



(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)



1-0052628

(51)^{2021.01} H04S 1/00; H04S 3/00

(13) B

(21) 1-2022-06582

(22) 12/03/2021

(86) PCT/EP2021/056358 12/03/2021

(87) WO 2021/180935 16/09/2021

(30) 20163159.5 13/03/2020 EP

(45) 27/10/2025 451

(43) 26/12/2022 417A

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

(72) HERRE, Jürgen (DE); ADAMI, Alexander (DE); ANEMÜLLER, Carlotta (DE).

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

(54) THIẾT BỊ VÀ PHƯƠNG PHÁP TỔNG HỢP NGUỒN ÂM MỞ RỘNG THEO
KHÔNG GIAN

(21) 1-2022-06582

(57) Sáng chế đề cập đến thiết bị và phương pháp tổng hợp nguồn âm mở rộng theo không gian. Thiết bị tổng hợp nguồn âm mở rộng theo không gian, thiết bị bao gồm: giao diện thông tin không gian (100) để nhận chỉ báo phạm vi không gian biểu thị phạm vi không gian giới hạn cho nguồn âm mở rộng theo không gian trong phạm vi không gian tối đa (600); bộ cung cấp thông tin tín hiệu điều khiển (200) để cung cấp một hoặc nhiều mục thông tin tín hiệu điều khiển phản hồi phạm vi không gian giới hạn; và bộ xử lý âm thanh (300) để xử lý tín hiệu âm thanh biểu diễn nguồn âm mở rộng theo không gian sử dụng một hoặc nhiều mục thông tin tín hiệu điều khiển.

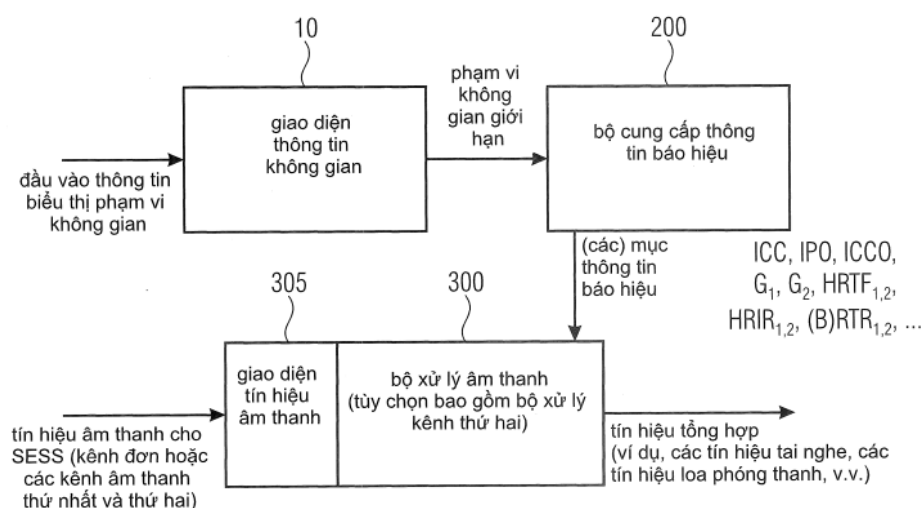


Fig. 1a

Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến việc xử lý tín hiệu âm thanh và, cụ thể, đến sự tái tạo một hoặc nhiều nguồn âm mở rộng theo không gian.

Tình trạng kỹ thuật của sáng chế

Đối với nhiều ứng dụng khác nhau, sự tái tạo nguồn âm qua một số loa phát thanh hoặc ống nghe được yêu cầu. Các ứng dụng này bao gồm các ứng dụng thực tế ảo, hỗn hợp hoặc bổ sung 6 mức tự do (6-Degrees-of-Freedom - 6DoF). Cách đơn giản nhất để tái tạo các nguồn âm qua các thiết lập như vậy là kết xuất chúng như các nguồn điểm. Tuy nhiên, khi nhắm đến tái tạo các nguồn âm vật lý với mức không gian nghe có thể bỏ qua, mô hình này là không đủ. Các ví dụ cho các nguồn âm như vậy là đàn piano lớn, giàn đồng ca hoặc thác nước, mà đều có “kích thước” nhất định.

Sự tái tạo thực tế của các nguồn âm với mức không gian trở thành mục đích của nhiều phương pháp tái tạo âm. Điều này bao gồm sự tái tạo lập thể, sử dụng các ống nghe, cũng như sự tái tạo thông thường, sử dụng các thiết lập loa phát thanh nằm trong khoảng từ 2 loa (“lập thể”) đến nhiều loa được sắp xếp trong mặt phẳng nằm ngang (“âm bao quanh”) và nhiều loa bao quanh người nghe trong cả ba chiều (“âm thanh 3D”). Sau đây, phần mô tả về các phương án hiện có được đưa ra. Các phương pháp khác nhau đã được nhóm lại thành các phương pháp cân nhắc chiều rộng nguồn lần lượt trong không gian 2D và 3D.

Các phương pháp được mô tả mà thuộc về việc kết xuất SESS trên bề mặt 2D đối diện với điểm nhìn của người nghe. Ví dụ, điều này có thể là ví dụ trong phạm vi góc phương vị nhất định tại góc nâng bằng 0 (giống như trong trường hợp âm lập thể/bao quanh thông thường) hoặc trong các phạm vi nhất định của góc phương vị và góc nâng (tương tự trường hợp trong âm thanh 3D hoặc thực tế ảo (Virtual Reality - VR) với 3 mức tự do (3-Degrees-of-Freedom - 3DoF) của sự di chuyển người dùng, tức là, việc quay đầu trong trục độ cao/lắc ngang/lăn).

Việc tăng chiều rộng rõ ràng của đối tượng âm thanh mà được quét giữa hai hoặc

nhiều hơn hai loa phát thanh (tạo ra cái được gọi là ảnh ảo hoặc nguồn ảo) có thể đạt được bằng cách giảm sự tương quan của các tín hiệu kênh tham gia [1, p.241-257].

Với sự tương quan giảm, sự lan nguồn ảo tăng, cho đến khi các giá trị tương quan gần bằng 0, nó bao phủ phạm vi giữa các loa phát thanh. Các phiên bản giải tương quan của tín hiệu nguồn thu được bằng các dẫn xuất và áp dụng các bộ lọc giải tương quan phù hợp. Lauridsen [2] đề xuất để cộng/trừ phiên bản trễ và định tỉ lệ thời gian của tín hiệu nguồn cho bản thân nó để thu được hai phiên bản giải tương quan của tín hiệu. Các phương pháp phức tạp hơn là ví dụ được đề xuất bởi Kendall [3]. Tác giả suy ra lặp lại sự giải tương quan được bắt cặp các bộ lọc thông qua tất cả dựa trên tổ hợp gồm các chuỗi số ngẫu nhiên. Faller và cộng sự đề xuất các bộ lọc giải tương quan phù hợp (“các bộ khuếch tán”) trong [4, 5]. Also, Zotter et al. [6] các cặp bộ lọc được suy ra mà phụ thuộc vào tần số hoặc các chênh lệch biên độ được sử dụng để đạt được sự mở rộng của nguồn ảo. Alary và cộng sự [7] đề xuất các bộ lọc giải tương quan dựa trên nhiễu mịn mà còn được tối ưu hóa bởi Schlecht và cộng sự. [8].

