tính tổng quát, tín hiệu micrô thứ nhất được tham chiếu đến như tín hiệu micrô tham chiếu $P_{\text{ref}}(k, n) = P_1(k, n)$.

Trong phương án này, tín hiệu micrô tham chiếu $P_{\text{ref}}(k, n)$ được sử dụng trong khối (105) để tính toán tín hiệu âm thanh trực tiếp được biểu thị bởi $P_{\text{dir}}(k, n)$ và tín hiệu âm thanh khuếch tán được biểu thị bởi $P_{\text{diff}}(k, n)$. Trong khối (105), tín hiệu âm thanh trực tiếp $P_{\text{dir}}(k, n)$ có thể được tính toán ví dụ bằng cách áp dụng bộ lọc kênh đơn $W_{\text{dir}}(k, n)$ cho tín hiệu micrô tham chiếu, tức là

$$P_{\text{dir}}(k, n) = W_{\text{dir}}(k, n)P_{\text{ref}}(k, n).$$

Có những khả năng khác nhau tồn tại trong tài liệu tham khảo để tính toán bộ lọc kênh đơn tối ưu $W_{\text{dir}}(k, n)$. Ví dụ, bộ lọc căn bậc hai Wiener đã biết có thể được sử dụng, mà đã được chỉ rõ ví dụ trong mục tham khảo [Victaulic] như

$$W_{\text{dir}}(k, n) = \sqrt{\frac{\text{SDR}(k, n)}{\text{SDR}(k, n) + 1}}$$

trong đó $\text{SDR}(k, n)$ là tỉ số tín hiệu đến khuếch tán (signal-to-diffuse ratio - SDR) tại khoảng thời gian n và chỉ số tần số k mà mô tả tỉ số năng lượng giữa âm thanh trực tiếp và âm thanh khuếch tán như được mô tả trong mục tham khảo [VirtualMic]. SDR có thể được ước lượng sử dụng hai micrô bất kỳ có nhiều tín hiệu micrô $P_{1..M}(k, n)$ với bộ ước lượng SDR trong tình trạng kỹ thuật có sẵn trong tài liệu tham khảo, ví dụ các bộ ước lượng được đề xuất trong mục tham khảo [SDRestim] mà dựa trên sự gắn kết không gian giữa hai tín hiệu loa phát thanh tùy ý. Trong khối (105), tín hiệu âm thanh

khuếch tán $P_{\text{diff}}(k, n)$ có thể được tính toán ví dụ bằng cách áp dụng bộ lọc đơn kênh $W_{\text{diff}}(k, n)$ cho tín hiệu micro tham chiếu, tức là

$$P_{\text{diff}}(k, n) = W_{\text{diff}}(k, n)P_{\text{ref}}(k, n).$$

Có những khả năng khác nhau tồn tại trong tài liệu tham khảo để tính toán bộ lọc kênh đơn tối ưu $W_{\text{diff}}(k, n)$. Ví dụ, bộ lọc Wiener căn bậc hai đã được biết đến có thể được sử dụng, mà được định rõ ví dụ trong mục tham khảo [VirtualMic] như

$$W_{\text{diff}}(k, n) = \sqrt{\frac{1}{\text{SDR}(k, n) + 1}}$$

trong đó $\text{SDR}(k, n)$ là SDR mà có thể được ước lượng như được thảo luận ở trên.

Trong phương án này, tín hiệu âm thanh trực tiếp $P_{\text{dir}}(k, n)$ được xác định trong khối (105) được kết hợp chẳng hạn như được nhân 115a trên từng thời gian và tần số với đáp ứng $G_l^m(k, n)$ của hàm cơ sở trong không gian được xác định trong khối (103), tức là

$$B_{\text{dir}, l}^m(k, n) = P_{\text{dir}}(k, n)G_l^m(k, n),$$

dẫn đến thành phần âm thanh trực tiếp của âm thanh bao quanh $B_{\text{dir}, l}^m(k, n)$ có bậc (mức) l và kiểu m cho ô thời gian-tần số (k, n) . Hơn nữa, tín hiệu âm thanh khuếch tán $P_{\text{diff}}(k, n)$ được xác định trong khối (105) được kết hợp như được nhân 115b trên từng thời gian và tần số với đáp ứng trung bình $D_l^m(k)$ của hàm cơ sở trong không gian được xác định trong khối (106), tức là

$$B_{\text{diff}, l}^m(k, n) = P_{\text{diff}}(k, n)D_l^m(k),$$

thu được thành phần âm thanh khuếch tán của âm thanh bao quanh $B_{\text{diff}, l}^m(k, n)$ có bậc (mức) l và dạng m cho ô thời gian-tần số (k, n) .

Cuối cùng, thành phần âm thanh bao quanh của âm thanh trực tiếp $B_{\text{dir}, l}^m(k, n)$ và thành phần âm thanh bao quanh của âm thanh khuếch tán $B_{\text{diff}, l}^m(k, n)$ được kết hợp, ví dụ, qua hoạt động lấy tổng (109), để thu được thành phần âm thanh bao quanh cuối cùng $B_l^m(k, n)$ có bậc (mức) l và kiểu m mong muốn cho ô thời gian-tần số (k, n) , tức là

$$B_l^m(k, n) = B_{\text{dir},l}^m(k, n) + B_{\text{diff},l}^m(k, n).$$

Các thành phần âm thanh bao quanh thu được $B_l^m(k, n)$ cuối cùng có thể được biến đổi trở lại thành miền thời gian sử dụng bộ lọc nghịch đảo hoặc STFT nghịch đảo, được lưu trữ, được truyền, hoặc được sử dụng ví dụ cho việc tái tạo âm thanh trong không gian. Thực tế, có thể tính toán các thành phần âm thanh bao quanh cho tất cả các bậc và các kiểu mong muốn để thu được tín hiệu âm thanh bao quanh mong muốn coa bậc (mức) cực đại mong muốn.

Điều quan trọng là phải nhấn mạnh rằng việc biến đổi lại thành miền thời gian sử dụng ví dụ giàn lọc nghịch đảo hoặc STFT nghịch đảo có thể được thực hiện trước khi tính toán $B_l^m(k, n)$, tức là trước hoạt động (109). Điều này có nghĩa là, trước tiên ta có thể biến đổi $B_{\text{dir},l}^m(k, n)$ và $B_{\text{diff},l}^m(k, n)$ lại thành miền thời gian và sau đó lấy tổng cả hai thành phần với thao tác (109) để thu được thành phần âm thanh bao quanh cuối cùng B_l^m . Việc này là có khả năng vì giàn lọc nghịch đảo hoặc STFT nghịch đảo là ở trong các phép toán tuyến tính thông thường.

Lưu ý là thuật toán trong phương án này có thể được tạo cấu hình sao cho các thành phần âm thanh bao quanh của âm thanh trực tiếp $B_{\text{dir},l}^m(k, n)$ và thành phần âm thanh bao quanh của âm thanh khuếch tán $B_{\text{diff},l}^m(k, n)$ được tính toán cho các kiểu (các bậc) khác nhau l . Ví dụ, $B_{\text{dir},l}^m(k, n)$ có thể được tính toán lên đến bậc $l = 4$ trong đó $B_{\text{diff},l}^m(k, n)$ có thể được tính toán chỉ lên đến bậc $l = 1$ (trong trường hợp này, $B_{\text{diff},l}^m(k, n)$ sẽ là không cho các bậc lớn hơn $l = 1$). Việc này có các ưu điểm cụ thể như được giải thích trong phương án 4. Nếu được mong muốn ví dụ để tính toán chỉ $B_{\text{dir},l}^m(k, n)$ mà không phải là $B_{\text{diff},l}^m(k, n)$ cho bậc (mức) l hoặc kiểu m cụ thể, sau đó ví dụ khối (105) có thể được tạo cấu hình sao cho tín hiệu âm thanh khuếch tán $P_{\text{diff}}(k, n)$ trở thành bằng không. Việc này có thể đạt được ví dụ bằng cách thiết lập bộ lọc $W_{\text{diff}}(k, n)$ trong phương trình trước khi đến 0 và bộ lọc $W_{\text{dir}}(k, n)$ đến 1. Một cách thay thế, có thể thiết lập thủ công SDR trong các phương trình có trước đến giá trị rất cao.

Phương án 4

Fig.5 thể hiện phương án khác của sáng chế mà cho phép tổng hợp thành phần âm thanh bao quanh có bậc (mức) l và kiểu m mong muốn từ các tín hiệu của nhiều (hai hoặc nhiều hơn hai) micrô. Phương án này tương tự như phương án 3 nhưng chứa thêm các bộ giải tương quan cho các thành phần âm thanh khuếch tán.

Như trong phương án 3, đầu vào là các tín hiệu của nhiều (hai hoặc nhiều hơn hai) micrô. Các micrô có thể được sắp xếp ở dạng hình học tùy ý, ví dụ, như thiết lập trùng nhau, dãy tịnh tiến, dãy dạng mặt phẳng hai chiều, dãy dạng không gian ba chiều. Hơn nữa, từng micrô có thể xử lý hướng tính đồng hướng hoặc theo hướng tùy ý. Các hướng tính của các micrô khác nhau có thể khác nhau.

Như trong phương án 3, nhiều tín hiệu micrô được biến đổi thành miền thời gian-tần số trong khối (101) sử dụng ví dụ giàn lọc hoặc phép biến đổi Fourier thời gian ngắn (STFT). Đầu ra của phép biến đổi thời gian-tần số (101) là các tín hiệu micrô trong miền thời gian-tần số, mà được biểu thị bởi $P_{l,m}(k,n)$. Việc xử lý tiếp theo được thực hiện một cách riêng rẽ cho các ô thời gian-tần số (k,n) .

Như trong phương án 3, việc ước lượng hướng âm thanh được thực hiện trong khối (102) cho từng thời gian và tần số sử dụng hai hoặc nhiều hơn hai tín hiệu micrô $P_{l,m}(k,n)$. Các bộ ước lượng tương ứng được thảo luận trong phương án 1. Đầu ra của bộ ước lượng hướng âm thanh (102) là hướng âm thanh trên từng khoảng thời gian n và chỉ số tần số k . Hướng âm thanh có thể được biểu diễn ví dụ trong giới hạn vector đơn vị định mức $\mathbf{n}(k,n)$ hoặc trong giới hạn góc phương vị $\varphi(k,n)$ và/hoặc góc ngẩng $\theta(k,n)$, mà có liên quan như được giải thích trong phương án 1.

Như trong phương án 3, đáp ứng của hàm cơ sở trong không gian có bậc (mức) l và kiểu m mong muốn được xác định trong khối (103) trên từng thời gian và tần số sử dụng thông tin hướng âm thanh được ước lượng. Đáp ứng của hàm cơ sở trong không gian được biểu thị bởi $G_l^m(k,n)$. Ví dụ, ta có thể xem xét các sóng hài hình cầu được định giá trị thực với chuẩn hóa N3D như hàm cơ sở trong không gian $G_l^m(k,n)$ có thể được xác định như được giải thích trong phương án 1.

Như trong phương án 3, đáp ứng trung bình của hàm cơ sở trong không gian có bậc (mức) l và kiểu m mong muốn, mà độc lập với chỉ số thời gian n , được thu từ khối (106). Đáp ứng trung bình này được biểu thị bởi $D_l^m(k)$ và mô tả đáp ứng của hàm cơ sở trong không gian cho âm thanh đến từ mọi hướng có thể (như âm thanh khuếch tán hoặc âm thanh trong môi trường xung quanh). Đáp ứng trung bình $D_l^m(k)$ có thể được thu như được mô tả trong phương án 3.

Như trong phương án 3, không mất tính tổng quát, tín hiệu micrô thứ nhất được tham chiếu đến như tín hiệu micrô tham chiếu $P_{\text{ref}}(k, n) = P_1(k, n)$.

Như trong phương án 3, tín hiệu micrô tham chiếu $P_{\text{ref}}(k, n)$ được sử dụng trong khối (105) để tính toán tín hiệu âm thanh trực tiếp được biểu thị bởi $P_{\text{dir}}(k, n)$ và tín hiệu âm thanh khuếch tán được biểu thị bởi $P_{\text{diff}}(k, n)$. Việc tính toán $P_{\text{dir}}(k, n)$ và $P_{\text{diff}}(k, n)$ iđược giải thích trong phương án 3.

Như trong phương án 3, tín hiệu âm thanh trực tiếp $P_{\text{dir}}(k, n)$ được xác định trong khối (105) được kết hợp như được nhân 115a trên từng thời gian và tần số với đáp ứng $G_l^m(k, n)$ của hàm cơ sở trong không gian được xác định trong khối (103) dẫn đến thành phần âm thanh bao quanh của âm thanh trực tiếp $B_{\text{dir}, l}^m(k, n)$ có bậc (mức) l và kiểu m cho ô thời gian-tần số (k, n) . Hơn nữa, tín hiệu âm thanh khuếch tán $P_{\text{diff}}(k, n)$ được xác định trong khối (105) được kết hợp như được nhân 115b trên thời gian và tần số với đáp ứng trung bình $D_l^m(k)$ của hàm cơ sở trong không gian được xác định trong khối (106) dẫn đến thành phần âm thanh bao quanh của âm thanh khuếch tán $B_{\text{diff}, l}^m(k, n)$ có bậc (mức) l và kiểu m cho ô thời gian-tần số (k, n) .

Trong phương án này, thành phần âm thanh bao quanh của âm thanh khuếch tán được tính toán $B_{\text{diff}, l}^m(k, n)$ được giải tương quan trong khối (107) sử dụng bộ giải tương quan dẫn đến thành phần âm thanh bao quanh của âm thanh khuếch tán được giải tương quan, được biểu thị bởi $\tilde{B}_{\text{diff}, l}^m(k, n)$. Đối với các kỹ thuật giải tương quan trong tình trạng kỹ thuật giải tương quan có thể được sử dụng. Những cách giải tương quan khác nhau hoặc các cách thực hiện khác nhau của bộ giải tương quan thường được áp

dụng cho thành phần âm thanh bao quanh của âm thanh khuếch tán $B_{\text{diff},l}^m(k,n)$ có bậc (mức) l và kiểu m khác nhau mà các thành phần âm thanh bao quanh của âm thanh khuếch tán được giải tương quan thu được $\tilde{B}_{\text{diff},l}^m(k,n)$ có mức và kiểu khác nhau là không tương quan với nhau. Để làm điều này, các thành phần âm thanh bao quanh của âm thanh khuếch tán $\tilde{B}_{\text{diff},l}^m(k,n)$ xử lý hành vi vật lý được kỳ vọng, cụ thể mà các thành phần âm thanh bao quanh có các bậc khác nhau và các dạng khác nhau không tương quan lẫn nhau nếu trường âm thanh là âm thanh trong môi trường xung quanh hoặc khuếch tán trong mục tham khảo [SpCoherence]. Lưu ý là thành phần âm thanh bao quanh của âm thanh khuếch tán $B_{\text{diff},l}^m(k,n)$ có thể được biến đổi lại thành miền thời gian sử dụng giàn lọc nghịch đảo hoặc STFT nghịch đảo trước khi áp dụng bộ giải tương quan (107).

Cuối cùng, thành phần âm thanh bao quanh của âm thanh trực tiếp $B_{\text{dir},l}^m(k,n)$ và thành phần âm thanh bao quanh của âm thanh khuếch tán đã giải tương quan $\tilde{B}_{\text{diff},l}^m(k,n)$ được kết hợp, ví dụ, thông qua việc lấy tổng (109), để thu được thành phần âm thanh bao quanh cuối cùng $B_l^m(k,n)$ có bậc (mức) l và kiểu m mong muốn cho ô thời gian-tần số (k,n) , tức là

$$B_l^m(k,n) = B_{\text{dir},l}^m(k,n) + \tilde{B}_{\text{diff},l}^m(k,n).$$

Các thành phần âm thanh bao quanh thu được $B_l^m(k,n)$ cuối cùng có thể được biến đổi lại thành miền thời gian sử dụng ví dụ giàn lọc nghịch đảo hoặc STFT nghịch đảo, được lưu trữ, được truyền, hoặc được sử dụng ví dụ cho việc tái tạo âm thanh trong không gian. Trong thực tế, có thể tính toán các thành phần âm thanh bao quanh có tất cả các bậc và các kiểu mong muốn để thu được tín hiệu âm thanh bao quanh mong muốn có bậc cực đại mong muốn.