Ngoài việc giảm sự tương quan của các tín hiệu kênh tương ứng của nguồn ảo, chiều rộng nguồn có thể cũng tăng bằng cách tăng số lượng nguồn ảo đặc trưng cho đối tượng âm thanh. Trong tài liệu [9], chiều rộng nguồn được điều khiển bằng cách quét tín hiệu giống nhau đến các hướng (hơi) khác nhau. Phương pháp được đề xuất ban đầu để ổn định sự trải rộng nguồn ảo được nhận biết của các tín hiệu nguồn được quét VBAP [10] khi chúng được di chuyển trong ngữ cảnh âm. Điều này có lợi vì phụ thuộc vào hướng của nguồn, nguồn được kết xuất được tái tạo bởi hai hoặc nhiều hơn hai loa, mà có thể dẫn đến sự thay thế không cần của chiều rộng nguồn được nhận biết.

DirAC thế giới ảo [11] là sự mở rộng của phương pháp mã hóa âm thanh định hướng truyền thống (Directional Audio Coding - DirAC) [12] cho sự tổng hợp âm trong các thế giới ảo. Để kết xuất mức không gian, các thành phần âm định hướng của nguồn được quét ngẫu nhiên trong phạm vi nhất định quanh hướng ban đầu của nguồn, trong đó các hướng quét thay đổi với thời gian và tần số.

Phương pháp tương tự được theo đuổi trong tài liệu [13], trong đó mức không gian đạt được bằng cách phân bố ngẫu nhiên băng tần của tín hiệu âm vào các hướng không gian khác nhau. Đây là phương pháp nhắm đến việc sản xuất âm được phân bố theo

không gian và đường bao đến đều nhau từ tất cả các hướng hơn là điều khiển mức độ chính xác.

Verron và cộng sự đạt được mức không gian của nguồn bằng cách không sử dụng các tín hiệu tương quan được quét, mà bằng cách tổng hợp nhiều phiên bản rời rạc của tín hiệu nguồn, phân bố chúng đồng đều trên vòng tròn quanh người nghe, và trộn giữa chúng [14]. Số lượng và độ khuếch đại của các nguồn hoạt động đồng thời xác định mật độ của hiệu quả làm rộng. Phương pháp này được thực thi như sự mở rộng không gian đến bộ tổng hợp cho các âm môi trường.

Các phương pháp được mô tả có liên quan đến việc kết xuất các nguồn âm được mở rộng trong không gian 3D, tức là, theo cách về thể tích như được yêu cầu cho VR với 6DoF của sự di chuyển người dùng. 6 mức tự do này bao gồm sự quay đầu theo các trục độ cao/lắc ngang/lăn cộng thêm 3 hướng di chuyển tịnh tiến x/y/z.

Potard và cộng sự mở rộng quan điểm về mức nguồn như tham số một chiều của nguồn (tức là, chiều rộng của chúng giữa hai loa phát thanh) bằng cách tìm hiểu về cảm quan của các dạng nguồn [15]. Họ tạo ra nhiều nguồn điểm rời rạc bằng cách sử dụng công nghệ giải tương quan (thay đổi theo thời gian) đến tín hiệu nguồn ban đầu và sau đó đặt các nguồn rời rạc đến các vị trí không gian khác nhau và bằng cách này cho chúng mức ba chiều [16].

Trong AudioBIFS tiên tiến MPEG-4 [17], các đối tượng/hình dạng về thể tích (dạng vỏ, hình hộp, hình elip và hình trụ) có thể được thêm các nguồn âm được phân bố và giải tương quan đồng đều để gợi lên mức nguồn ba chiều.

Gần đây, Schlecht và cộng sự [18] đề xuất phương pháp mà chiều viên lõi của hình học SESS về phía vị trí người nghe, điều này cho phép kết xuất SESS tại bất kì vị trí liên quan đến người nghe. Tương tự với AudioBIFS tiên tiến MPEG-4, một số nguồn điểm giải tương quan khi đó được đặt trong sự chiếu này.

Để tăng và điều khiển mức nguồn sử dụng Ambisonics, Schmele và cộng sự [19] đề xuất hỗn hợp giảm thứ tự Ambisonics của tín hiệu đầu vào, mà vốn tăng chiều rộng nguồn rõ ràng, và phân bố các bản giải tương quan của tín hiệu nguồn quanh không gian nghe.

Phương pháp khác được giới thiệu bởi Zotter và cộng sự, trong đó, họ sử dụng nguyên lý được đề xuất trong tài liệu [6] (tức là, suy ra các cặp bộ lọc mà giới thiệu phụ thuộc vào tần số và sự chênh lệch cường độ để đạt được mức nguồn trong các thiết lập tái tạo lập thể) cho Ambisonics [20].

Bất lợi chung của các phương pháp dựa trên viết quét (ví dụ, các tài liệu [10, 9, 12, 11]) là sự phụ thuộc của chúng vào vị trí của người nghe. Ngay cả độ lệch nhỏ từ điểm ngọt gây ra ảnh không gian sẽ sụp đổ vào loa phát thanh gần nhất với người nghe. Điều này hạn chế đáng kể ứng dụng của chúng trong ngữ cảnh của VR và thực tế tăng cường (Augmented Reality - AR) trong đó người nghe được cho là di chuyển tự do xung quanh. Thêm vào đó, việc phân bố các ngăn thời gian - tần số trong các phương pháp trên cơ sở DirAC (ví dụ, tài liệu [12, 11]) luôn luôn không đảm bảo sự kết xuất phù hợp của mức không gian của các nguồn ảo. Hơn nữa, nó suy giảm đáng kể điển hình âm sắc của tín hiệu nguồn.

Sự giải tương quan của các tín hiệu nguồn thường đạt được bởi một trong số các phương pháp sau đây: i) suy ra các cặp với cường độ bổ sung (ví dụ, [2]), hoặc ii) sử dụng các bộ lọc thông qua tất cả với cường độ không đổi nhưng pha xáo trộn (ngẫu nhiên) (ví dụ, [3, 16]). Hơn nữa, sự mở rộng tín hiệu nguồn thu được bằng cách phân bố ngẫu nhiên các ngăn thời gian - tần số của tín hiệu nguồn (ví dụ, [13]).

Tất cả các phương pháp đến với quan hệ mật thiết của chúng: Việc lọc bù tín hiệu nguồn theo i) thường dẫn đến âm sắc được nhận biết thay thế của các tín hiệu giải tương quan. Trong khi việc lọc thông qua tất cả như trong ii) bảo toàn âm sắc của tín hiệu nguồn, pha xáo trộn phá vỡ các mối quan hệ pha ban đầu và đặc biệt đối với các tín hiệu chuyển tiếp gây ra sự phân tán mạnh mẽ và các thành phần lạ nhòe. Các ngăn thời gian - tần số phân bố theo không gian chứng minh là hiệu quả cho một số tín hiệu, nhưng cũng thay thế các âm sắc được nhận biết khác. Được thể hiện rằng phụ thuộc tín hiệu cao và dẫn đến các thành phần lạ mạnh mẽ cho các tín hiệu xung.