Điều quan trọng phải nhấn mạnh là việc biến đổi lại thành miền thời gian sử dụng ví dụ giàn lọc nghịch đảo hoặc STFT nghịch đảo có thể được thực hiện trước khi tính toán $B_l^m(k,n)$ tức là trước thao tác (109). Có nghĩa là, trước tiên ta có thể biến đổi $B_{\text{dir},l}^m(k,n)$ và $\tilde{B}_{\text{diff},l}^m(k,n)$ lại thành miền thời gian và sau đó cộng tổng hai thành phần với thao tác (109) để thu được thành phần âm thanh bao quanh cuối cùng B_l^m . Việc này có

thể xảy ra vì giàn lọc nghịch đảo hoặc STFT nghịch đảo trong các phép toán tuyến tính thông thường. Theo cách thức tương tự, bộ giải tương quan (107) có thể được áp dụng cho thành phần âm thanh bao quanh của âm thanh khuếch tán $B_{\text{diff},l}^m$ sau khi biến đổi $B_{\text{diff},l}^m$ lại thành miền thời gian. Nó có thể có thuận lợi trong thực tiễn vì một số bộ giải tương quan hoạt động trên các tín hiệu miền thời gian.

Hơn nữa, có thể được lưu ý là khối có thể được thêm vào Fig.6, như bộ lọc nghịch đảo trước bộ giải tương quan, và giàn lọc nghịch đảo có thể được bổ sung bất kỳ đâu trong hệ thống..

Như được giải thích trong phương án 3, thuật toán trong phương án này có thể được tạo cấu hình sao cho các thành phần âm thanh bao quanh của âm thanh trực tiếp $B_{\text{dir},l}^m(k,n)$ và thành phần âm thanh bao quanh của âm thanh khuếch tán $B_{\text{diff},l}^m(k,n)$ được tính toán cho các kiểu (các bậc) l khác nhau. Ví dụ, $B_{\text{dir},l}^m(k,n)$ có thể được tính toán lên đến bậc $l = 4$ trong đó $B_{\text{diff},l}^m(k,n)$ có thể được tính toán lên đến bậc $l = 1$. Việc này có thể làm giảm sự phức tạp tính toán.

Phương án 5

Fig.6 thể hiện phương án khác của sáng chế mà cho phép tổng hợp thành phần âm thanh bao quanh có bậc (mức) l và kiểu m mong muốn từ các tín hiệu của nhiều (hai hoặc nhiều hơn hai) micrô. Phương án này tương tự như phương án 4 nhưng tín hiệu âm thanh trực tiếp và tín hiệu âm thanh khuếch tán được xác định từ nhiều tín hiệu micrô bằng cách khai thác thông tin hướng đến.

Như trong phương án 4, đầu vào là các tín hiệu của nhiều (hai hoặc nhiều hơn hai) micrô. Các micrô có thể được sắp xếp ở dạng hình học tùy ý, ví dụ, như thiết lập trùng nhau, dãy tịnh tiến, dãy dạng mặt phẳng hai chiều, dãy dạng không gian ba chiều. Hơn nữa, từng micrô có thể xử lý hướng tính đồng hướng hoặc theo hướng tùy ý. Các hướng tính của các micrô khác nhau có thể khác nhau.

Như trong phương án 4, nhiều tín hiệu micrô được biến đổi thành miền tần số trong khối (101) sử dụng ví dụ giàn lọc hoặc phép biến đổi Fourier thời gian ngắn (STFT). Đầu ra của phép biến đổi thời gian-tần số (101) là các tín hiệu micrô trong

miền thời gian-tần số, mà được biểu thị bởi $P_{1-M}(k, n)$. Việc xử lý sau được thực hiện một cách riêng rẽ cho các ô thời gian-tần số (k, n) .

Như trong phương án 4, việc ước lượng hướng âm thanh được thực hiện trong khối (102) cho từng thời gian và tần số sử dụng hai hoặc nhiều hơn hai tín hiệu micrô $P_{1-M}(k, n)$. Các bộ ước lượng tương ứng được thảo luận trong phương án 1. Đầu ra của bộ ước lượng hướng âm thanh (102) là hướng âm thanh trên từng khoảng thời gian n và chỉ số tần số k . Hướng âm thanh có thể được biểu diễn ví dụ trong giới hạn vector đơn vị định mức $\mathbf{n}(k, n)$ hoặc trong giới hạn góc phương vị $\varphi(k, n)$ và/hoặc góc ngẩng $\vartheta(k, n)$, mà có liên quan như được giải thích trong phương án 1.

Như trong phương án 4, đáp ứng của hàm cơ sở không gian có bậc (mức) l và kiểu m mong muốn được xác định trong khối (103) trên từng thời gian và tần số sử dụng thông tin hướng âm thanh được ước lượng. Đáp ứng của hàm cơ sở trong không gian được biểu thị bởi $G_l^m(k, n)$. Ví dụ, ta có thể xem xét các sóng hài hình cầu được định giá trị thực với chuẩn hóa N3D như hàm cơ sở trong không gian $G_l^m(k, n)$ có thể được xác định như được giải thích trong phương án 1.

Như trong phương án 4, đáp ứng trung bình của hàm cơ sở trong không gian có bậc (mức) l và kiểu m mong muốn, mà độc lập với chỉ số thời gian n , được thu từ khối (106). Đáp ứng trung bình này được biểu thị bởi $D_l^m(k)$ và mô tả đáp ứng của hàm cơ sở trong không gian cho âm thanh đến từ mọi hướng có thể (như âm thanh khuếch tán hoặc âm thanh trong môi trường xung quanh). Đáp ứng trung bình $D_l^m(k)$ có thể được thu như được mô tả trong phương án 3.

Trong phương án này, tín hiệu âm thanh trực tiếp $P_{\text{dir}}(k, n)$ và tín hiệu âm thanh khuếch tán $P_{\text{diff}}(k, n)$ được xác định trong khối (110) trên chỉ số thời gian n và chỉ số tần số k từ hai hoặc nhiều hơn hai tín hiệu micrô có sẵn $P_{1-M}(k, n)$. Với mục đích này, khối (110) thường khai thác thông tin hướng âm thanh mà được xác định trong khối (102). Sau đây, các ví dụ khác nhau của khối (110) được giải thích mà mô tả xác định thế nào $P_{\text{dir}}(k, n)$ và $P_{\text{diff}}(k, n)$.

Trong ví dụ thứ nhất của khối (110), tín hiệu micrô tham chiếu được biểu thị bởi $P_{\text{ref}}(k, n)$ được xác định từ nhiều tín hiệu micrô $P_{1:M}(k, n)$ dựa trên thông tin hướng âm thanh được cung cấp bởi khối (102). Tín hiệu micrô tham chiếu $P_{\text{ref}}(k, n)$ có thể được xác định bằng cách lựa chọn tín hiệu micrô mà gần nhất với hướng âm thanh được ước lượng cho từng thời gian và tần số được xem xét. Quy trình lựa chọn này để xác định tín hiệu micrô tham chiếu $P_{\text{ref}}(k, n)$ được giải thích trong phương án 2. Sau khi xác định $P_{\text{ref}}(k, n)$, tín hiệu âm thanh trực tiếp $P_{\text{dir}}(k, n)$ và tín hiệu âm thanh khuếch tán $P_{\text{diff}}(k, n)$ có thể được tính toán ví dụ bằng cách áp dụng lần lượt các bộ lọc đơn kênh $W_{\text{dir}}(k, n)$ và $W_{\text{diff}}(k, n)$, cho tín hiệu micrô tham chiếu $P_{\text{ref}}(k, n)$. Cách tiếp cận và tính toán này của các bộ lọc đơn kênh tương ứng được giải thích trong phương án 3.

Trong ví dụ thứ hai của khối (110), ta xác định tín hiệu micrô tham chiếu $P_{\text{ref}}(k, n)$ như trong ví dụ trước và tính toán $P_{\text{dir}}(k, n)$ bằng cách áp dụng bộ lọc đơn kênh $W_{\text{dir}}(k, n)$ cho $P_{\text{ref}}(k, n)$. Tuy nhiên, để xác định tín hiệu khuếch tán, ta lựa chọn tín hiệu tham chiếu thứ hai $P_{\text{ref},l}^m(k, n)$ và áp dụng bộ lọc đơn kênh $W_{\text{diff}}(k, n)$ cho tín hiệu tham chiếu thứ hai $P_{\text{ref},l}^m(k, n)$, tức là

$$P_{\text{diff}}(k, n) = W_{\text{diff}}(k, n)P_{\text{ref},l}^m(k, n).$$

Bộ lọc $W_{\text{diff}}(k, n)$ có thể được tính toán như được giải thích ví dụ trong phương án 3. Tín hiệu tham chiếu thứ hai $P_{\text{ref},l}^m(k, n)$ tương ứng với một trong số tín hiệu micrô có sẵn $P_{1:M}(k, n)$. Tuy nhiên, đối với các bậc l và các kiểu m khác nhau ta có thể sử dụng các tín hiệu micrô khác nhau như tín hiệu tham chiếu thứ hai. Ví dụ, với mức $l = 1$ và kiểu $m = -1$, ta có thể sử dụng tín hiệu micrô thứ nhất như tín hiệu micrô thứ hai, tức là, $P_{\text{ref},1}^{-1}(k, n) = P_1(k, n)$. Với mức $l = 1$ và dạng $m = 0$, ta có thể sử dụng tín hiệu micrô thứ hai, tức là $P_{\text{ref},1}^0(k, n) = P_2(k, n)$. Với mức $l = 1$ và dạng $m = 1$, ta có thể sử dụng tín hiệu tham chiếu thứ ba, tức là $P_{\text{ref},1}^1(k, n) = P_3(k, n)$. Các tín hiệu micrô có sẵn $P_{1:M}(k, n)$ có thể được ấn định một cách ngẫu nhiên ví dụ cho tín hiệu tham chiếu thứ hai $P_{\text{ref},l}^m(k, n)$ cho các bậc và các dạng khác. Đây là cách tiếp cận hợp lý trong thực tiễn vì đối với các trạng thái ghi âm khuếch tán hoặc trong môi trường xung quanh, tất

cả các tín hiệu micro thường chứa năng lượng âm thanh giống nhau. Việc lựa chọn các tín hiệu micro tham chiếu thứ hai khác nhau cho các bậc khác nhau và các dạng khác nhau có ưu điểm là các tín hiệu âm thanh khuếch tán thu được và thường (ít nhất một phần) không tương quan với nhau cho các bậc và các dạng khác nhau.

Trong ví dụ thứ ba của khối (110), tín hiệu âm thanh trực tiếp $P_{\text{dir}}(k, n)$ được xác định bằng cách áp dụng bộ lọc đa kênh được biểu thị bởi $\mathbf{w}_{\text{dir}}(\mathbf{n})$ cho nhiều tín hiệu micro $P_{1:M}(k, n)$, tức là

$$P_{\text{dir}}(k, n) = \mathbf{w}_{\text{dir}}^H(\mathbf{n})\mathbf{p}(k, n),$$

trong đó bộ lọc đa kênh $\mathbf{w}_{\text{dir}}(\mathbf{n})$ phụ thuộc vào hướng âm thanh được ước lượng và vector $\mathbf{p}(k, n) = [P_1(k, n), \dots, P_M(k, n)]^T$ chứa nhiều tín hiệu micro. Có nhiều bộ lọc đa kênh tối ưu khác nhau $\mathbf{w}_{\text{dir}}(\mathbf{n})$ trong tài liệu tham khảo mà có thể được sử dụng để tính toán $P_{\text{dir}}(k, n)$ từ thông tin hướng âm thanh, ví dụ các bộ lọc được suy ra ra trong mục tham khảo [InformedSF]. Một cách tương tự, tín hiệu âm thanh khuếch tán $P_{\text{diff}}(k, n)$ được xác định bằng cách áp dụng bộ lọc đa kênh được biểu thị bởi $\mathbf{w}_{\text{diff}}(\mathbf{n})$ cho nhiều tín hiệu micro $P_{1:M}(k, n)$, tức là

$$P_{\text{diff}}(k, n) = \mathbf{w}_{\text{diff}}^H(\mathbf{n})\mathbf{p}(k, n),$$

trong đó bộ lọc đa kênh $\mathbf{w}_{\text{diff}}(\mathbf{n})$ phụ thuộc vào hướng âm thanh được ước lượng. Có nhiều bộ lọc đa kênh tối ưu khác nhau $\mathbf{w}_{\text{diff}}(\mathbf{n})$ trong tài liệu tham khảo mà có thể được sử dụng để tính toán $P_{\text{diff}}(k, n)$, ví dụ bộ lọc mà được suy ra trong mục tham khảo [DiffuseBF].

Trong ví dụ thứ tư của khối (110), ta xác định $P_{\text{dir}}(k, n)$ và $P_{\text{diff}}(k, n)$ như trong ví dụ trước bằng cách lần lượt áp dụng bộ lọc đa kênh $\mathbf{w}_{\text{dir}}(\mathbf{n})$ và $\mathbf{w}_{\text{diff}}(\mathbf{n})$, cho các tín hiệu micro $\mathbf{p}(k, n)$. Tuy nhiên, ta sử dụng các bộ lọc khác nhau $\mathbf{w}_{\text{diff}}(\mathbf{n})$ cho các bậc l và kiểu m khác nhau sao cho các tín hiệu âm thanh khuếch tán thu được $P_{\text{diff}}(k, n)$ cho các bậc l và các kiểu m khác nhau là không tương quan với nhau. Các bộ lọc khác nhau này

$w_{\text{diff}}(n)$ mà tối thiểu hóa sự tương quan giữa các tín hiệu đầu ra có thể được tính toán ví dụ như được giải thích trong mục tham khảo [CovRender].

Như trong phương án 4, tín hiệu âm thanh trực tiếp $P_{\text{dir}}(k, n)$ được xác định trong khối (105) được kết hợp như được nhân 115a trên từng thời gian và tần số với đáp ứng $G_l^m(k, n)$ của hàm cơ sở trong không gian được xác định trong khối (103) dẫn đến thành phần âm thanh bao quanh của âm thanh trực tiếp $B_{\text{dir},l}^m(k, n)$ có bậc (mức) l và kiểu m cho ô thời gian-tần số (k, n) . Hơn nữa, tín hiệu âm thanh khuếch tán $P_{\text{diff}}(k, n)$ được xác định trong khối (105) được kết hợp như được nhân 115b trên từng thời gian và tần số với đáp ứng trung bình $D_l^m(k)$ của hàm cơ sở trong không gian trong khối (106) dẫn đến thành phần âm thanh bao quanh của âm thanh khuếch tán $B_{\text{diff},l}^m(k, n)$ có bậc (mức) l và kiểu m cho ô thời gian-tần số (k, n) .

Như trong phương án 3, thành phần âm thanh bao quanh của âm thanh trực tiếp được tính toán $B_{\text{dir},l}^m(k, n)$ và thành phần âm thanh bao quanh của âm thanh khuếch tán $B_{\text{diff},l}^m(k, n)$ được kết hợp, ví dụ, thông qua thao tác lấy tổng (109), để thu được thành phần âm thanh bao quanh cuối cùng $B_l^m(k, n)$ có bậc (mức) l và kiểu m mong muốn cho từng ô thời gian-tần số (k, n) . Các thành phần âm thanh bao quanh thu được $B_l^m(k, n)$ cuối cùng có thể biến đổi trở lại thành miền thời gian sử dụng bộ lọc nghịch đảo hoặc STFT nghịch đảo, được lưu trữ, được truyền, hoặc được sử dụng cho việc tái tạo âm thanh trong không gian. Trong thực tiễn, có thể tính toán các thành phần âm thanh bao quanh cho tất cả các bậc và các kiểu mong muốn. Như được giải thích trong phương án 3, phép biến đổi lại thành miền thời gian có thể được thực hiện trước khi tính toán $B_l^m(k, n)$, tức là trước thao tác (109).

Lưu ý là thuật toán trong phương án này có thể được tạo cấu hình sao cho các thành phần âm thanh bao quanh của âm thanh trực tiếp $B_{\text{dir},l}^m(k, n)$ và thành phần âm thanh bao quanh của âm thanh khuếch tán $B_{\text{diff},l}^m(k, n)$ được tính toán cho các dạng (bậc) l khác nhau. Ví dụ, $B_{\text{dir},l}^m(k, n)$ có thể được tính toán đến bậc $l = 4$ trong đó $B_{\text{diff},l}^m(k, n)$ có thể được tính toán chỉ đến bậc $l = 1$ (trong trường hợp này $B_{\text{diff},l}^m(k, n)$ sẽ bằng không

hoặc với các bậc lớn hơn $l = 1$). Nếu được mong muốn ví dụ để tính toán chỉ $B_{\text{dir},l}^m(k,n)$ mà không phải $B_{\text{diff},l}^m(k,n)$ cho bậc (mức) l hoặc kiểu m cụ thể, sau đó ví dụ khối (110) có thể được tạo cấu hình sao cho tín hiệu âm thanh khuếch tán $P_{\text{diff}}(k,n)$ trở thành bằng không. Việc này có thể đạt được ví dụ bằng cách thiết lập bộ lọc $W_{\text{diff}}(k,n)$ trong các phương trình trước đến 0 và bộ lọc $W_{\text{dir}}(k,n)$ đến 1. Tương tự, bộ lọc $w_{\text{diff}}^H(n)$ có thể được thiết lập đến không.