Việc làm giàu hình dạng thể tích với nhiều phiên bản giải tương quan của tín hiệu nguồn như được đề xuất trong AudioBIFS tiên tiến ([17, 15, 16]) giả định tính khả dụng của số lượng lớn bộ lọc mà sản xuất các tín hiệu đầu ra giải tương quan chung (điển hình, nhiều hơn mười nguồn điểm mỗi dạng thể tích được sử dụng). Tuy nhiên, tìm kiếm

các bộ lọc như vậy không phải nhiệm vụ bình thường và trở nên khó hơn vì cần đến nhiều bộ lọc như vậy. Nếu các tín hiệu nguồn không được giải tương quan đầy đủ và người nghe di chuyển quanh hình dạng, ví dụ, trong ngữ cảnh VR, các khoảng cách nguồn cá nhân đến người nghe tương ứng với các độ trễ khác nhau của tín hiệu nguồn. Siêu vị trí của chúng tại tai của người nghe sẽ dẫn đến việc lọc lược phụ thuộc vào vị trí, có tiềm năng dẫn đến sự tô màu không ổn định hiện nhiều của tín hiệu nguồn. Hơn nữa, ứng dụng của nhiều bộ lọc giải tương quan có nghĩa nhiều độ phức tạp tính toán.

Các cân nhắc tương tự sử dụng cho phương pháp được mô tả trong tài liệu [18], trong đó số lượng các nguồn điểm được giải tương quan được đặt trên phần chiếu viên lồi của hình học SESS. Trong khi các tác giả không đề cập đến bất cứ thứ gì về số lượng yêu cầu của nguồn hỗ trợ giải tương quan, có tiềm năng là số lượng lớn cần để đạt được mức nguồn thuyết phục. Điều này dẫn đến các trở ngại đã được thảo luận trong đoạn trước.

Việc điều khiển chiều rộng nguồn sử dụng công nghệ trên cơ sở Ambisonics được mô tả trong tài liệu [19] bằng cách hạ thấp thứ tự Ambisonics được thể hiện để có hiệu quả nghe chỉ cho sự chuyển tiếp từ thứ hai đến thứ nhất hoặc thứ không. Các chuyển tiếp này không chỉ được tiếp nhận như sự mở rộng nguồn mà cũng thường xuyên như sự di chuyển của nguồn ảo. Trong khi thêm các phiên bản giải tương quan của tín hiệu nguồn có thể giúp ổn định cảm quan của chiều rộng nguồn rõ ràng, cũng dẫn đến các hiệu quả bộ lọc lược, mà thay thế âm sắc của nguồn ảo.

Tài liệu tham khảo

- [1] J. Blauert, *Spatial Hearing: Psychophysics of Human Sound Localization*, 3rd ed. Cambridge, Mass: MIT Press, 2001.
- [2] H. Lauridsen, "Experiments Concerning Different Kinds of Room-Acoustics Recording," *Ingenioren*, 1954.
- [3] G. Kendall, "The Decorrelation of Audio Signals and Its Impact on Spatial Imagery," *Computer Music Journal*, vol. 19, no. 4, pp. 71–87, 1995.

- [4] C. Faller and F. Baumgarte, "Binaural cue coding-Part II: Schemes and applications," *IEEE Transactions on Speech and Audio Processing*, vol. 11, no. 6, pp. 520–531, Nov. 2003.
- [5] F. Baumgarte and C. Faller, "Binaural cue coding-Part I: Psychoacoustic fundamentals and design principles," *IEEE Transactions on Speech and Audio Processing*, vol. 11, no. 6, pp. 509–519, Nov. 2003.
- [6] F. Zotter and M. Frank, "Efficient Phantom Source Widening," *Archives of Acoustics*, vol. 38, pp. 27–37, Mar. 2013.
- [7] B. Alary, A. Politis, and V. V'alima'ki, "Velvet-noise decorrelator," *Proc. DAFx-17*, Edinburgh, UK, pp. 405–411, 2017.
- [8] S. Schlecht, B. Alary, V. V'alima'ki, and E. Habets, "Optimized velvet-noise decorrelator," Sep. 2018.
- [9] V. Pulkki, "Uniform spreading of amplitude panned virtual sources," *Proceedings of the 1999 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics. WASPAA'99* (Cat. No.99TH8452), pp. 187–190, 1999.
- [10] —, "Virtual Sound Source Positioning Using Vector Base Amplitude Panning," *Journal of the Audio Engineering Society*, vol. 45, no. 6, pp. 456–466, Jun. 1997.
- [11] V. Pulkki, M.-V. Laitinen, and C. Erkut, "Efficient Spatial Sound Synthesis for Virtual Worlds," *Audio Engineering Society*, Feb. 2009.
- [12] V. Pulkki, "Spatial Sound Reproduction with Directional Audio Coding," *Journal of the Audio Engineering Society*, vol. 55, no. 6, pp. 503–516, Jun. 2007.
- [13] T. Pihlajam'aki, O. Santala, and V. Pulkki, "Synthesis of Spatially Extended Virtual Source with Time-Frequency Decomposition of Mono Signals," *Journal of the Audio Engineering Society*, vol. 62, no. 7/8, pp. 467–484, Aug. 2014.
- [14] C. Verron, M. Aramaki, R. Kronland-Martinet, and G. Pallone, "A 3-D Immersive Synthesizer for Environmental Sounds," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 18, pp. 1550–1561, Sep. 2010.

[15] G. Potard and I. Burnett, “A study on sound source apparent shape and wideness,” pp. 6–9, Aug. 2003.

[16] —, “Decorrelation techniques for the rendering of apparent sound source width in 3D audio displays,” Jan. 2004, pp. 280–208.

[17] J. Schmidt and E. F. Schroeder, “New and Advanced Features for Audio Presentation in the MPEG-4 Standard.” Audio Engineering Society, May 2004.

[18] S. Schlecht, A. Adami, E. Habets, and J. Herre, “Apparatus and Method for Reproducing a Spatially Extended Sound Source or Apparatus and Method for Generating a Bitstream from a Spatially Extended Sound Source,” Patent Application PCT/EP2019/085 733.

[19] T. Schmele and U. Sayin, “Controlling the Apparent Source Size in Ambisonics Using Decorrelation Filters.” Audio Engineering Society, Jul. 2018.

[20] F. Zotter, M. Frank, M. Kronlachner, and J.-W. Choi, “Efficient Phantom Source Widening and Diffuseness in Ambisonics,” Jan. 2014.

[21] C. Borß, “An Improved Parametric Model for the Design of Virtual Acoustics and its Applications,” Ph.D. dissertation, Ruhr-Universität Bochum, Jan. 2011.

Bản chất kỹ thuật của sáng chế

Mục đích của sáng chế là cung cấp giải pháp được cải thiện về tổng hợp nguồn âm mở rộng theo không gian.

Mục đích này đạt được bởi thiết bị để tổng hợp nguồn âm mở rộng theo không gian theo điểm 1, phương pháp tổng hợp nguồn âm mở rộng theo không gian theo điểm 23, hoặc chương trình máy tính theo điểm 24.