Phương án 6

Fig.7 thể hiện phương án khác của sáng chế mà cho phép tổng hợp thành phần âm thanh bao quanh có bậc (mức) l và kiểu m mong muốn từ các tín hiệu của nhiều (hai hoặc nhiều hơn hai) micrô. Phương án này tương tự như phương án 5 nhưng chứa thêm các bộ giải tương quan cho các thành phần âm thanh bao quanh khuếch tán.

Như trong phương án 5, đầu vào là các tín hiệu của nhiều (hai hoặc nhiều hơn hai) micrô. Các micrô có thể được sắp xếp ở dạng hình học tùy ý, ví dụ, như thiết lập trùng nhau, dãy tịnh tiến, dãy dạng mặt phẳng hai chiều, dãy dạng không gian ba chiều. Hơn nữa, từng micrô có thể xử lý hướng tính đồng hướng hoặc theo hướng tùy ý. Các hướng tính của các micrô khác nhau có thể khác nhau.

Như trong phương án 5, nhiều tín hiệu micrô được biến đổi thành miền tần số trong khối (101) sử dụng ví dụ giàn lọc hoặc phép biến đổi Fourier thời gian ngắn (STFT). Đầu ra của phép biến đổi thời gian-tần số (101) là các tín hiệu micrô trong miền thời gian-tần số, mà được biểu thị bởi $P_{1..M}(k,n)$. Việc xử lý sau được thực hiện một cách riêng rẽ cho các ô thời gian-tần số (k,n) .

Như trong phương án 5, việc ước lượng hướng âm thanh được thực hiện trong khối (102) cho từng thời gian và tần số sử dụng hai hoặc nhiều hơn hai tín hiệu micrô $P_{1..M}(k,n)$. Các bộ ước lượng tương ứng được thảo luận trong phương án 1. Đầu ra của bộ ước lượng hướng âm thanh (102) là hướng âm thanh trên từng khoảng thời gian n và chỉ số tần số k . Hướng âm thanh có thể được biểu diễn ví dụ trong giới hạn vector

đơn vị định mức $n(k, n)$ hoặc trong giới hạn góc phương vị $\varphi(k, n)$ và/hoặc góc ngẩng $\theta(k, n)$, mà có liên quan như được giải thích trong phương án 1.

Như trong phương án 5, đáp ứng của hàm cơ sở không gian có bậc (mức) l và kiểu m mong muốn được xác định trong khối (103) trên từng thời gian và tần số sử dụng thông tin hướng âm thanh được ước lượng. Đáp ứng của hàm cơ sở trong không gian được biểu thị bởi $G_l^m(k, n)$. Ví dụ, ta có thể xem xét các sóng hài hình cầu được định giá trị thực với chuẩn hóa N3D như hàm cơ sở trong không gian $G_l^m(k, n)$ có thể được xác định như được giải thích trong phương án 1.

Như trong phương án 5, đáp ứng trung bình của hàm cơ sở trong không gian có bậc (mức) l và kiểu m mong muốn, mà độc lập với chỉ số thời gian n , được thu từ khối (106). Đáp ứng trung bình này được biểu thị bởi $D_l^m(k)$ và mô tả đáp ứng của hàm cơ sở trong không gian cho âm thanh đến từ mọi hướng có thể (như âm thanh khuếch tán hoặc âm thanh trong môi trường xung quanh). Đáp ứng trung bình $D_l^m(k)$ có thể được thu như được mô tả trong phương án 3.

Như trong phương án 5, tín hiệu âm thanh trực tiếp $P_{\text{dir}}(k, n)$ và tín hiệu âm thanh khuếch tán $P_{\text{diff}}(k, n)$ được xác định trong khối (110) trên chỉ số thời gian n và chỉ số tần số k từ hai hoặc nhiều hơn hai tín hiệu âm micro sẵn có $P_{1:M}(k, n)$. Với mục đích này, khối (110) thường khai thác thông tin hướng âm thanh mà được xác định trong khối (102). Các ví dụ khác nhau của khối (110) được giải thích trong phương án 5.

Như trong phương án 5, tín hiệu âm thanh trực tiếp $P_{\text{dir}}(k, n)$ được xác định trong khối (105) được kết hợp như được nhân 115a trên thời gian và tần số với đáp ứng $G_l^m(k, n)$ của hàm cơ sở trong không gian được xác định trong khối (103) dẫn đến thành phần âm thanh bao quanh của âm thanh trực tiếp $B_{\text{dir},l}^m(k, n)$ có bậc (mức) l và dạng m cho ô thời gian-tần số (k, n) . Hơn nữa, tín hiệu âm thanh khuếch tán $P_{\text{diff}}(k, n)$ được xác định trong khối (105) được kết hợp như được nhân 115b trên thời gian và tần số với đáp ứng trung bình $D_l^m(k)$ của hàm cơ sở trong không gian trong khối (106) dẫn

đến thành phần âm thanh bao quanh của âm thanh khuếch tán $B_{\text{diff},l}^m(k,n)$ có bậc (mức) l và kiểu m cho ô thời gian-tần số (k,n) .

Như trong phương án 4, thành phần âm thanh bao quanh của âm thanh khuếch tán được tính toán $B_{\text{diff},l}^m(k,n)$ được giải tương quan trong khối (107) sử dụng bộ giải tương quan dẫn đến thành phần âm thanh bao quanh của âm thanh khuếch tán được giải tương quan, được biểu thị bởi $\tilde{B}_{\text{diff},l}^m(k,n)$. Lập luận và các phương pháp đằng sau phép giải tương quan được thảo luận trong phương án 4. Như trong phương án 4, thành phần âm thanh bao quanh của âm thanh khuếch tán $B_{\text{diff},l}^m(k,n)$ có thể được biến đổi lại thành miền thời gian sử dụng giàn lọc nghịch đảo hoặc STFT nghịch đảo trước khi sử dụng bộ giải tương quan (107).

Như trong phương án 4, thành phần âm thanh bao quanh của âm thanh trực tiếp $B_{\text{dir},l}^m(k,n)$ và thành phần âm thanh bao quanh của âm thanh khuếch tán được giải tương quan $\tilde{B}_{\text{diff},l}^m(k,n)$ được kết hợp, ví dụ, thao tác tính tổng (109), để thu được thành phần âm thanh bao quanh cuối cùng $B_l^m(k,n)$ có bậc (mức) l và kiểu m mong muốn cho ô thời gian-tần số (k,n) . Các thành phần âm thanh bao quanh thu được $B_l^m(k,n)$ cuối cùng có thể biến đổi trở lại thành miền thời gian sử dụng bộ lọc nghịch đảo hoặc STFT nghịch đảo, được lưu trữ, được truyền, hoặc được sử dụng cho việc tái tạo âm thanh trong không gian. Trong thực tế, có thể tính toán các thành phần âm thanh bao quanh cho tất cả các bậc và các kiểu mong muốn để thu được tín hiệu âm thanh bao quanh có bậc (mức) cực đại được mong muốn. Như được giải trong phương án 4, phép biến đổi lại thành miền thời gian có thể được thực hiện trước khi tính toán $B_l^m(k,n)$, tức là trước thao tác (109).

Như trong phương án 4, thuật toán trong phương án này có thể được tạo cấu hình sao cho các thành phần âm thanh bao quanh của âm thanh trực tiếp $B_{\text{dir},l}^m(k,n)$ và thành phần âm thanh bao quanh của âm thanh khuếch tán $B_{\text{diff},l}^m(k,n)$ được tính toán cho các dạng (các bậc) l khác nhau. Ví dụ, $B_{\text{dir},l}^m(k,n)$ có thể được tính toán lên đến bậc $l = 4$ trong đó $B_{\text{diff},l}^m(k,n)$ có thể được tính toán lên đến bậc $l = 1$.

Phương án 7

Fig.8 thể hiện các phương án khác của sáng chế mà cho phép tổng hợp thành phần âm thanh bao quanh có bậc (mức) l và kiểu m mong muốn từ các tín hiệu của nhiều (hai hoặc nhiều hơn hai) micrô. Phương án này tương tự như phương án 1 nhưng có thêm khối (111) mà áp dụng hoạt động làm nhẵn cho đáp ứng được tính toán $G_l^m(k, n)$ của hàm cơ sở trong không gian.

Như trong phương án 1, đầu vào là các tín hiệu của nhiều (hai hoặc nhiều hơn hai) micrô. Các micrô có thể được sắp xếp ở dạng hình học tùy ý, ví dụ, như thiết lập trùng nhau, dãy tịnh tiến, dãy dạng mặt phẳng hai chiều, dãy dạng không gian ba chiều. Hơn nữa, từng micrô có thể xử lý hướng tính đồng hướng hoặc theo hướng tùy ý. Các hướng tính của các micrô khác nhau có thể khác nhau.

Như trong phương án 1, nhiều tín hiệu micrô được biến đổi thành miền tần số trong khối (101) sử dụng ví dụ giàn lọc hoặc phép biến đổi Fourier thời gian ngắn (STFT). Đầu ra của phép biến đổi thời gian-tần số (101) là các tín hiệu micrô trong miền thời gian-tần số, mà được biểu thị bởi $P_{1-M}(k, n)$. Việc xử lý sau được thực hiện một cách riêng rẽ cho các ô thời gian-tần số (k, n) .

Như trong phương án 1, không mất tính tổng quát, tín hiệu micrô thứ nhất được tham chiếu đến như tín hiệu micrô tham chiếu $P_{ref}(k, n) = P_1(k, n)$.

Như trong phương án 1, việc ước lượng hướng âm thanh được thực hiện trong khối (102) cho từng thời gian và tần số sử dụng hai hoặc nhiều hơn hai tín hiệu micrô $P_{1-M}(k, n)$. Các bộ ước lượng tương ứng được thảo luận trong phương án 1. Đầu ra của bộ ước lượng hướng âm thanh (102) là hướng âm thanh trên từng khoảng thời gian n và chỉ số tần số k . Hướng âm thanh có thể được biểu diễn ví dụ trong giới hạn vector đơn vị định mức $\mathbf{n}(k, n)$ hoặc trong giới hạn góc phương vị $\varphi(k, n)$ và/hoặc góc ngẩng $\vartheta(k, n)$, mà có liên quan như được giải thích trong phương án 1.

Như trong phương án 1, đáp ứng của hàm cơ sở không gian có bậc (mức) l và kiểu m mong muốn được xác định trong khối (103) trên từng thời gian và tần số sử

dụng thông tin hướng âm thanh được ước lượng. Đáp ứng của hàm cơ sở trong không gian được biểu thị bởi $G_l^m(k, n)$. Ví dụ, ta có thể xem xét các sóng hài hình cầu được định giá trị thực với chuẩn hóa N3D như hàm cơ sở trong không gian $G_l^m(k, n)$ có thể được xác định như được giải thích trong phương án 1.

Ngược lại với phương án 1, đáp ứng $G_l^m(k, n)$ được sử dụng như đầu vào đến khối (111) mà áp dụng hoạt động làm nhẵn cho $G_l^m(k, n)$. Đầu ra của khối (111) là hàm đáp ứng được làm nhẵn được biểu thị như $\bar{G}_l^m(k, n)$. Mục đích của hoạt động làm nhẵn là để làm giảm sai số ước lượng không mong muốn của các giá trị của $G_l^m(k, n)$, mà có thể xảy ra trong thực tiễn ví dụ nếu các hướng âm thanh $\varphi(k, n)$ và/hoặc $\theta(k, n)$ được ước lượng trong khối (102) là nhiễu âm. Việc làm nhẵn được áp dụng cho $G_l^m(k, n)$ ví dụ có thể được thực hiện theo thời gian và/hoặc tần số. Ví dụ, việc làm nhẵn theo thời gian có thể đạt được bằng cách sử dụng bộ lọc lấy trung bình đệ quy đã được biết đến

$$\bar{G}_l^m(k, n) = \alpha G_l^m(k, n) + (1 - \alpha) \bar{G}_l^m(k, n - 1),$$

trong đó $G_l^m(k, n - 1)$ là hàm đáp ứng được tính toán trong khung thời gian có trước. Hơn nữa, α là số được định giá trị thực giữa 0 và 1 mà kiểm soát chiều dài của việc làm nhẵn theo thời gian. Với các giá trị bằng α gần bằng 0, việc lấy trung bình theo thời gian mạnh được thực hiện cho các giá trị bằng α gần bằng 1, việc lấy trung bình theo thời gian ngắn được thực hiện. Trong các giá trị thực tiễn, giá trị bằng α phụ thuộc vào ứng dụng và có thể được thiết lập không đổi, ví dụ $\alpha = 0.5$. Một cách thay thế, việc làm nhẵn phổ cũng có thể được thể hiện trong khối (111), mà có nghĩa là đáp ứng $G_l^m(k, n)$ được lấy trung bình theo nhiều băng tần số. Việc làm nhẵn phổ này, ví dụ bên trong cái gọi là các băng ERB, được mô tả ví dụ trong mục tham khảo [ERBsmooth].

Trong phương án này, tín hiệu micrô tham chiếu $P_{\text{ref}}(k, n)$ cuối cùng được kết hợp như được nhân 115 trên thời gian và tần số với đáp ứng được làm nhẵn $\bar{G}_l^m(k, n)$ có hàm cơ sở trong không gian được xác định trong khối (111) dẫn đến thành phần âm thanh bao quanh mong muốn $B_l^m(k, n)$ có bậc (mức) l và kiểu m cho ô thời gian-tần số

(k, n) . Các thành phần âm thanh bao quanh thu được $B_l^m(k, n)$ cuối cùng có thể biến đổi trở lại thành miền thời gian sử dụng bộ lọc nghịch đảo hoặc STFT nghịch đảo, được lưu trữ, được truyền, hoặc được sử dụng cho việc tái tạo âm thanh trong không gian. Trong thực tế, có thể tính toán các thành phần âm thanh bao cho tất cả các bậc và các kiểu mong muốn để thu được tín hiệu âm thanh bao quanh mong muốn có bậc (mức) cực đại mong muốn.

Một cách rõ ràng, việc làm nhẵn độ khuếch đại trong khối (111) có thể được áp dụng cũng trong tất cả các phương án khác của sáng chế này.

Phương án 8

Sáng chế có thể được sử dụng cũng trong trường hợp gọi là nhiều bước sóng, trong đó nhiều hơn một hướng âm thanh được xem xét trên ô thời gian-tần số. Ví dụ phương án 2 được minh họa trên Fig.3b có thể được hiểu rõ trong trường hợp nhiều bước sóng. Trong trường hợp này, khối (102) ước lượng các hướng âm thanh J trên từng thời gian và tần số, trong đó J là giá trị nguyên lớn hơn một, ví dụ, $J = 2$. Để ước lượng nhiều hướng âm thanh, các bộ ước lượng trong tình trạng kỹ thuật có thể được sử dụng, ví dụ ESPRIT hoặc Root MUSIC, mà được mô tả trong mục tham khảo [ESPRIT, RootMUSIC1]. Trong trường hợp này, đầu ra của khối (102) là nhiều hướng âm thanh, được biểu thị ví dụ trong các giới hạn của nhiều góc phương vị $\varphi_{1-J}(k, n)$ và hoặc các góc ngẩng $\vartheta_{1-J}(k, n)$.

Nhiều hướng âm thanh sau đó được sử dụng trong khối (103) để tính toán nhiều đáp ứng $G_{l,1-J}^m(k, n)$, một đáp ứng cho từng hướng âm thanh được ước lượng như được thảo luận ví dụ trong phương án 1. Hơn nữa, các hướng đa âm được tính toán trong khối (102) được sử dụng trong khối (104) để tính toán nhiều tín hiệu tham chiếu $P_{\text{ref},1-J}(k, n)$, một cho từng hướng trong số nhiều hướng âm thanh. Từng tín hiệu trong số nhiều tín hiệu tham chiếu có thể được tính toán bằng cách áp dụng bộ lọc đa kênh $\mathbf{w}_{1-J}(\mathbf{n})$ cho nhiều tín hiệu micro, một cách tương tự như được giải thích trong phương án 2. Ví dụ, tín hiệu tham chiếu thứ nhất $P_{\text{ref},1}(k, n)$ có thể được thu bằng cách áp dụng bộ lọc đa kênh thuộc tình trạng kỹ thuật $\mathbf{w}_1(\mathbf{n})$, mà có thể trích xuất các âm

thanh từ hướng $\varphi_1(k, n)$ và/hoặc $\vartheta_1(k, n)$ trong khi làm giảm các âm thanh từ tất cả các hướng âm thanh khác. Bộ lọc này có thể được tính toán ví dụ như bộ lọc LCMV được báo mà được giải thích trong mục tham khảo [InformedSF]. Nhiều tín hiệu tham chiếu $P_{\text{ref},1 \dots J}(k, n)$ sau đó được nhân với nhiều đáp ứng tương ứng $G_{i,1 \dots J}^m(k, n)$ để thu được nhiều thành phần âm thanh bao quanh $B_{i,1 \dots J}^m(k, n)$. Ví dụ thành phần âm thanh bao quanh thứ j lần lượt tương ứng với hướng âm thanh thứ j và tín hiệu tham chiếu, được tính toán như sau

$$B_{i,j}^m(k, n) = P_{\text{ref},j}(k, n)G_{i,j}^m(k, n).$$

Cuối cùng, các thành phần âm thanh bao quanh j được tính tổng để thu được các thành phần âm thanh bao quanh mong muốn $B_i^m(k, n)$ có bậc (mức) i và kiểu m cho ô thời gian-tần số (k, n) , tức là:

$$B_i^m(k, n) = \sum_{j=1}^J B_{i,j}^m(k, n).$$

Một cách rõ ràng, cũng trong các phương án nêu trên có thể mở rộng cho trường hợp nhiều bước sóng. Ví dụ, trong phương án 5 và phương án 6 ta có thể tính toán nhiều âm thanh trực tiếp $P_{\text{dir},1 \dots J}(k, n)$, một cho từng hướng trong số nhiều hướng âm thanh, sử dụng các bộ lọc đa kênh giống nhau như được đề cập trong phương án này. Nhiều âm thanh trực tiếp sau đó được nhân với nhiều đáp ứng tương ứng $G_{i,1 \dots J}^m(k, n)$ dẫn đến nhiều thành phần âm thanh bao quanh của âm thanh trực tiếp $B_{\text{dir},i,1 \dots J}^m(k, n)$ mà có thể được tính tổng để thu được thành phần âm thanh bao quanh của âm thanh trực tiếp mong muốn $B_{\text{dir},i}^m(k, n)$.