Sáng chế dựa trên phát hiện rằng sự tái tạo của nguồn âm mở rộng theo không gian có thể đạt được một cách hiệu quả bằng cách sử dụng sự chỉ báo phạm vi không gian biểu thị phạm vi đích không gian giới hạn cho nguồn âm mở rộng theo không gian trong phạm vi không gian tối đa. Dựa trên bộ chỉ báo phạm vi không gian và, cụ thể, dựa trên phạm vi không gian giới hạn, một hoặc nhiều mục thông tin tín hiệu điều khiển được

cung cấp và, bộ xử lý xử lý tín hiệu âm thanh biểu diễn nguồn âm mở rộng theo không gian sử dụng một hoặc nhiều mục tín hiệu điều khiển.

Quy trình này đạt được sự xử lý hiệu quả cao của nguồn âm mở rộng theo không gian. Đối với sự tái tạo ống nghe, ví dụ, chỉ hai kênh lập thể, tức là, kênh lập thể bên trái hoặc kênh lập thể bên phải, được yêu cầu. Đối với sự tái tạo lập thể, chỉ hai kênh được yêu cầu. Do đó, ngược với việc tổng hợp nguồn âm mở rộng theo không gian sử dụng số lượng đáng kể các nguồn âm hình cầu điền đầy thể tích thực tế hoặc diện tích của nguồn âm mở rộng theo không gian hoặc, nhìn chung, điền đầy phạm vi không gian giới hạn do sự dịch chuyển riêng lẻ, điều này không được yêu cầu phù hợp với sáng chế, vì nguồn âm mở rộng theo không gian không được kết xuất sử dụng số lượng đáng kể các nguồn âm riêng lẻ được đặt trong thể tích, nhưng nguồn âm mở rộng theo không gian được kết xuất sử dụng hai hoặc, chắc chắn, ba kênh mà có các tín hiệu điều khiển nhất định với nhau mà có thể thu được, khi số lượng lớn các nguồn âm riêng lẻ hình cầu được nhận biết tại hai hoặc ba địa điểm.

Do đó, trái với các phương pháp khác nhau đang có và nhắm đến tái tạo theo hiện thực các nguồn âm mở rộng theo không gian (spatially extended sound sources - SESS), trong đó các phương pháp hiện có này về cơ bản yêu cầu số lượng lớn các tín hiệu đầu vào được giải tương quan, sáng chế đi theo hướng khác. Việc tạo ra các tín hiệu đầu vào được giải tương quan như vậy có thể tương đối tốn kém về độ phức tạp tính toán. Các phương pháp hiện có trước đó có thể cũng làm giảm sút chất lượng của âm qua sự chênh lệch âm sắc hoặc sự nhòe âm sắc. Và tìm ra số lượng lớn các bộ giải tương quan trực giao chung nhìn chung không dễ để giải quyết vấn đề. Do đó, các quy trình trước đó luôn dẫn đến sự đánh đổi giữa độ giải tương quan chung và sự suy giảm tín hiệu được đưa vào ngoài các tài nguyên tính toán cao được yêu cầu.

Trái ngược với điều đó, sáng chế tổng hợp số lượng thấp các kênh ví dụ như kênh bên trái thu được và kênh bên phải thu được cho nguồn âm được mở rộng theo không gian sử dụng chỉ hai tín hiệu đầu vào được giải tương quan. Tốt hơn là, kết quả tổng hợp là tín hiệu tại bên trái và bên phải cho sự tái tạo tai nghe. Tuy nhiên, đối với các loại ngữ cảnh tái tạo khác, như sự kết xuất loa phát thanh hoặc sự kết xuất loa phát thanh tái tạo nhiều xuyên âm, sáng chế có thể cũng được áp dụng. Thay vì đặt nhiều tín hiệu âm giải

tương quan khác nhau tại các vị trí khác nhau trong thể tích cho nguồn âm được mở rộng theo không gian, tín hiệu âm lượng cho nguồn âm được mở rộng theo không gian bao gồm một hoặc nhiều kênh được xử lý sử dụng một hoặc nhiều mục thông tin tín hiệu điều khiển được suy ra từ bộ cung cấp thông tin tín hiệu điều khiển phản hồi chỉ báo phạm vi không gian giới hạn được nhận từ giao diện thông tin không gian.

Các phương án được ưu tiên nhằm đến tổng hợp hiệu quả SESS cho sự tái tạo ống nghe. Sự tổng hợp nhờ đó dựa trên mô hình nằm dưới mô tả SESS bởi số vô hạn (lý tưởng) của các nguồn điểm giải tương quan được đặt cách dày đặc được phân bố qua toàn bộ phạm vi mức nguồn. Phạm vi mức nguồn mong muốn có thể được thể hiện như hàm của góc phương vị và góc nâng, mà khiến phương pháp theo sáng chế có thể được áp dụng cho các ứng dụng 3DoF. Tuy nhiên, sự mở rộng của các ứng dụng 6DoF là khả thi, bằng cách liên tục chiếu hình học SESS theo hướng về phía vị trí người nghe hiện thời như được mô tả trong tài liệu [18]. Như ví dụ cụ thể, mức nguồn mong muốn được mô tả sau đây về phạm vi góc phương vị và góc nâng.

Các phương án được ưu tiên khác phụ thuộc vào việc sử dụng giá trị hiệu chỉnh liên kênh như thông tin tín hiệu điều khiển hoặc còn sử dụng sự chênh lệch pha liên kênh, sự chênh lệch thời gian liên kênh, sự chênh lệch liên mức và thừa số khuếch đại hoặc cặp mục thông tin thừa số độ khuếch đại thứ nhất và thứ hai. Do đó, các mức tuyệt đối của các kênh có thể được thiết lập bởi hai thừa số khuếch đại hoặc thừa số khuếch đại đơn và sự chênh lệch mức liên kênh, bất kỳ hàm bộ lọc âm thanh thay vì các mục tín hiệu điều khiển thực tế hoặc, ngoài các mục tín hiệu điều khiển thực tế có thể cũng được cung cấp như các mục thông tin tín hiệu điều khiển từ bộ cung cấp thông tin tín hiệu điều khiển đến bộ xử lý âm thanh để bộ xử lý âm thanh vận hành bằng cách tổng hợp, ví dụ, hai kênh đầu ra như hai kênh đầu ra lập thể hoặc cặp kênh đầu ra bên trái và bên phải sử dụng ứng dụng của mục tín hiệu điều khiển thực tế, tùy chọn, lọc sử dụng hàm chuyển liên quan phân đầu cho mỗi kênh như mục thông tin tín hiệu điều khiển hoặc sử dụng hàm phản hồi xung liên quan phân đầu như mục thông tin tín hiệu điều khiển hoặc sử dụng hàm phản hồi xung phòng lập thể hoặc (không lập thể) như mục thông tin tín hiệu điều khiển. Nhìn chung, chỉ thiết lập mục tín hiệu điều khiển đơn có thể đủ, nhưng trong các phương án phức tạp hơn, nhiều hơn một mục tín hiệu điều khiển có hoặc không có các bộ lọc sẽ được áp đặt trên các tín hiệu âm thanh bởi bộ xử lý âm thanh.