Cần lưu ý là sáng chế không chỉ được áp dụng cho các kỹ thuật âm thanh bao quanh ba chiều (hình cầu) hoặc hai chiều (hình trụ) mà còn được áp dụng cho các kỹ thuật khác bất kỳ dựa trên các hàm cơ sở trong không gian để tính toán các thành phần trường âm thanh bất kỳ.

Các phương án của sáng chế được liệt kê:

Biến đổi nhiều tín hiệu micrô thành miền thời gian tần số.

Tính toán một hoặc nhiều hướng âm thanh trên từng thời gian và tần số từ nhiều tín hiệu micro.

Tính toán cho từng thời gian và tần số một hoặc nhiều hàm đáp ứng phụ thuộc vào một hoặc nhiều hướng âm thanh.

Đối với từng thời gian và tần số thu được một hoặc nhiều tín hiệu micro tham chiếu.

Đối với từng thời gian và tần số, việc nhân một hoặc nhiều tín hiệu micro tham chiếu với một hoặc nhiều hàm đáp ứng để thu được một hoặc nhiều thành phần âm thanh bao quanh có bậc và kiểu mong muốn.

Nếu nhiều thành phần âm thanh bao quanh được thu cho bậc và kiểu mong muốn, cộng tổng các thành phần âm thanh bao quanh tương ứng để thu được thành phần âm thanh bao quanh mong muốn cuối cùng.

Trong một số phương án, tính toán trong bước 4 một hoặc nhiều âm thanh trực tiếp và các âm thanh khuếch tán từ nhiều tín hiệu micro thay vì một hoặc nhiều tín hiệu micro tham chiếu.

Nhân một hoặc nhiều âm thanh trực tiếp và các âm thanh khuếch tán với một hoặc nhiều đáp ứng âm thanh trực tiếp tương ứng và các đáp ứng âm thanh khuếch tán để thu được một hoặc nhiều thành phần âm thanh bao quanh của âm thanh khuếch tán cho bậc và kiểu mong muốn.

Các thành phần âm thanh khuếch tán có thể được giải tương quan bổ sung cho các bậc và các kiểu khác nhau.

Cộng tổng các thành phần âm thanh bao quanh của âm thanh trực tiếp và các thành phần âm thanh bao quanh của âm thanh khuếch tán để thu được thành phần âm thanh bao quanh mong muốn cuối cùng có bậc và kiểu mong muốn.

Mặc dù một số khía cạnh đã được mô tả trong phạm vi của thiết bị, rõ ràng rằng các khía cạnh này cũng thể hiện phần mô tả của phương pháp tương ứng, trong đó khối hoặc thiết bị tương ứng với bước phương pháp hoặc dấu hiệu của bước phương pháp. Tương tự, các khía cạnh được mô tả trong phạm vi bước thuộc phương pháp cũng biểu diễn phần mô tả khối hoặc mục hoặc dấu hiệu tương ứng của thiết bị tương ứng.

Tín hiệu đã được tách ra có thể được lưu trữ trong phương tiện lưu trữ số hoặc có thể được truyền trong môi trường truyền như môi trường truyền dẫn không dây hoặc môi trường truyền dẫn có dây như Internet.

Phụ thuộc vào các yêu cầu thực hiện nhất định, các phương án của sáng chế có thể được thực hiện trong phần cứng hoặc trong phần mềm. Phương án có thể được thực hiện 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 đọc được nhờ điện tử, chúng có khả năng kết hợp với hệ thống máy tính lập trình được, do đó một trong các phương pháp được mô tả ở đây được thực hiện.

Thông thường, các phương án theo sáng chế có thể được thực hiện như sản phẩm chương trình máy tính với mã chương trình, mã chương trình hoạt động để thực hiện một trong 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 vật ghi có thể đọc được bằng máy tính bao gồm 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 phương tiện lưu trữ số, hoặc vật ghi có thể đọc được bằng máy tính) chứa chương trình máy tính đã được ghi lại trên đó để thực hiện một trong số các phương pháp được mô tả ở đây.

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

Phương án nữa bao gồm máy tính được thiết lập ở trên đó chương trình máy tính 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 ưu tiên được thực hiện bởi các 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 ra phần mô tả trường âm thanh có phép biểu diễn các thành phần trường âm thanh, thiết bị bao gồm:

bộ xác định hướng (102) để xác định một hoặc nhiều hướng âm thanh cho từng ô thời gian-tần số thuộc nhiều ô thời gian-tần số của nhiều tín hiệu micrô;

bộ đánh giá hàm cơ sở trong không gian (103) để đánh giá, cho từng ô thời gian-tần số thuộc nhiều ô thời gian-tần số, một hoặc nhiều hàm cơ sở trong không gian sử dụng một hoặc nhiều hướng âm thanh; và

bộ tính toán thành phần trường âm thanh (201) để tính toán, cho từng ô thời gian-tần số thuộc nhiều ô thời gian-tần số, một hoặc nhiều thành phần trường âm thanh tương ứng với một hoặc nhiều hàm cơ sở trong không gian sử dụng một hoặc nhiều hàm cơ sở trong không gian được đánh giá bằng cách sử dụng một hoặc nhiều hướng âm thanh và sử dụng tín hiệu tham chiếu cho ô thời gian-tần số tương ứng, tín hiệu tham chiếu được suy ra từ một hoặc nhiều tín hiệu micrô thuộc nhiều tín hiệu micrô.

2. Thiết bị theo điểm 1, thiết bị còn bao gồm:

bộ tính toán thành phần khuếch tán (301) để tính toán, cho từng ô thời gian-tần số thuộc nhiều ô thời gian-tần số, một hoặc nhiều thành phần âm thanh khuếch tán; và

bộ kết hợp (401) để kết hợp thông tin âm thanh khuếch tán với thông tin trường âm thanh trực tiếp để thu được phép biểu diễn trong miền tần số hoặc phép biểu diễn trong miền thời gian của các thành phần trường âm thanh.

3. Thiết bị theo điểm 2, trong đó bộ tính toán thành phần khuếch tán (301) còn bao gồm bộ giải tương quan (107) để giải tương quan thông tin âm thanh khuếch tán.

4. Thiết bị theo điểm bất kỳ trong số các điểm nêu trên, thiết bị còn bao gồm bộ chuyển đổi thời gian-tần số (101) để chuyển đổi từng tín hiệu trong số nhiều tín hiệu micrô trong miền thời gian thành phép biểu diễn tần số có nhiều ô thời gian-tần số.

5. Thiết bị theo một trong số các điểm nêu trên, thiết bị còn bao gồm bộ chuyển đổi tần số-thời gian (20) để chuyển đổi một hoặc nhiều thành phần trường âm thanh hoặc kết hợp của một hoặc nhiều thành phần trường âm thanh với các thành phần âm thanh

khuếch tán thành phép biểu diễn trong miền thời gian của các thành phần trường âm thanh.

6. Thiết bị theo điểm 5,

trong đó bộ chuyển đổi tần số-thời gian (20) được tạo cấu hình để xử lý một hoặc nhiều thành phần trường âm thanh để thu được nhiều thành phần trường âm thanh trong miền thời gian, trong đó bộ chuyển đổi tần số-thời gian được tạo cấu hình để xử lý các thành phần âm thanh khuếch tán để thu được nhiều thành phần khuếch tán trong miền thời gian, và trong đó bộ kết hợp (401) được tạo cấu hình để thực hiện phép kết hợp các thành phần trường âm thanh trong miền thời gian với các thành phần khuếch tán trong miền thời gian trong miền thời gian; hoặc

trong đó bộ kết hợp (401) được tạo cấu hình để kết hợp một hoặc nhiều thành phần trường âm thanh cho ô thời gian-tần số và các thành phần âm thanh khuếch tán cho ô thời gian-tần số tương ứng trong miền tần số, và trong đó bộ chuyển đổi tần số-thời gian (20) được tạo cấu hình để xử lý kết quả của bộ kết hợp (401) để thu được các thành phần trường âm thanh trong miền thời gian.

7. Thiết bị theo một trong số các điểm nêu trên, còn bao gồm bộ tính toán tín hiệu tham chiếu (104) để tính toán tín hiệu tham chiếu từ nhiều tín hiệu micrô

sử dụng một hoặc nhiều hướng âm thanh,

sử dụng việc lựa chọn tín hiệu micrô cụ thể từ nhiều tín hiệu micrô dựa trên một hoặc nhiều hướng âm thanh, hoặc

sử dụng bộ lọc đa kênh được áp dụng cho hai hoặc nhiều hơn hai tín hiệu micrô, bộ lọc đa kênh phụ thuộc vào một hoặc nhiều hướng âm thanh và các vị trí riêng lẻ của các micrô, từ đó nhiều tín hiệu micrô được thu.

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

trong đó bộ đánh giá hàm cơ sở trong không gian (103) được tạo cấu hình để sử dụng cho hàm cơ sở trong không gian, phép biểu diễn được tham số hóa, trong đó tham số của phép biểu diễn được tham số hóa là hướng âm thanh, và để chèn tham số

tương ứng với hướng âm thanh vào trong phép biểu diễn được tham số hóa để thu được kết quả đánh giá cho từng hàm cơ sở trong không gian; hoặc

trong đó bộ đánh giá hàm cơ sở trong không gian (103) được tạo cấu hình để sử dụng bảng tìm kiếm cho từng hàm cơ sở trong không gian có sự nhận dạng hàm cơ sở trong không gian làm đầu vào, và hướng âm thanh, và có kết quả đánh giá làm đầu ra, và trong đó bộ đánh giá hàm cơ sở trong không gian (103) được tạo cấu hình để xác định, cho một hoặc nhiều hướng âm thanh được xác định bởi bộ xác định hướng, hướng âm thanh, hướng âm thanh tương ứng của đầu vào của bảng tìm kiếm hoặc để tính toán số trung bình được gán trọng số hoặc không được gán trọng số giữa hai đầu vào của bảng tìm kiếm gần kề với một hoặc nhiều hướng âm thanh được xác định bởi bộ xác định hướng; hoặc

trong đó bộ đánh giá hàm cơ sở trong không gian (103) được tạo cấu hình để sử dụng cho hàm cơ sở trong không gian, phép biểu diễn được tham số hóa, trong đó tham số của phép biểu diễn được tham số hóa là hướng âm thanh, hướng âm thanh là một chiều, như góc phương vị, trong trạng thái không gian hai chiều hoặc hai chiều, như góc phương vị và góc ngẩng, trong trạng thái không gian ba chiều, và để chèn tham số tương ứng với hướng âm thanh vào trong phép biểu diễn được tham số hóa để thu được kết quả đánh giá cho từng hàm cơ sở trong không gian.

9. Thiết bị theo một trong số các điểm nêu trên còn bao gồm:

bộ xác định âm thanh trực tiếp hoặc khuếch tán (105) để xác định phân trực tiếp hoặc phân khuếch tán của nhiều tín hiệu micro như tín hiệu tham chiếu,

trong đó bộ tính toán thành phần trường âm thanh (201) được tạo cấu hình để sử dụng phân trực tiếp chỉ để tính toán một hoặc nhiều thành phần trường âm thanh trực tiếp.

10. Thiết bị theo điểm 9, thiết bị còn bao gồm bộ xác định hàm cơ sở của đáp ứng trung bình (106) để xác định đáp ứng của hàm cơ sở trong không gian trung bình, bộ xác định bao gồm quy trình tính toán hoặc quy trình truy cập bảng tìm kiếm; và

bộ tính toán thành phần âm thanh khuếch tán (301), để tính toán một hoặc nhiều thành phần trường âm thanh khuếch tán chỉ sử dụng phần khuếch tán như tín hiệu tham chiếu cùng với đáp ứng của hàm cơ sở trong không gian trung bình.

11. Thiết bị theo điểm 10, còn bao gồm bộ kết hợp (109, 401) để kết hợp thành phần trường âm thanh trực tiếp

với thành phần trường âm thanh khuếch tán để thu được thành phần trường âm thanh.

12. Thiết bị theo một trong số các điểm từ 9 đến 11,

trong đó bộ tính toán thành phần âm thanh khuếch tán (301) được tạo cấu hình để tính toán các thành phần âm thanh khuếch tán lên đến số hoặc bậc thứ nhất được xác định trước,

trong đó bộ tính toán thành phần trường âm thanh (201) được tạo cấu hình để tính toán các thành phần trường âm thanh trực tiếp lên đến số hoặc bậc thứ hai được xác định trước,

trong đó số hoặc bậc thứ hai được xác định trước lớn hơn số hoặc bậc thứ nhất được xác định trước, và

trong đó số hoặc bậc thứ nhất được xác định trước là 1 hoặc lớn hơn 1.

13. Thiết bị theo một trong số các điểm từ 10 đến 12,

trong đó bộ tính toán thành phần tín hiệu khuếch tán (105) bao gồm bộ giải tương quan (107) để giải tương quan thành phần âm thanh khuếch tán trước hoặc sau khi kết hợp với đáp ứng trung bình của hàm cơ sở trong không gian trong phép biểu diễn trong miền tần số hoặc phép biểu diễn trong miền thời gian.

14. Thiết bị theo một điểm trong số các điểm từ 9 đến 13, trong đó bộ xác định âm thanh trực tiếp hoặc âm thanh khuếch tán (105) được tạo cấu hình để:

tính toán phần trực tiếp và phần khuếch tán từ tín hiệu micro đơn lẻ, và trong đó bộ tính toán thành phần âm thanh khuếch tán (301) được tạo cấu hình để tính toán một hoặc nhiều thành phần âm thanh khuếch tán sử dụng phần khuếch tán như tín hiệu tham chiếu, và trong đó bộ tính toán thành phần trường âm thanh (201) được tạo cấu

hình để tính toán một hoặc nhiều thành phần trường âm thanh trực tiếp sử dụng phần trực tiếp như tín hiệu tham chiếu; hoặc

tính toán phần khuếch tán từ tín hiệu micrô mà khác với tín hiệu micrô, mà từ đó phần trực tiếp được tính toán, và trong đó bộ tính toán thành phần âm thanh khuếch tán được tạo cấu hình để tính toán một hoặc nhiều thành phần âm thanh khuếch tán sử dụng phần khuếch tán như tín hiệu tham chiếu, và trong đó bộ tính toán thành phần trường âm thanh (201) được tạo cấu hình để tính toán một hoặc nhiều thành phần trường âm thanh trực tiếp sử dụng phần trực tiếp như tín hiệu tham chiếu; hoặc

tính toán phần khuếch tán cho hàm cơ sở trong không gian khác nhau sử dụng tín hiệu micrô khác nhau, và trong đó bộ tính toán thành phần trường âm thanh khuếch tán (301) được tạo cấu hình để sử dụng phần khuếch tán thứ nhất như tín hiệu tham chiếu cho đáp ứng của hàm cơ sở trong không gian trung bình tương ứng với số thứ nhất, và để sử dụng phần khuếch tán thứ hai khác nhau như tín hiệu tham chiếu tương ứng với đáp ứng của hàm cơ sở trong không gian trung bình số thứ hai, trong đó số thứ nhất khác với số thứ hai, và trong đó số thứ nhất và số thứ hai biểu thị bậc hoặc mức hoặc kiểu bất kỳ của một hoặc nhiều hàm cơ sở trong không gian; hoặc

tính toán phần trực tiếp sử dụng bộ lọc đa kênh thứ nhất được áp dụng cho nhiều tín hiệu micrô và tính toán phần khuếch tán sử dụng bộ lọc đa kênh thứ hai được áp dụng cho nhiều tín hiệu micrô, bộ lọc đa kênh thứ hai khác với bộ lọc đa kênh thứ nhất, và trong đó bộ tính toán thành phần âm thanh khuếch tán (301) được tạo cấu hình để tính toán một hoặc nhiều thành phần âm thanh khuếch tán sử dụng phần khuếch tán làm tín hiệu tham chiếu, và trong đó bộ tính toán thành phần trường âm thanh (201) được tạo cấu hình để tính toán một hoặc nhiều thành phần trường âm thanh trực tiếp sử dụng phần trực tiếp làm tín hiệu tham chiếu; hoặc

tính toán các phần khuếch tán cho các hàm cơ sở trong không gian khác nhau sử dụng các bộ lọc đa kênh khác nhau cho các hàm cơ sở trong không gian khác nhau, và trong đó bộ tính toán thành phần âm thanh khuếch tán (301) được tạo cấu hình để tính toán một hoặc nhiều thành phần âm thanh khuếch tán sử dụng phần khuếch tán làm tín hiệu tham chiếu, và trong đó bộ tính toán thành phần trường âm thanh (201) được tạo

cấu hình để tính toán một hoặc nhiều thành phần trường âm thanh trực tiếp sử dụng phần trực tiếp làm tín hiệu tham chiếu.

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

trong đó bộ đánh giá hàm cơ sở trong không gian (103) bao gồm bộ làm nhẵn độ khuếch đại (111) hoạt động theo hướng thời gian hoặc hướng tần số, để làm nhẵn các kết quả đánh giá, và

trong đó bộ tính toán thành phần trường âm thanh (201) được tạo cấu hình để sử dụng các kết quả của bộ đánh giá được làm nhẵn trong việc tính toán một hoặc nhiều thành phần trường âm thanh.

16. Thiết bị theo một trong số các điểm nêu trên, trong đó bộ đánh giá hàm cơ sở trong không gian (103) được tạo cấu hình để tính toán, đối với từng hướng âm thanh thuộc ít nhất hai hướng âm thanh cho ô thời gian-tần số mà được xác định bởi bộ xác định hướng, kết quả đánh giá đối với từng hàm cơ sở trong không gian của một hoặc nhiều hơn hai hàm cơ sở trong không gian,

trong đó bộ tính toán tín hiệu tham chiếu (104) được tạo cấu hình để tính toán, cho từng hướng âm thanh, các tín hiệu tham chiếu riêng rẽ,

trong đó bộ tính toán thành phần trường âm thanh (103) được tạo cấu hình để tính toán thành phần trường âm thanh cho từng hướng sử dụng kết quả đánh giá cho hướng âm thanh và tín hiệu tham chiếu cho hướng âm thanh, và

trong đó bộ tính toán thành phần trường âm thanh được tạo cấu hình để bổ sung các thành phần trường âm thanh cho các hướng khác nhau được tính toán sử dụng hàm cơ sở trong không gian để thu được thành phần trường âm thanh cho hàm cơ sở trong không gian trong ô thời gian-tần số.

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

trong đó bộ đánh giá hàm cơ sở trong không gian (103) được tạo cấu hình để sử dụng một hoặc nhiều hàm cơ sở trong không gian cho âm thanh bao quanh trong trạng thái không gian hai chiều hoặc ba chiều.

18. Thiết bị theo điểm 17,

trong đó bộ tính toán hàm cơ sở trong không gian (103) được tạo cấu hình để sử dụng ít nhất các hàm cơ sở trong không gian có ít nhất hai bậc hoặc hai mức hoặc ít nhất hai kiểu.

19. Thiết bị theo điểm 18,

trong đó bộ tính toán thành phần trường âm thanh (201) được tạo cấu hình để tính toán thành phần trường âm thanh cho ít nhất hai mức của nhóm có các mức bao gồm mức 0, mức 1, mức 2, mức 3, mức 4, hoặc

trong đó bộ tính toán thành phần trường âm thanh (201) được tạo cấu hình để tính toán các thành phần trường âm thanh cho ít nhất hai kiểu của nhóm có các kiểu bao gồm kiểu -4, kiểu -3, kiểu -2, kiểu -1, kiểu 0, kiểu 1, kiểu 2, kiểu 3, kiểu 4.

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

bộ tính toán thành phần khuếch tán (301) để tính toán, cho từng ô thời gian-tần số thuộc nhiều ô thời gian-tần số, một hoặc nhiều thành phần âm thanh khuếch tán; và

bộ kết hợp (401) để kết hợp thông tin âm thanh khuếch tán với thông tin trường âm thanh trực tiếp để thu được phép biểu diễn trong miền tần số hoặc phép biểu diễn trong miền thời gian của các thành phần trường âm thanh,

trong đó bộ tính toán thành phần khuếch tán hoặc bộ kết hợp được tạo cấu hình để tính toán hoặc để kết hợp thành phần khuếch tán cho đến khi bậc nhất định hoặc số nhất định nhỏ hơn bậc hoặc số lên đến mức mà bộ tính toán thành phần trường âm thanh (201) được tạo cấu hình để tính toán thành phần trường âm thanh trực tiếp.

21. Thiết bị theo điểm 20, trong đó bậc hoặc số nhất định là một hoặc không, và bậc hoặc số lên đến mức mà bộ tính toán thành phần trường âm thanh (201) được tạo cấu hình để tính toán thành phần trường âm thanh là 2 hoặc lớn hơn 2.

22. Thiết bị theo điểm bất kì trong số các điểm nêu trên,

trong đó bộ tính toán thành phần trường âm thanh (201) được tạo cấu hình để nhân (115) tín hiệu trong ô thời gian-tần số của tín hiệu tham chiếu với kết quả đánh giá thu được từ hàm cơ sở trong không gian để thu được thông tin về thành phần trường âm thanh được kết hợp với hàm cơ sở trong không gian, và để nhân (115) tín

hiệu trong ô thời gian-tần số của tín hiệu tham chiếu với kết quả đánh giá nữa được thu từ hàm cơ sở trong không gian nữa để thu được thông tin thông tin về thành phần trường âm thanh nữa được kết hợp với hàm cơ sở trong không gian nữa.

23. Phương pháp tạo ra phần mô tả trường âm thanh có phép biểu diễn của các thành phần trường âm thanh, phương pháp bao gồm các bước:

xác định (102) một hoặc nhiều hướng âm thanh cho từng ô thời gian-tần số thuộc nhiều ô thời gian -tần số của nhiều tín hiệu micro;

đánh giá (103) một hoặc nhiều hàm cơ sở trong không gian bằng cách sử dụng một hoặc nhiều hướng âm thanh cho từng ô thời gian-tần số thuộc nhiều ô thời gian-tần số; và

tính toán (201) một hoặc nhiều thành phần trường âm thanh tương ứng với một hoặc nhiều hàm cơ sở trong không gian cho từng ô thời gian-tần số thuộc nhiều ô thời gian-tần số bằng cách sử dụng một hoặc nhiều hàm cơ sở trong không gian được đánh giá bằng cách sử dụng một hoặc nhiều hướng âm thanh và sử dụng tín hiệu tham chiếu cho ô thời gian-tần số tương ứng, tín hiệu tham chiếu được suy ra từ một hoặc nhiều tín hiệu micro thuộc nhiều tín hiệu micro.

24. Vật ghi có thể đọc được bằng máy tính bao gồm chương trình máy tính để thực hiện phương pháp tạo ra phần mô tả trường âm thanh có các thành phần trường âm thanh theo điểm 23, khi chạy trên máy tính hoặc bộ xử lý.