Do đó, trong phương án, khi giá trị tương quan liên kênh được cung cấp như mục thông tin tín hiệu điều khiển, và trong đó tín hiệu âm thanh bao gồm kênh âm thanh thứ nhất và kênh âm thanh thứ hai cho nguồn âm mở rộng theo không gian, hoặc trong đó tín hiệu âm thanh bao gồm kênh âm thanh thứ nhất và kênh âm thanh thứ hai được suy ra từ kênh âm thanh thứ nhất bởi việc thực thi bộ xử lý kênh thứ hai, ví dụ, xử lý giải tương quan hoặc xử lý mạng neuron hoặc bất kỳ sự xử lý để suy ra tín hiệu mà có thể được xem như tín hiệu được giải tương quan, bộ xử lý âm thanh được tạo cấu hình để áp đặt sự tương quan giữa kênh âm thanh thứ nhất và kênh âm thanh thứ hai sử dụng giá trị tương quan liên kênh và thêm vào hoặc trước hoặc sau sự xử lý này, các hàm bộ lọc âm thanh có thể được áp dụng để thu được hai kênh đầu ra mà có sự tương quan liên kênh đích được biểu thị bởi giá trị tương quan liên kênh và ngoài ra có các mối quan hệ được biểu thị bởi các hàm bộ lọc riêng lẻ hoặc các mục tín hiệu điều khiển thực tế khác.

Bộ cung cấp thông tin tín hiệu điều khiển có thể được thực thi bằng tìm kiếm bao gồm bộ nhớ hoặc mô hình hỗn hợp Gausơ hoặc máy vectơ hỗ trợ hoặc bảng mã vectơ, khớp hàm đa chiều hoặc một số thiết bị khác cung cấp một cách hiệu quả các tín hiệu điều khiển yêu cầu phản hồi sự chỉ báo phạm vi không gian.

Có thể, ví dụ, trong ví dụ bảng tìm kiếm, hoặc trong bảng mã vectơ hoặc ví dụ khớp hàm đa chiều hoặc trong ví dụ GMM hoặc SVM, để cung cấp các hiểu biết trước để nhiệm vụ chính của giao diện thông tin không gian là để thực sự tìm phạm vi không gian ứng viên được khớp mà khớp, trong số tất cả các phạm vi không gian ứng viên có sẵn, tốt nhất có thể với thông tin chỉ báo phạm vi không gian đầu vào. Thông tin này có thể được cung cấp trực tiếp qua người dùng hoặc có thể được tính toán sử dụng thông tin trên nguồn âm mở rộng theo không gian và vị trí người nghe hoặc sự định hướng người nghe (như ví dụ, được xác định bởi bộ truy vết phần đầu hoặc thiết bị như vậy) bởi một vài loại phép tính sự chiếu. Hình học hoặc kích thước của đối tượng và khoảng cách giữa người nghe và đối tượng có thể đủ để suy ra góc mở, và, do đó, phạm vi không gian giới hạn cho sự kết xuất của nguồn âm. Trong các phương án khác, giao diện thông tin không gian chỉ là đầu vào để nhận phạm vi không gian giới hạn và để chuyển tiếp dữ liệu này đến bộ cung cấp thông tin tín hiệu điều khiển, khi dữ liệu được nhận bởi giao diện đã ở định dạng có thể sử dụng được bởi bộ cung cấp thông tin tín hiệu điều khiển.

Mô tả vắn tắt các hình vẽ

Sau đây, các phương án ưu tiên của sáng chế được thảo luận với việc đối chiếu với các hình vẽ đi kèm trong đó:

Fig.1a minh họa sự thực thi được ưu tiên của thiết bị tổng hợp nguồn âm mở rộng theo không gian;

Fig.1b minh họa phương án khác của bộ xử lý âm thanh và bộ cung cấp thông tin tín hiệu điều khiển;

Fig.2 minh họa phương án ưu tiên của bộ xử lý kênh thứ hai nằm trong bộ xử lý âm thanh của Fig.1a;

Fig.3 minh họa sự thực thi ưu tiên của thiết bị thực hiện sự điều chỉnh ICC;

Fig.4 minh họa phương án ưu tiên của sáng chế trong đó các mục thông tin tín hiệu điều khiển dựa vào các mục và các bộ lọc tín hiệu điều khiển thực tế;

Fig.5 minh họa phương án khác dựa thêm vào các bộ lọc và mục tương quan liên kênh;

Fig.6 minh họa sơ đồ phân giải lược minh họa phạm vi không gian tối đa trong tình huống hai chiều hoặc ba chiều và các phần riêng hoặc các phạm vi không gian giới hạn có thể, ví dụ, được sử dụng như các phân ứng viên;

Fig.7 minh họa sự thực thi của giao diện thông tin không gian;

Fig.8 minh họa sự thực thi khác của giao diện thông tin không gian phụ thuộc vào các quy trình tính toán sự chiếu;

Fig.9a và Fig.9b minh họa các phương án để thực hiện phép tính sự chiếu và sự xác định phạm vi không gian;

Fig.10 minh họa sự thực thi được ưu tiên khác của giao diện thông tin không gian;

Fig.11 minh họa sự thực thi khác của giao diện thông tin không gian liên quan đến sự thực thi bộ giải mã;

Fig.12 minh họa phép tính của phạm vi không gian giới hạn cho nguồn âm mở rộng theo không gian hình cầu;

Fig.13 minh họa các phép tính khác của các phạm vi không gian giới hạn cho nguồn âm mở rộng theo không gian hình elip;

Fig.14 minh họa phép tính khác của phạm vi không gian giới hạn cho nguồn âm mở rộng theo không gian dạng đường;

Fig.15 minh họa sự minh họa khác cho phép tính của phạm vi không gian giới hạn cho nguồn âm mở rộng theo không gian hình hộp phẳng;

Fig.16 minh họa ví dụ khác để tính phạm vi không gian giới hạn cho nguồn âm mở rộng theo không gian hình cầu;

Fig.17 minh họa nguồn âm mở rộng theo không gian dạng đàn piano với dạng elip theo tham số gần đúng; và

Fig.18 minh họa các điểm để định nghĩa phạm vi không gian giới hạn cho sự kết xuất của nguồn âm mở rộng theo không gian dạng đàn piano.

Mô tả chi tiết sáng chế

Fig.1a minh họa sự thực thi được ưu tiên của thiết bị tổng hợp nguồn âm mở rộng theo không gian. Thiết bị bao gồm giao diện thông tin không gian 10 để nhận đầu vào thông tin chỉ báo phạm vi không gian biểu thị phạm vi không gian giới hạn cho nguồn âm mở rộng theo không gian trong phạm vi không gian tối đa. Phạm vi không gian giới hạn được nhập vào bộ cung cấp thông tin tín hiệu điều khiển 200 được tạo cấu hình để cung cấp một hoặc nhiều mục thông tin tín hiệu điều khiển phản hồi phạm vi không gian giới hạn được cho bởi giao diện thông tin không gian 10. Mục thông tin tín hiệu điều khiển hoặc một số mục thông tin tín hiệu điều khiển được cung cấp đến bộ xử lý âm thanh 300 được tạo cấu hình để xử lý tín hiệu âm thanh biểu diễn nguồn âm mở rộng theo không gian sử dụng một hoặc nhiều mục thông tin tín hiệu điều khiển được cung cấp bởi bộ cung cấp thông tin tín hiệu điều khiển 200. Tín hiệu âm thanh cho nguồn âm mở rộng theo không gian (spatially extended sound source - SESS) có thể là kênh đơn hoặc có thể là kênh âm thanh thứ nhất và kênh âm thanh thứ hai hoặc có thể là nhiều hơn hai kênh âm thanh. Tuy nhiên, dành cho mục đích có tải xử lý thấp, số lượng nhỏ các kênh cho nguồn âm mở rộng theo không gian hoặc, đối với tín hiệu âm thanh biểu diễn nguồn âm mở rộng theo không gian được ưu tiên. Tín hiệu âm thanh được nhập

vào giao diện tín hiệu âm thanh 305 của bộ xử lý âm thanh 300 và bộ xử lý âm thanh 300 xử lý tín hiệu âm thanh đầu vào được nhận bởi giao diện tín hiệu âm thanh hoặc, khi số lượng các kênh âm thanh đầu vào nhỏ hơn được yêu cầu như chỉ một, bộ xử lý âm thanh bao gồm bộ xử lý kênh thứ hai 310 được minh họa trên Fig.2 bao gồm, ví dụ, bộ giải tương quan để tạo ra kênh âm thanh thứ hai S_2 được giải tương quan từ kênh âm thanh thứ nhất S mà cũng được minh họa trên Fig.2 như S_1 . Các mục thông tin tín hiệu điều khiển có thể là các mục tín hiệu điều khiển thực tế như các mục tương quan liên kênh, các mục chênh lệch pha liên kênh, các mục chênh lệch mức liên kênh và độ khuếch đại, các mục thừa số khuếch đại G_1 , G_2 , cùng biểu diễn sự chênh lệch mức liên kênh và/hoặc biên độ tuyệt đối hoặc các mức công suất hoặc năng lượng, ví dụ, hoặc các mục thông tin tín hiệu điều khiển có thể cũng là các hàm bộ lọc thực tế như các hàm chuyển liên quan đến phần đầu với số lượng như yêu cầu bởi số lượng thực tế các kênh đầu ra sẽ được tổng hợp trong tín hiệu tổng hợp. Do đó, khi tín hiệu tổng hợp là có hai kênh như hai kênh lập thể hoặc hai kênh loa phát thanh, một hàm chuyển liên quan đến phần đầu cho mỗi kênh được yêu cầu. Thay vì hàm chuyển liên quan đến phần đầu, các hàm phản hồi xung liên quan đến phần đầu (head related impulse response functions - HRIR) hoặc các hàm phản hồi xung phòng lập thể hoặc không lập thể (B)RIR là cần thiết. Như được minh họa trên Fig.1a, một hàm chuyển như vậy được yêu cầu cho mỗi kênh và Fig.1a minh họa sự thực thi của có hai kênh để các chỉ số biểu thị “1” và “2”.

Trong phương án, bộ cung cấp thông tin tín hiệu điều khiển 200 được tạo cấu hình để cung cấp, như mục thông tin tín hiệu điều khiển, giá trị tương quan liên kênh. Bộ xử lý âm thanh 300 được tạo cấu hình để nhận thực tế, qua giao diện tín hiệu âm thanh 305, kênh âm thanh thứ nhất và kênh âm thanh thứ hai. Tuy nhiên, khi giao diện tín hiệu âm thanh 305 chỉ nhận kênh đơn, bộ xử lý kênh thứ hai được cung cấp tùy chọn tạo ra, ví dụ, bằng quy trình trên Fig.2, kênh âm thanh thứ hai. Bộ xử lý âm thanh thực hiện sự xử lý tương quan để áp đặt sự tương quan giữa kênh âm thanh thứ nhất và kênh âm thanh thứ hai sử dụng giá trị tương quan liên kênh.

Ngoài ra, hoặc cách khác, mục thông tin tín hiệu điều khiển khác có thể được cung cấp như mục chênh lệch pha liên kênh, mục chênh lệch thời gian liên kênh, sự chênh lệch mức liên kênh và mục độ khuếch đại, và mục thông tin thừa số khuếch đại thứ nhất và mục thông tin thừa số khuếch đại thứ hai. Các mục có thể cũng là các giá trị tương

quan liên tai (IACC), tức là, các giá trị tương quan liên kênh cụ thể hơn, hoặc các mục chênh lệch pha liên tai (interaural phase difference - IAPD), tức là, nhiều giá trị chênh lệch pha liên kênh cụ thể hơn.

Trong phương án được ưu tiên, sự tương quan được áp đặt bởi bộ xử lý âm thanh 300 phản hồi mục thông tin tín hiệu điều khiển tương quan, trước các sự điều chỉnh ICPD, ICTD hoặc ICLD được thực hiện hoặc, trước, HRTF hoặc các sự xử lý hàm bộ lọc chuyển khác được thực hiện. Tuy nhiên, như trường hợp có thể, thứ tự có thể được thiết lập khác nhau.

Trong phương án được ưu tiên, bộ xử lý âm thanh bao gồm bộ nhớ để lưu trữ thông tin trên các mục thông tin tín hiệu điều khiển khác nhau liên quan đến các chỉ báo phạm vi không gian khác nhau. Trong trường hợp này, bộ cung cấp thông tin tín hiệu điều khiển còn bao gồm giao diện đầu ra để truy hồi, từ bộ nhớ, một hoặc nhiều thông tin tín hiệu điều khiển được kết hợp với đầu vào chỉ báo phạm vi không gian vào bộ nhớ tương ứng. Bảng tìm kiếm 210 như vậy, ví dụ, được minh họa trên Fig.1b, Fig.4 hoặc Fig.5, trong đó bảng tìm kiếm bao gồm bộ nhớ và giao diện đầu ra để xuất ra các mục thông tin tín hiệu điều khiển tương ứng. Cụ thể, bộ nhớ có thể không chỉ chứa IACC, IAPD hoặc các giá trị G_l và G_r như được minh họa trên Fig.1b, mà bộ nhớ trong bảng tìm kiếm có thể cũng lưu trữ các hàm như được minh họa trên khối 220 của Fig.4 và Fig.5 được biểu thị như “lựa chọn HRTF”. Trong phương án này, mặc dù được minh họa riêng lẻ trên Fig.4 và Fig.5, các khối 210, 220 có thể bao gồm bộ nhớ giống nhau trong đó, khi kết hợp với phạm vi không gian tương ứng được biểu thị như các góc phương vị và các góc nâng, các mục thông tin tín hiệu điều khiển tương ứng như IACC và, tùy chọn, IAPD và các hàm chuyển cho các bộ lọc như $HRTF_l$ cho kênh đầu ra bên trái và $HRTF_r$ cho kênh đầu ra bên phải được lưu trữ, trong đó các kênh đầu ra bên trái và bên phải được biểu thị như S_l và S_r trên Fig.4 hoặc Fig.5 hoặc Fig.1b.