1/16

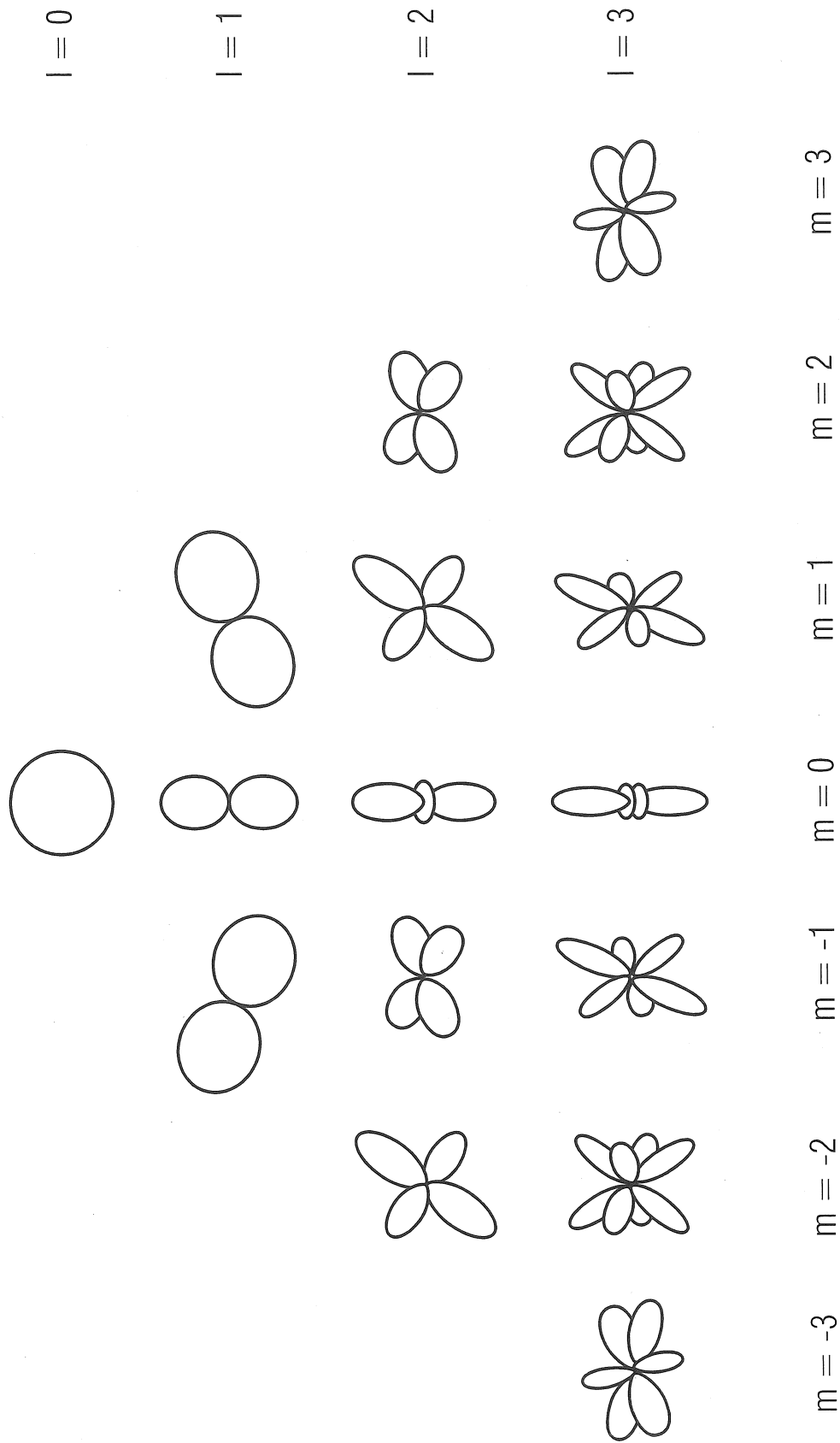


Fig. 1a

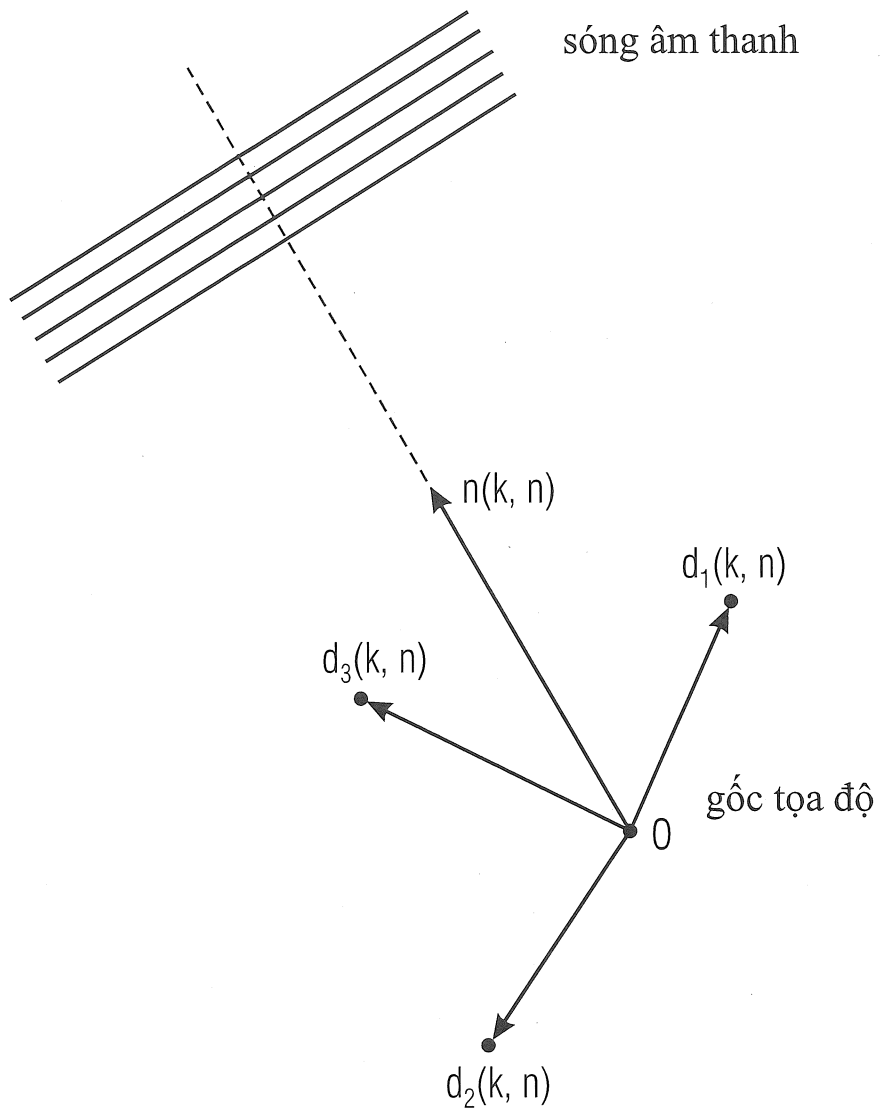


Fig. 1b

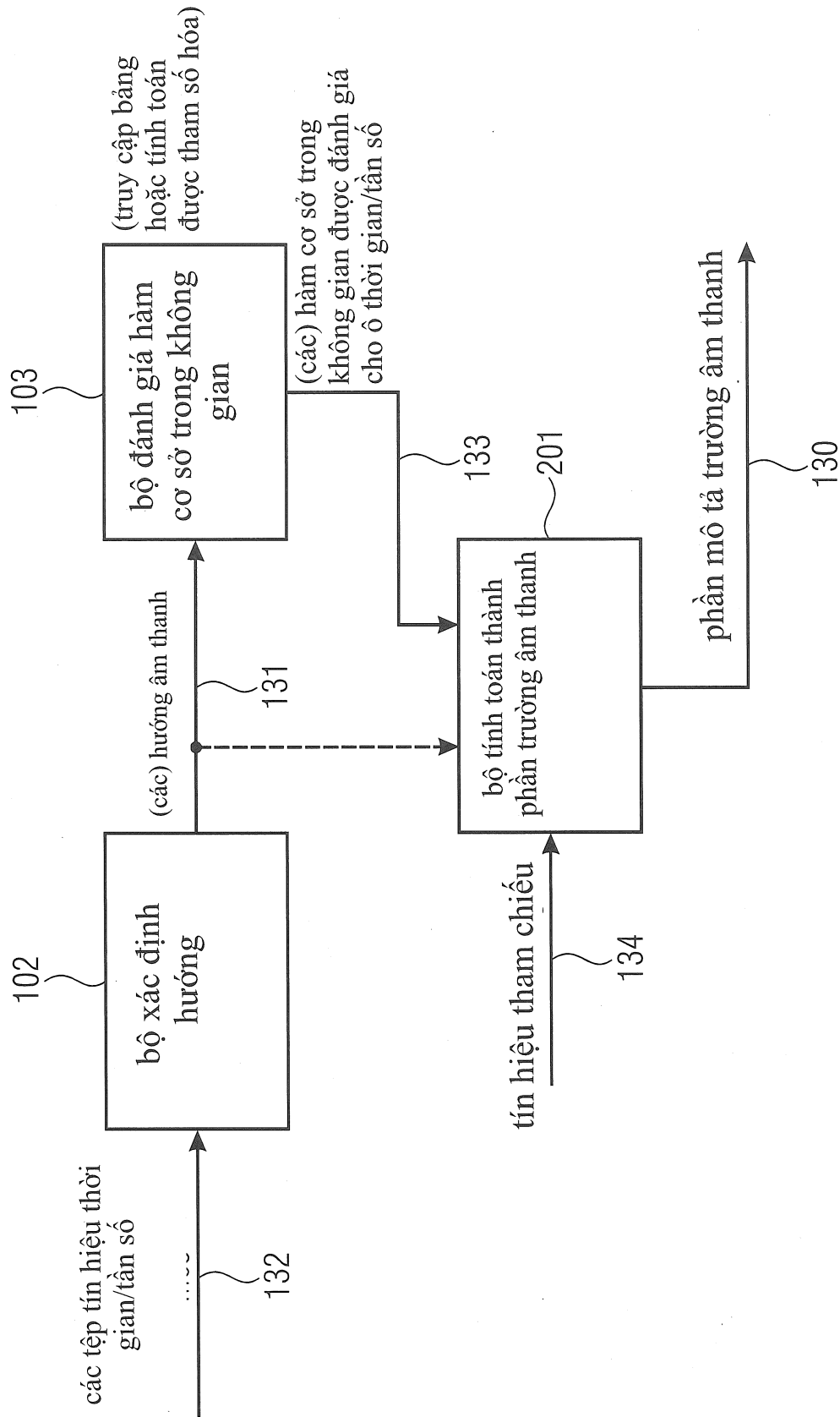
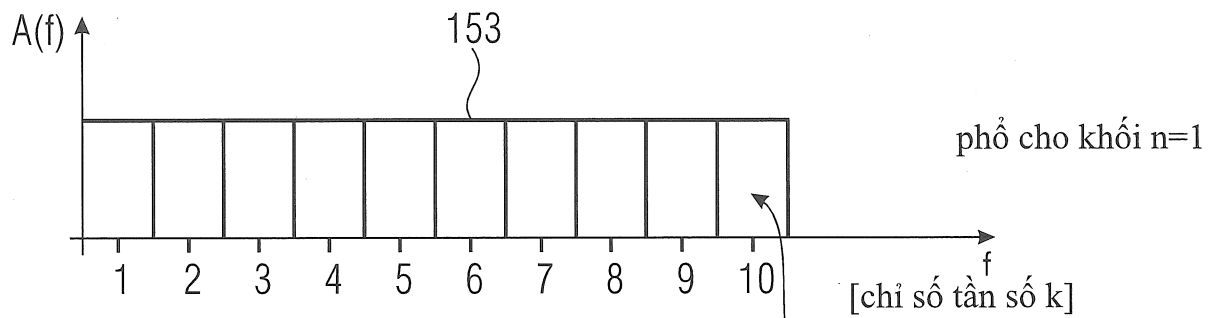
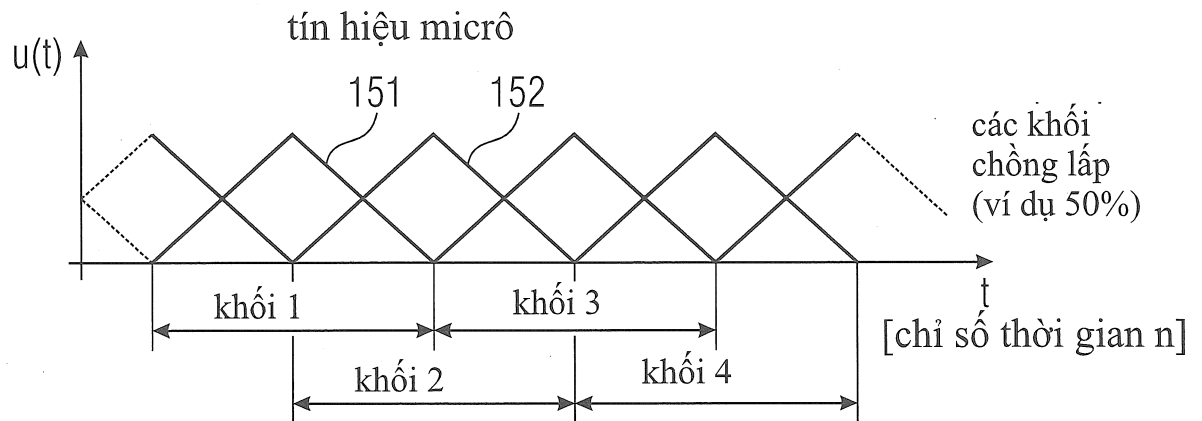
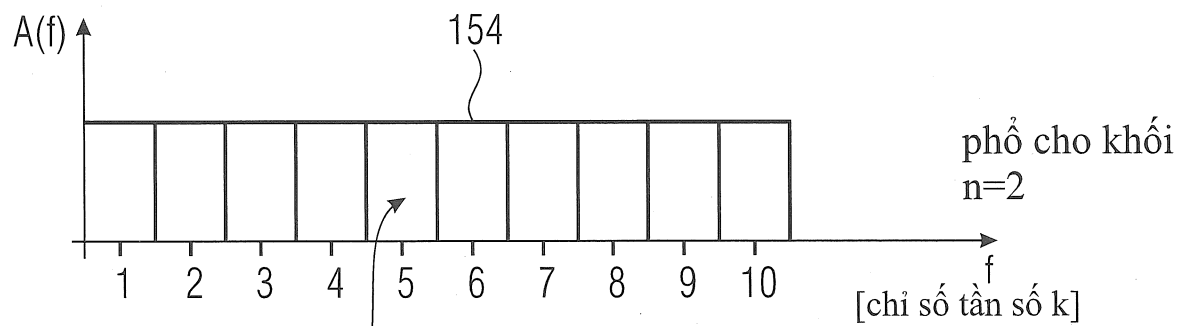


Fig. 1c



ô thời gian-tần số (k, n)
k = 10, n = 1



ô thời gian-tần số (k, n)
k = 5, n = 2

Fig. 1d

5/16

- ô thời gian - tần số (10, 1)

có DOA $\vec{n}$ (10, 1) $\vec{n}$: vectơ đơn vị chuẩn

- ô thời gian - tần số (5, 2) $\vec{n} = \begin{pmatrix} n_x \\ n_y \\ n_z \end{pmatrix}$ trường hợp không gian 3-D

có DOA $\vec{n}$ (5, 2) $\vec{n} = \begin{pmatrix} n_x \\ n_y \end{pmatrix}$ trường hợp không gian 2-D

- các hàm cơ sở trong không gian Y_i (DOA)

e.g. $i = 1$: đẳng hướng $i = 2$: hướng theo chuớng x $i = 3$: hướng theo hướng y $i = 4$: hướng theo hướng z

⋮



các hàm cơ sở trong không gian được đánh giá $G_i(k, n)$
cho các ô thời gian-tần số

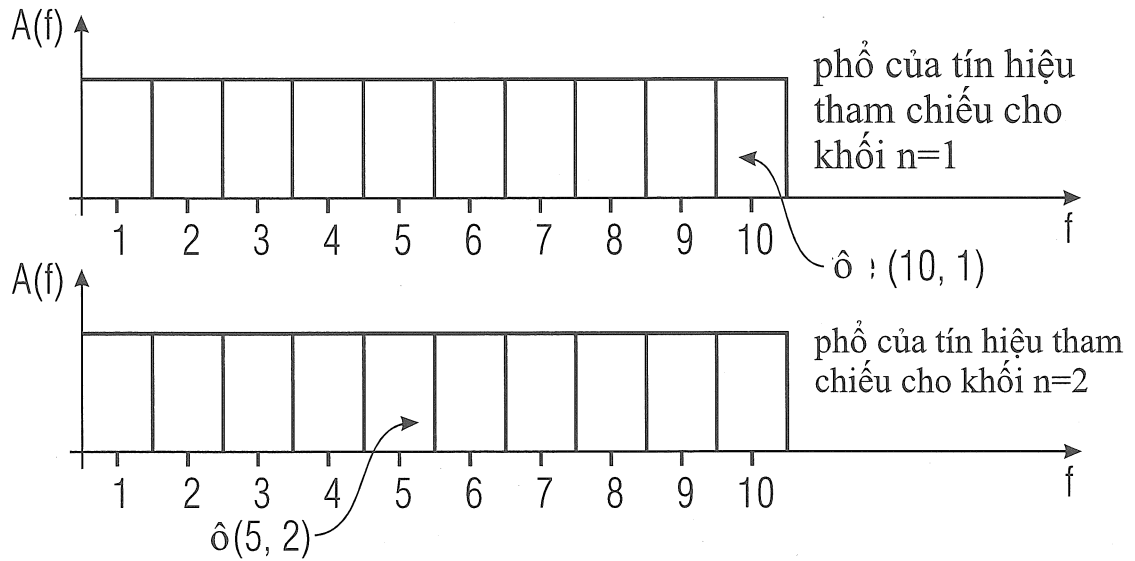
...	...	$G_1(10, 1)$	...	...	$G_1(5, 2)$	...	...
		$G_2(10, 1)$			$G_2(5, 2)$		
		$G_3(10, 1)$			$G_3(5, 2)$		
		$G_4(10, 1)$			$G_4(5, 2)$		
		f cho khối n=1			cho khối n=2		

Fig. 1e

6/16

tính toán các thành phần trường âm thanh

- tín hiệu tham chiếu trong phép biểu diễn thời gian-tần số



$$B_i = f(P, G_i)$$

155

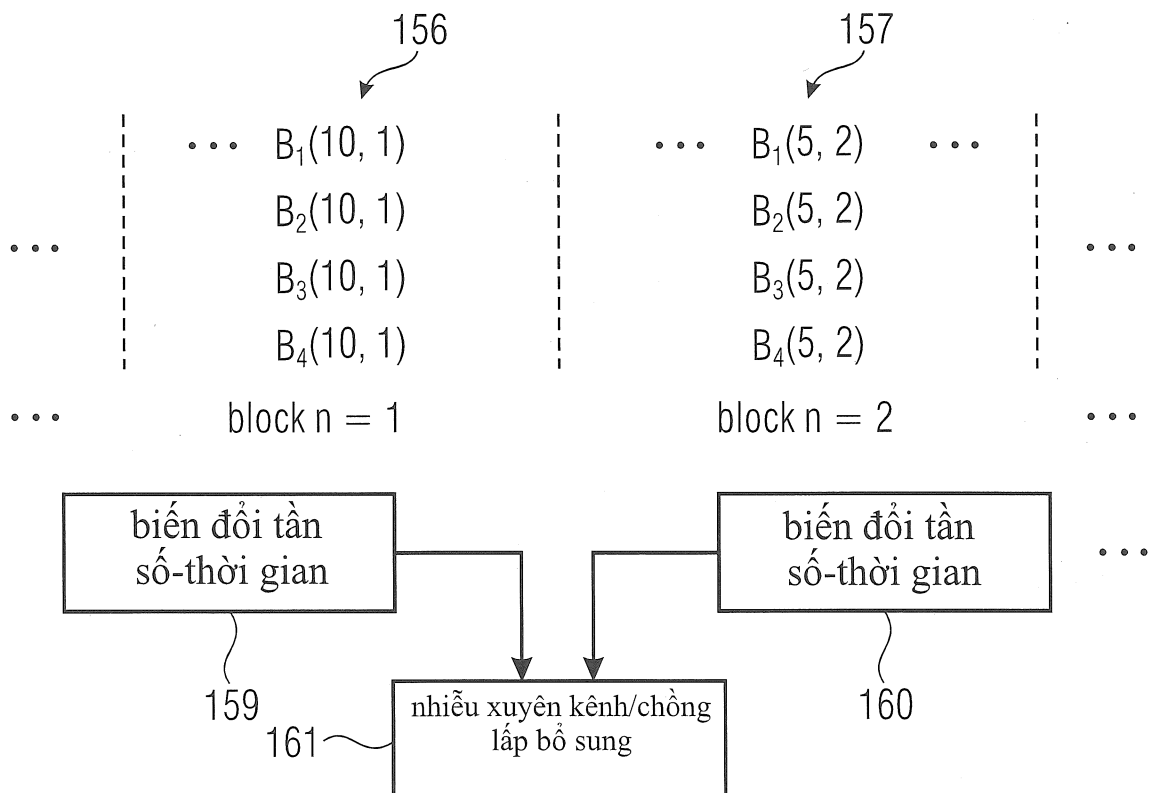
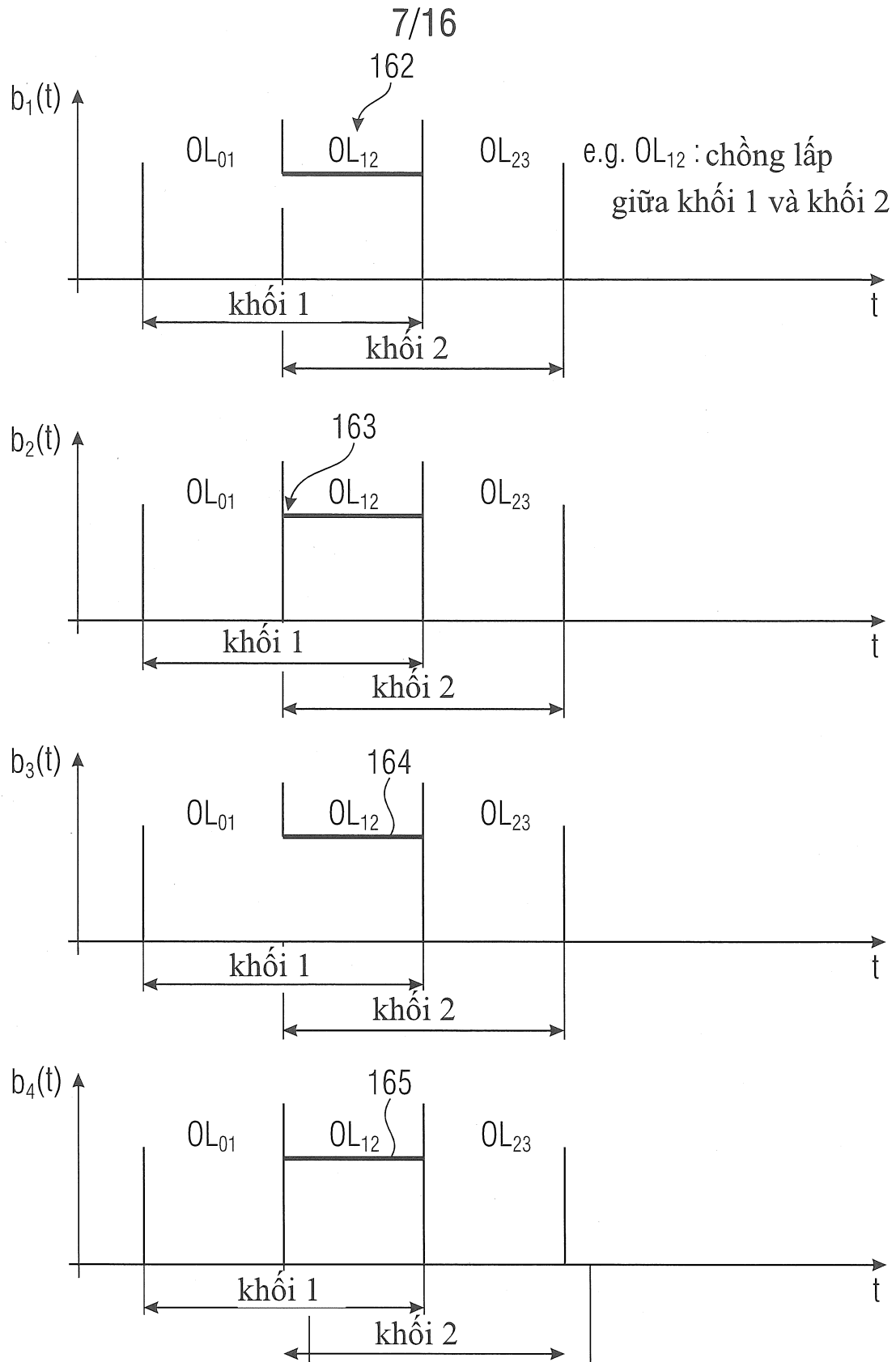
phép biểu diễn (phổ) miền tần số của các thành phần trường âm thanh B_i 

Fig. 1f



phép biểu diễn miền thời gian của các thành phần trường âm thanh bi

Fig. 1g

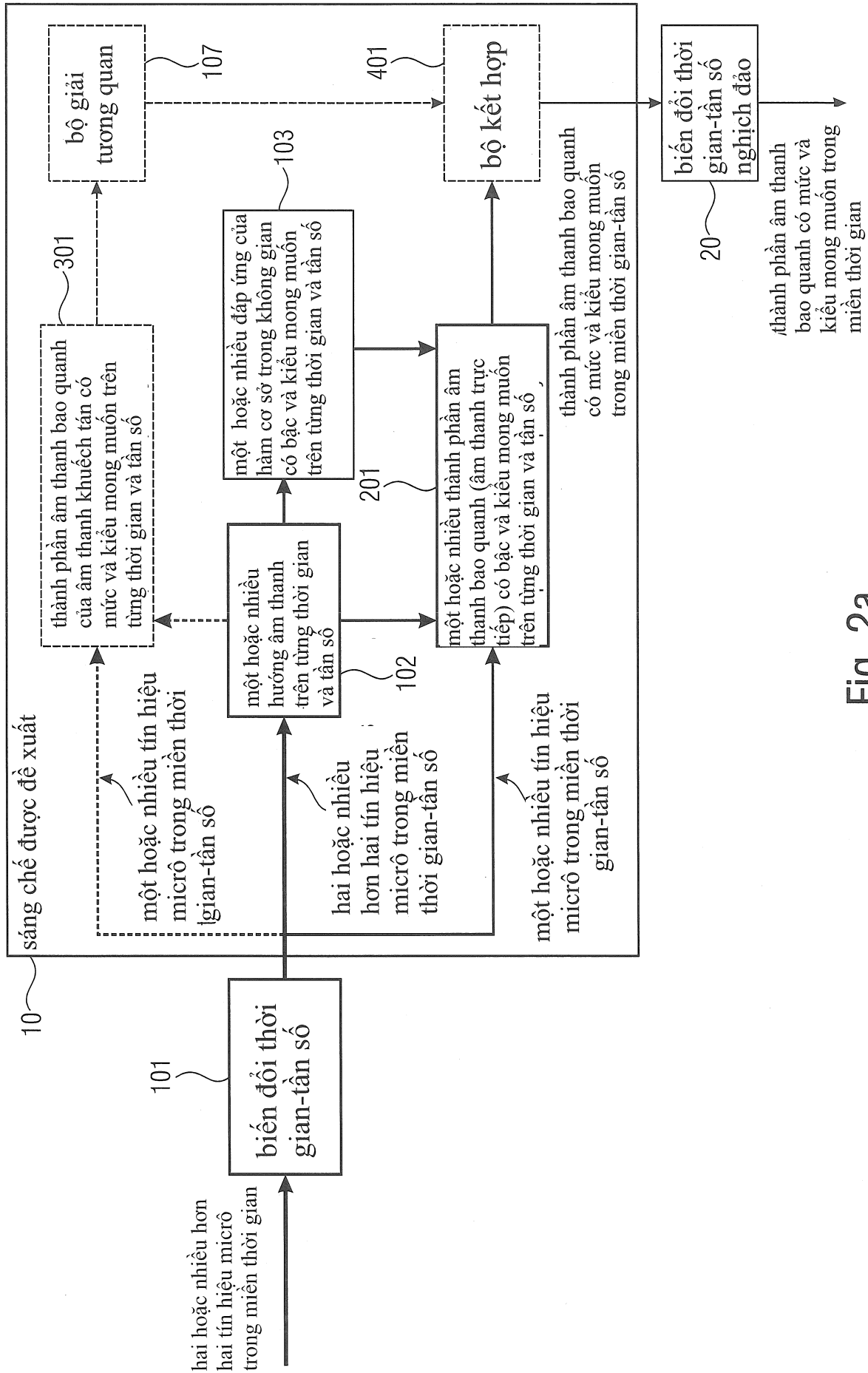


Fig. 2a

9/16

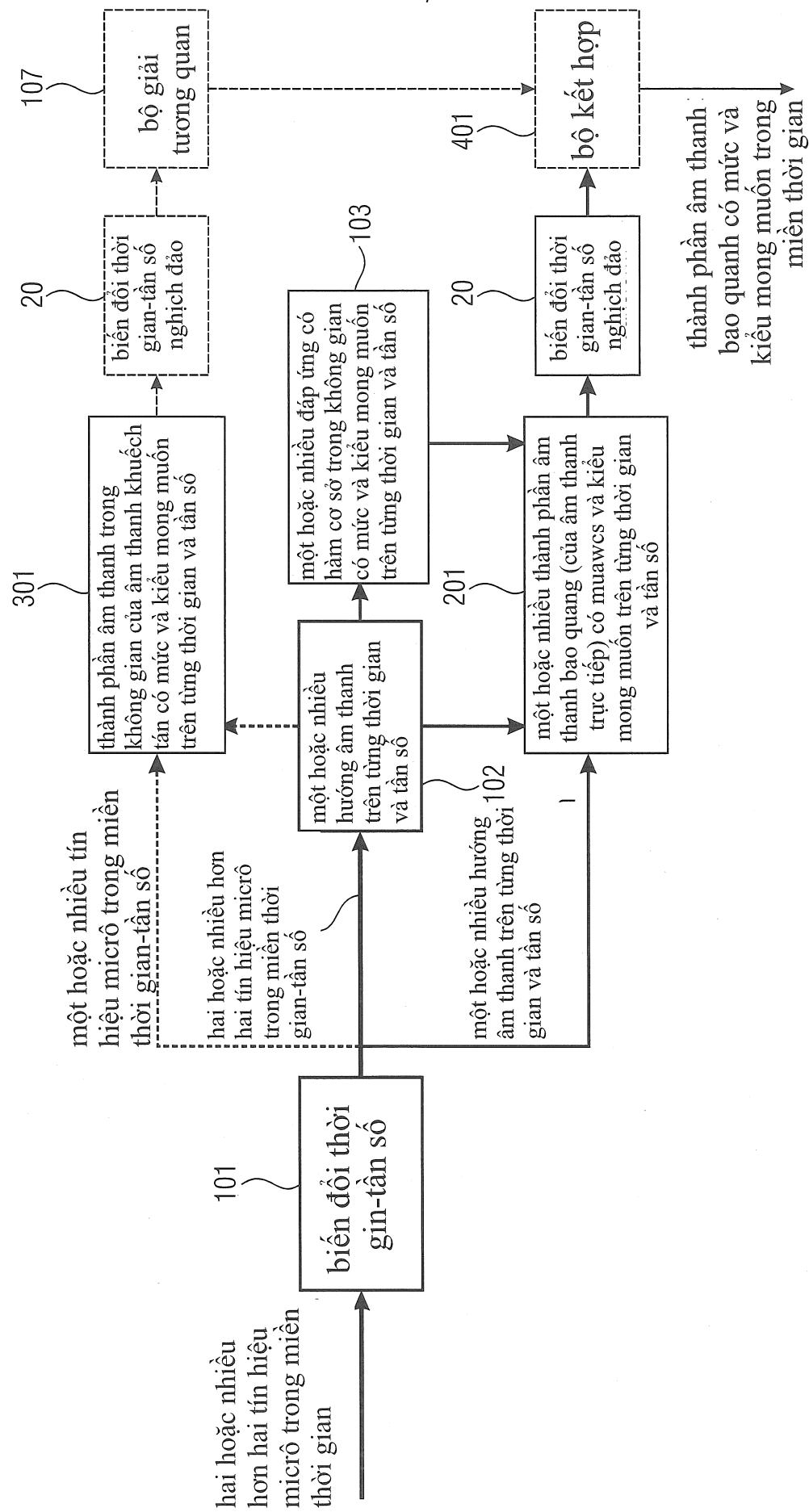


Fig. 2b

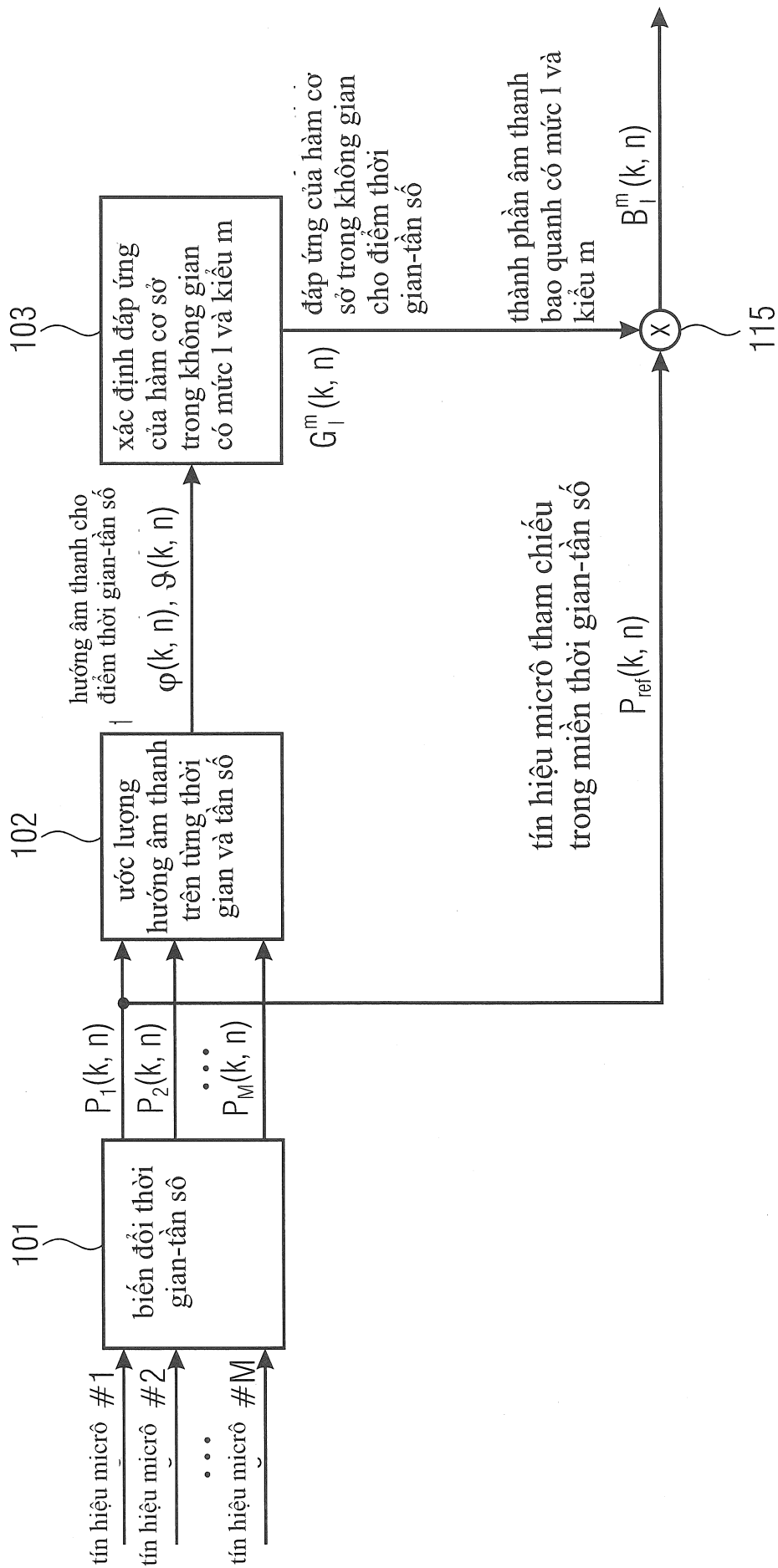


Fig. 3a

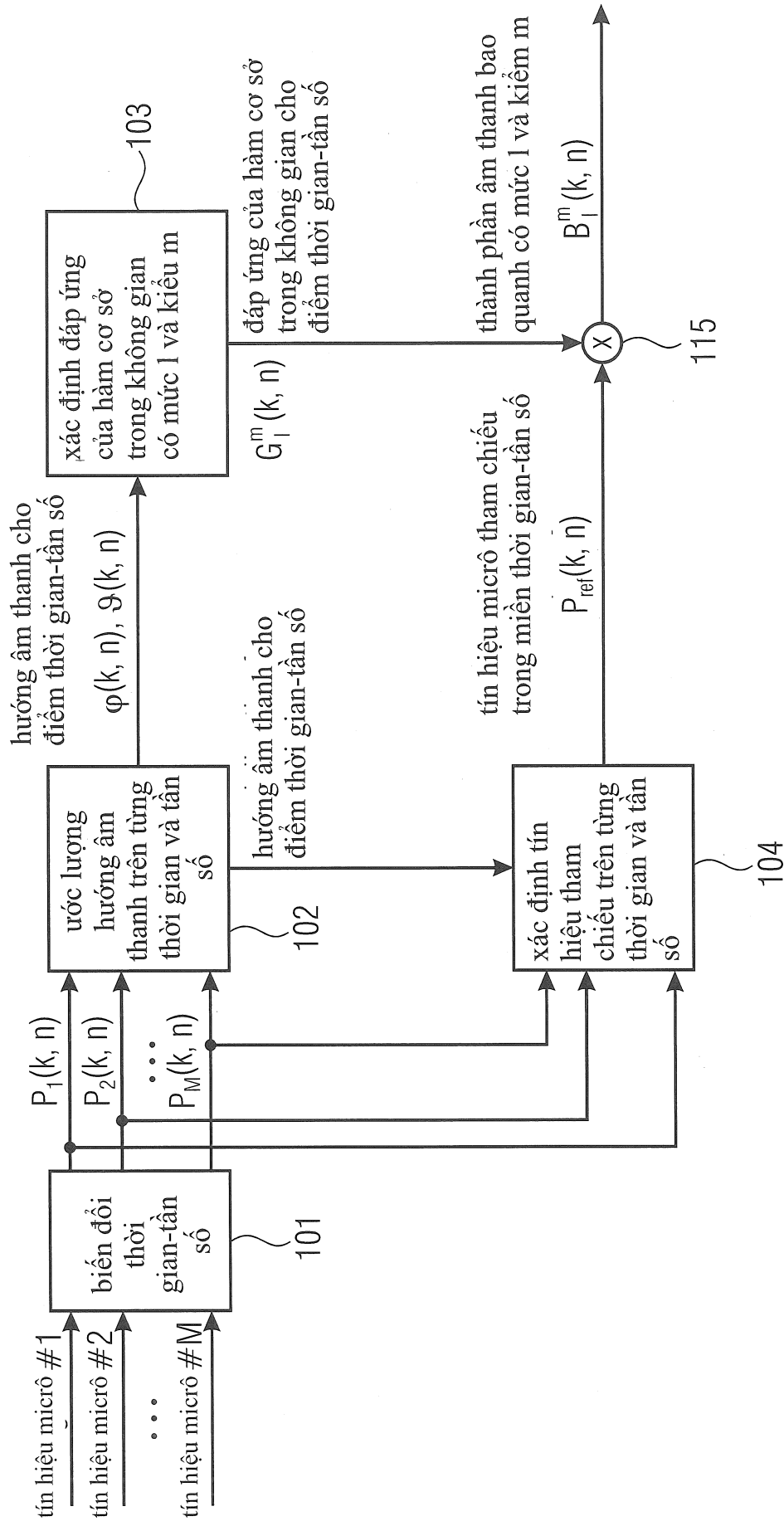