Bộ nhớ được sử dụng bởi bảng tìm kiếm 210 hoặc khối hàm lựa chọn 220 có thể cũng sử dụng thiết bị lưu trữ trong đó, dựa trên các mã phần hoặc các góc phần hoặc các phạm vi góc phần nhất định, các tham số tương ứng có sẵn. Cách khác, bộ nhớ có thể lưu trữ bảng mã vectơ, hoặc chương trình khớp hàm đa chiều hoặc mô hình hỗn hợp

Gauxơ (Gaussian Mixture Model - GMM) hoặc máy véctơ hỗ trợ (Support Vector Machine - SVM) như trường hợp có thể.

Cho phạm vi mức nguồn mong muốn, SESS được tổng hợp sử dụng hai tín hiệu đầu vào được giải tương quan. Các tín hiệu đầu vào này được xử lý theo cách mà các tín hiệu điều khiển thính giác quan trọng về cảm quan được tái tạo một cách chính xác. Điều này bao gồm các tín hiệu điều khiển liên tai sau đây: Tương quan chéo liên tai (Interaural Cross Correlation - IACC), sự chênh lệch pha liên tai (Interaural Phase Differences - IAPD)¹ và sự chênh lệch mức liên tai (Interaural Level Differences - IALD). Bên cạnh đó, các tín hiệu điều khiển phổ đơn âm được tái tạo. Những điều này chủ yếu quan trọng với sự xác định vị trí nguồn âm trong mặt phẳng dọc. Trong khi IAPD và IALD chủ yếu quan trọng cho mục đích xác định vị trí, IACC được biết đến là tín hiệu điều khiển quyết định cho cảm quan chiều rộng nguồn trong mặt phẳng ngang. Trong thời gian chạy, các giá trị đích của các tín hiệu điều khiển này được truy hồi từ bộ nhớ được tính toán trước. Sau đây, bảng tìm kiếm được sử dụng cho mục đích này. Tuy nhiên, mỗi ý nghĩa khác của dữ liệu đa chiều lưu trữ, ví dụ, bảng mã véctơ hoặc khớp hàm đa chiều, có thể được sử dụng. Ngoài phạm vi mức nguồn được cân nhắc, tất cả các tín hiệu điều khiển phụ thuộc chỉ vào tập hợp dữ liệu hàm chuyển liên quan đến phần đầu (Head-Related Transfer Function - HRTF) được sử dụng. Sau đó, đạo hàm các tín hiệu điều khiển nghe khác nhau được đưa ra.

Trên Fig.1b, sơ đồ khối chung của phương pháp được thể hiện. $[\Phi_1, \Phi_2]$ mô tả mức nguồn mong muốn về phạm vi góc phương vị. $[\theta_1, \theta_2]$ là mức nguồn mong muốn về phạm vi góc nâng. $S_1(\omega)$ và $S_2(\omega)$ biểu thị hai tín hiệu đầu vào, với ω mô tả chỉ số tần số. Đối với $S_1(\omega)$ và $S_2(\omega)$ phương trình sau đây giữ:

$$\mathbb{E}\{S_1(\omega) \cdot S_2^*(\omega)\} = 0 \quad (1)$$

Ngoài ra, cả tín hiệu đầu vào được yêu cầu có mật độ phổ công suất giống nhau. Như sự thay thế, có thể chỉ đưa một tín hiệu đầu vào, $S(\omega)$. Tín hiệu đầu vào thứ hai được tạo ra bên trong sử dụng bộ giải tương quan như được thể hiện trên Fig.2. Cho

$S_l(\omega)$ và $S_r(\omega)$, nguồn âm được mở rộng được tổng hợp bằng cách điều chỉnh liên tiếp sự nhất quán liên kênh (Inter-Channel Coherence - ICC), các chênh lệch pha liên kênh (Inter-Channel Phase Differences - ICPD) và các chênh lệch mức liên kênh (Inter-Channel Level Differences - ICLD) để kết hợp tín hiệu liên tai tương ứng. Số lượng cần cho các bước xử lý này được đọc từ bảng tìm kiếm được tính toán trước. Các tín hiệu kênh bên trái và bên phải thu được, $S_l(\omega)$ và $S_r(\omega)$ có thể được phát lại qua các ống nghe và giống với SESS. Nên lưu ý rằng sự điều chỉnh ICC phải được thực hiện đầu tiên, tuy nhiên, các khối điều chỉnh ICPD và ICLD có thể được thay đổi lẫn nhau. Thay vì IAPD, các chênh lệch thời gian liên tai (Interaural Time Differences - IATD) tương ứng có thể cũng được tái tạo. Tuy nhiên, sau đây, chỉ IAPD được xem xét thêm.

Trong khối điều chỉnh ICC, sự tương quan chéo giữa cả hai tín hiệu đầu vào được điều chỉnh đến giá trị mong muốn $|IACC(\omega)|$ sử dụng các công thức sau đây [21]:

$$\hat{S}_1(\omega) = H_\alpha(\omega) \cdot S_1(\omega) + H_\beta(\omega) \cdot S_2(\omega), \quad (2)$$

$$\hat{S}_2(\omega) = H_\alpha(\omega) \cdot S_2(\omega) + H_\beta(\omega) \cdot S_1(\omega), \quad (3)$$

$$H_\beta(\omega) = H_\alpha(\omega) \sqrt{\frac{1}{2} (1 - \sqrt{1 - |IACC(\omega)|^2})}, \quad (4)$$

$$H_\alpha(\omega) = \sqrt{1 - H_\beta^2(\omega)}. \quad (5)$$

Áp dụng các kết quả công thức này trong sự tương quan chéo, miễn là các tín hiệu đầu vào $S_1(\omega)$ và $S_2(\omega)$ được giải tương quan đầy đủ. Ngoài ra, mật độ phổ công suất của chúng cần giống nhau. Sơ đồ khối tương ứng được thể hiện trên Fig.3.

Khối điều chỉnh ICPD được mô tả bởi các công thức sau đây:

$$\hat{S}'_1(\omega) = e^{j \cdot IAPD(\omega)} \cdot \hat{S}_1(\omega), \quad (6)$$

$$\hat{S}'_2(\omega) = \hat{S}_2(\omega). \quad (7)$$

Sau đây, sự điều chỉnh ICLD được thực hiện như sau:

$$S_l(\omega) = G_l(\omega) \cdot \hat{S}'_1(\omega), \quad (8)$$

$$S_r(\omega) = G_r(\omega) \cdot \hat{S}'_2(\omega), \quad (9)$$

trong đó $G_l(\omega)$ mô tả độ khuếch đại tai trái và $G_r(\omega)$ mô tả độ khuếch đại tai phải. Các kết quả này trong ICLD miễn là $\hat{S}'_1(\omega)$ và $\hat{S}'_2(\omega)$ có mật độ phổ công suất tương tự. Vì độ khuếch đại tai trái và phải được sử dụng trực tiếp, các tín hiệu điều khiển phổ một tai được tái tạo thêm vào IALD.