Fig. 3b

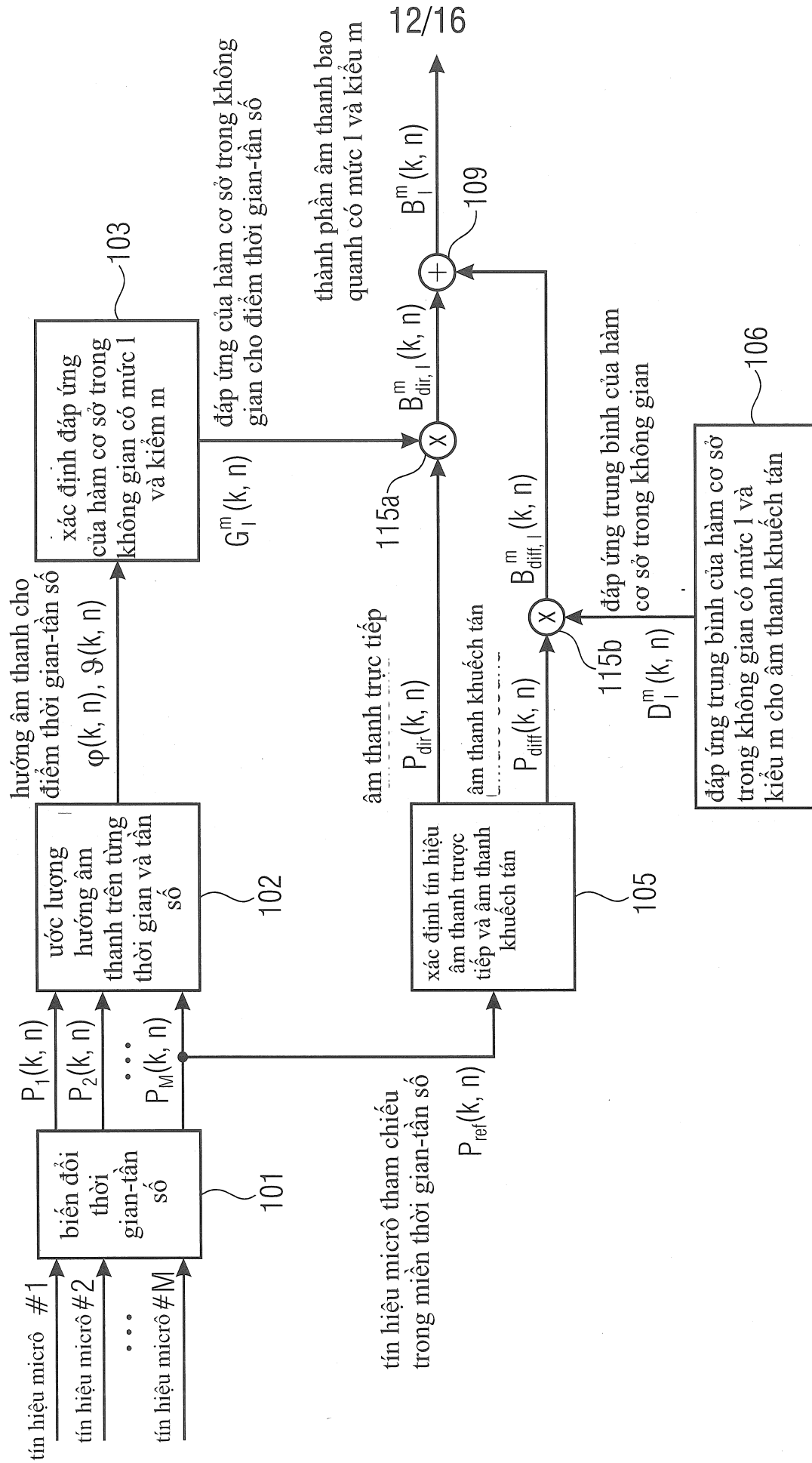


Fig. 4

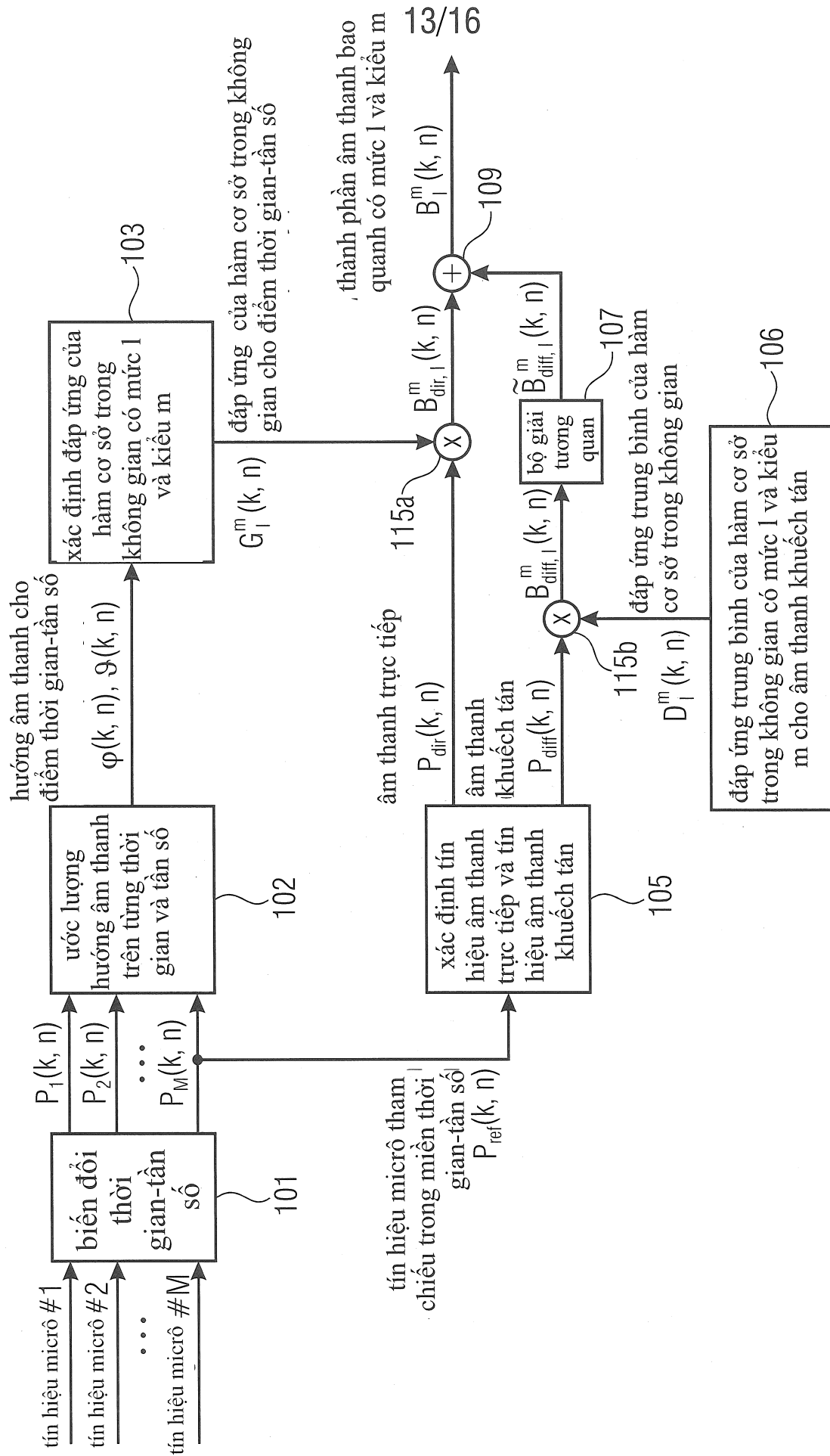


Fig. 5

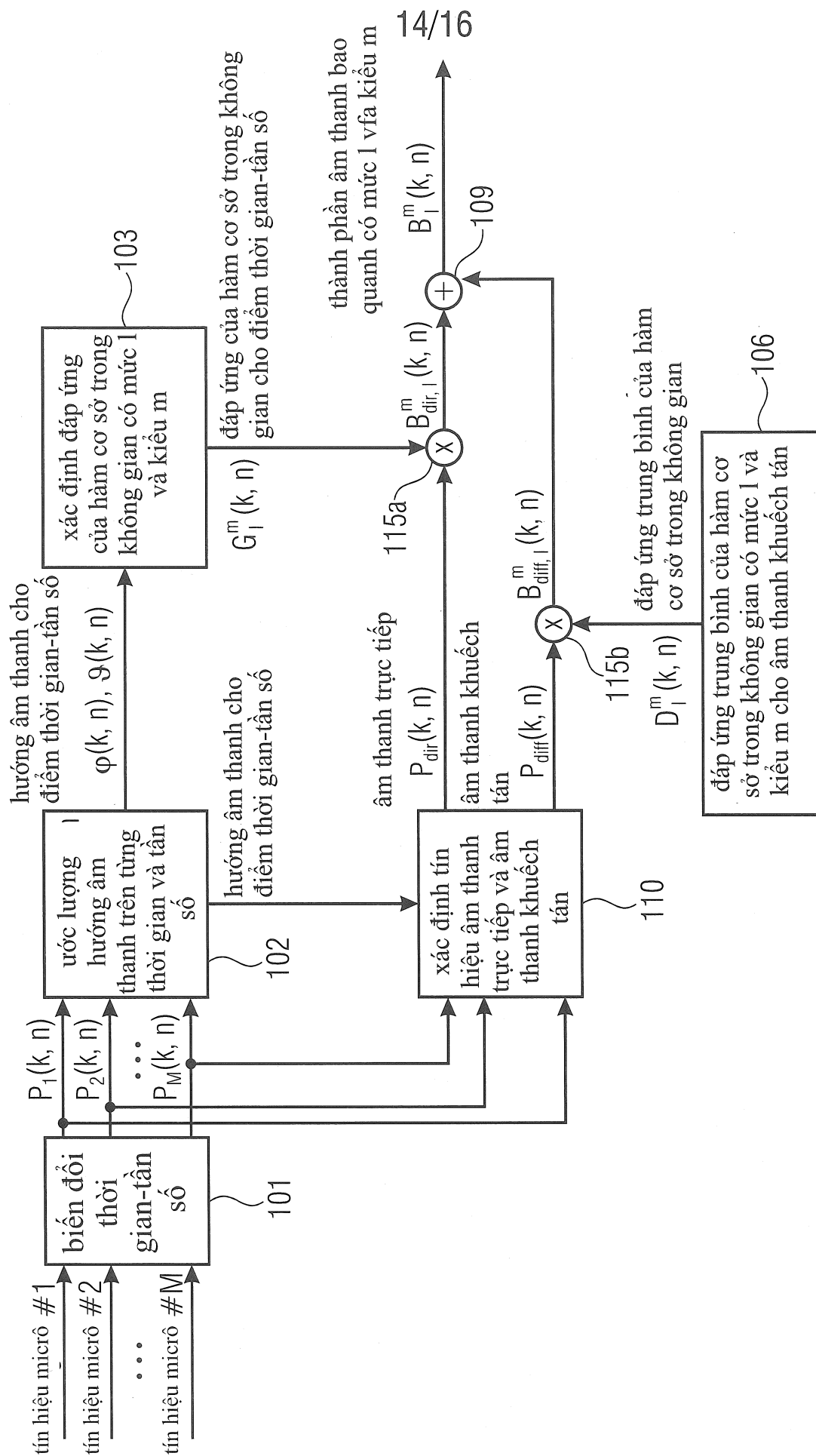


Fig. 6

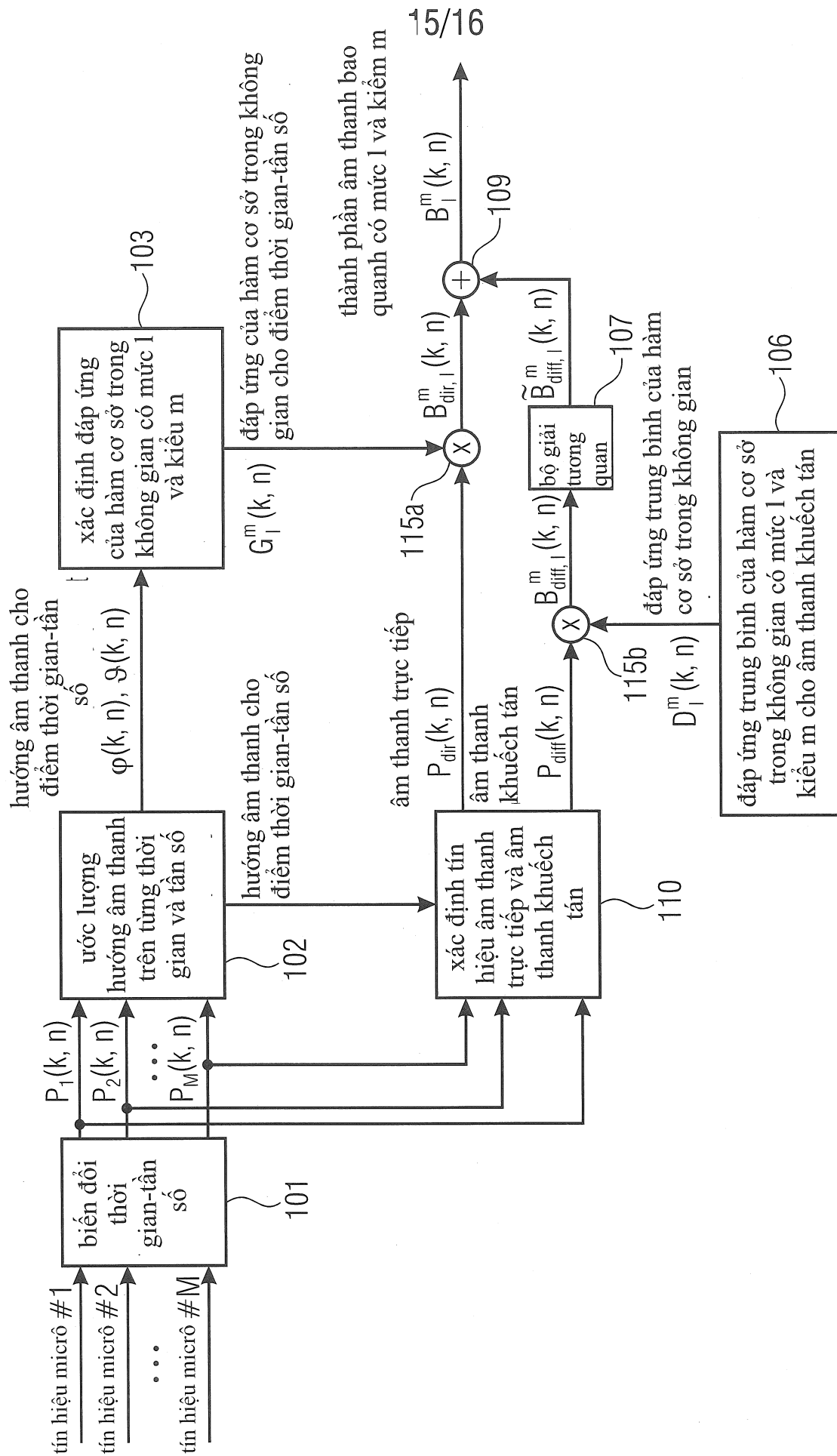


Fig. 7

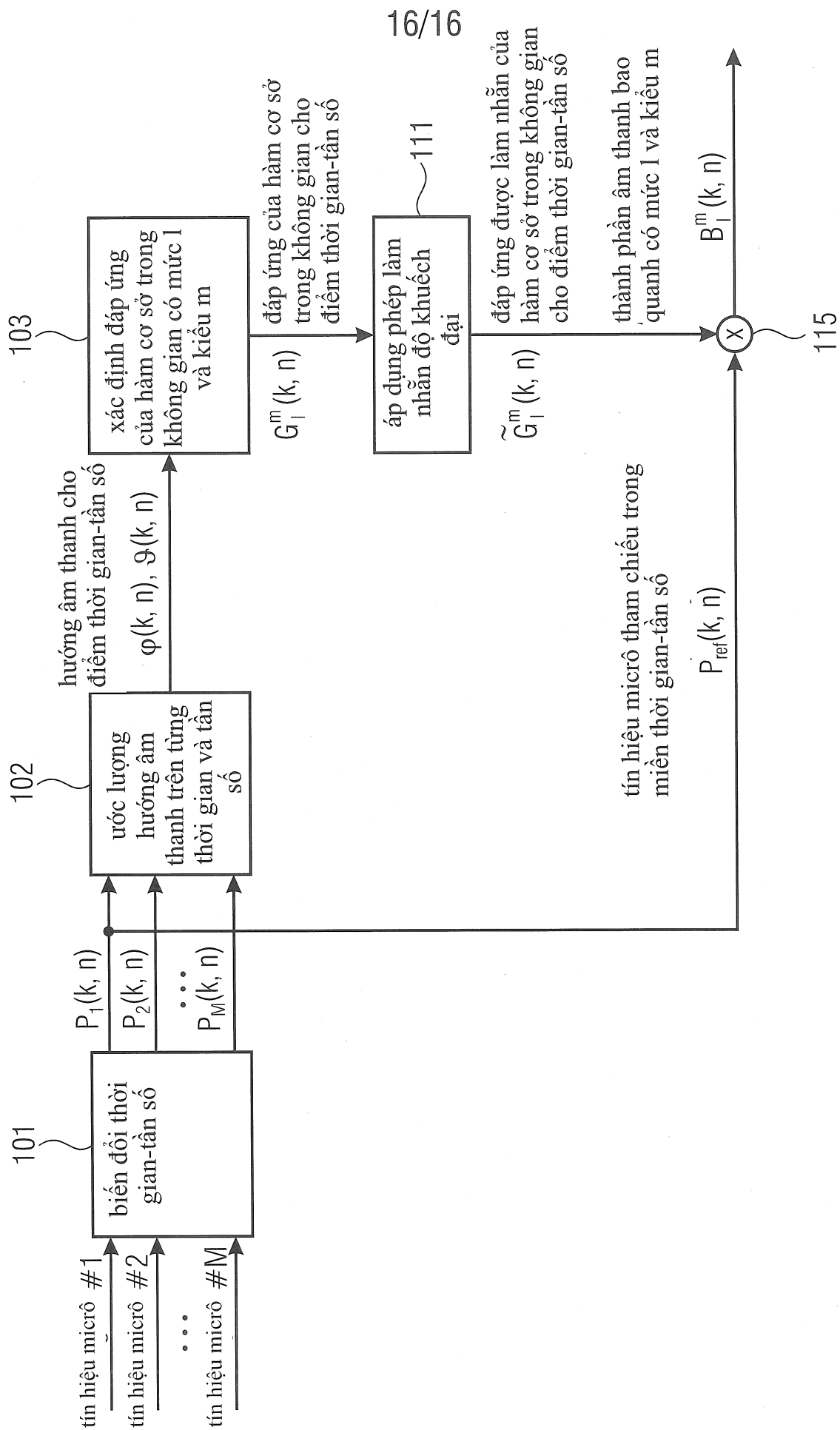


Fig. 8