Để đơn giản hóa thêm phương pháp được thảo luận trước đó, hai tùy chọn cho sự đơn giản hóa được mô tả. Như được đề cập trước đó, tín hiệu điều khiển liên tai chính ảnh hưởng đến mức không gian được tiếp nhận (trong mặt phẳng ngang) là IACC. Nhờ đó có thể nhận biết không sử dụng các giá trị IAPD và/hoặc IALD được tính toán trước, nhưng điều chỉnh chúng thông qua HRTF một cách trực tiếp. Đối với mục đích này, HRTF tương ứng với vị trí biểu diễn phạm vi mức nguồn mong muốn được sử dụng. Như vị trí này, trung bình của phạm vi góc phương vị/góc nâng mong muốn được chọn ở đây mà không tổn hại tính tổng quát. Sau đây, phần mô tả về cả hai tùy chọn được đưa ra.

Tùy chọn thứ nhất bao gồm sử dụng các giá trị IACC và IAPD. Tuy nhiên, ICLD được điều chỉnh sử dụng HRTF tương ứng với tâm của phạm vi mức nguồn.

Sơ đồ khối của tùy chọn thứ nhất được thể hiện trên Fig.4. $S_l(\omega)$ và $S_r(\omega)$ bây giờ được tính toán sử dụng các công thức sau:

$$S_l(\omega) = \hat{S}'_1(\omega) \cdot |\text{HRTF}_l(\omega, \bar{\phi}, \bar{\theta})|, \quad (10)$$

$$S_r(\omega) = \hat{S}'_2(\omega) \cdot |\text{HRTF}_r(\omega, \bar{\phi}, \bar{\theta})|, \quad (11)$$

với $\bar{\phi} = (\phi_1 + \phi_2)/2$ và $\bar{\theta} = (\theta_1 + \theta_2)/2$ mô tả vị trí của HRTF mà biểu diễn trung bình của phạm vi góc phương vị/góc nâng mong muốn. Lợi ích chính của tùy chọn thứ nhất bao gồm:

- Không có sự tạo hình/tô màu phổ khi mức nguồn tăng so với nguồn điểm trong tâm của phạm vi mức nguồn.

- Các yêu cầu bộ nhớ thấp hơn so với sự phát triển mạnh, như $G_l(\omega)$ và $G_r(\omega)$ không phải được lưu trữ trong bảng tìm kiếm.

Linh động hơn cho các thay đổi về tập hợp dữ liệu HRTF trong suốt thời gian hoạt động so với phương pháp phát triển mạnh, như ICC và ICPD thu được, nhưng không phải ICLD, phụ thuộc vào tập hợp dữ liệu HRTF được sử dụng trong khi tính toán trước.

Bất lợi chính của phiên bản đơn giản hóa này là nó sẽ thất bại bất cứ khi nào xảy ra thay đổi đáng kể trong IALD, so với nguồn được mở rộng khác. Trong trường hợp này, IALD sẽ không được tái tạo với độ chính xác đủ. Ví dụ, đây là trường hợp khi nguồn không được định tâm quanh góc phương vị 0° và cùng lúc đó, mức nguồn theo hướng ngang trở nên quá lớn.

Tùy chọn thứ hai bao gồm chỉ sử dụng các giá trị IACC được tính toán trước. ICPD và ICLD được điều chỉnh sử dụng HRTF tương ứng với tâm của phạm vi mức nguồn.

Sơ đồ khối của tùy chọn thứ hai được thể hiện trên Fig.5. $S_l(\omega)$ và $S_r(\omega)$ bây giờ được tính toán sử dụng các công thức sau:

$$S_l(\omega) = \hat{S}_1(\omega) \cdot \text{HRTF}_l(\omega, \bar{\phi}, \bar{\theta}), \quad (12)$$

$$S_r(\omega) = \hat{S}_2(\omega) \cdot \text{HRTF}_r(\omega, \bar{\phi}, \bar{\theta}). \quad (13)$$

Ngược với tùy chọn thứ nhất, pha và cường độ của HRTF được sử dụng bây giờ thay vì chỉ có cường độ. Điều này cho phép không chỉ điều chỉnh ICLD mà còn ICPD. Lợi ích chính của tùy chọn thứ hai bao gồm:

- Đối với tùy chọn thứ nhất, không có sự tạo hình/tô màu phổ xảy ra khi mức nguồn tăng so với nguồn điểm trong tâm của phạm vi mức nguồn.

- Thậm chí các yêu cầu bộ nhớ thấp hơn so với tùy chọn thứ nhất, vì $G_l(\omega)$ và $G_r(\omega)$ hoặc IAPD không phải được lưu trữ trong bảng tìm kiếm.

- So với tùy chọn thứ nhất, thậm chí linh hoạt hơn về các thay đổi trong tập hợp dữ liệu HRTF trong suốt thời gian vận hành. Chỉ ICC thu được phụ thuộc vào tập hợp dữ liệu HRTF được sử dụng trong khi tính toán trước.

- Sự tích hợp đủ vào các hệ thống kết xuất lập thể hiện có là khả thi, vì hai đầu vào khác nhau đơn giản, $\hat{S}_1(\omega)$ và $\hat{S}_2(\omega)$, phải được sử dụng cho việc tạo ra tín hiệu tai trái và phải.

Như đối với tùy chọn thứ nhất, phiên bản đơn giản hóa này sẽ sẽ thất bại bất cứ khi nào xảy ra thay đổi đáng kể trong IALD, so với nguồn được mở rộng khác. Ngoài ra, các thay đổi trong IAPD không nên quá lớn so với nguồn không được mở rộng. Tuy nhiên, vì IAPD của nguồn được mở rộng sẽ gần hơn với IAPD của nguồn điểm trong tâm của phạm vi mức nguồn, điều sau không mong đợi là vấn đề lớn.

Fig.6 minh họa sơ đồ phân giản lược ví dụ. Cụ thể, bản đồ phân giản lược được minh họa tại 600 và bản đồ phân giản lược 600 minh họa phạm vi không gian tối đa. Khi sơ đồ phân giản lược được xem xét là sự minh họa hai chiều của bề mặt ba chiều của hình cầu, mà được dự định bằng cách thể hiện các phạm vi góc phương vị và góc nâng nằm trong khoảng từ 0° đến 360° đối với góc phương vị và nằm trong khoảng từ -90° đến $+90^\circ$ đối với góc nâng, sẽ trở nên rõ ràng rằng, khi một người bao phủ sơ đồ phân giản lược lên hình cầu, và một người có thể đặt vị trí người nghe trong tâm của hình cầu, tất cả các phần cá nhân được minh họa dạng ví dụ bởi một số nấc, tức là S1 đến S24 có thể chia nhỏ toàn bộ bề mặt hình cầu thành các phần. Do đó, ví dụ, phần S3 mở rộng, so với phạm vi góc phương vị từ $\Phi_1 = 60^\circ$ cho tới Φ_2 đến 90° , khi chú giải của Fig.1b, Fig.4, Fig.5 được sử dụng. Phần S3 mở rộng một cách minh họa trong phạm vi góc nâng giữa -30° và 0° .

Tuy nhiên, sơ đồ phân giản lược 600 có thể cũng được sử dụng khi người nghe không được đặt trong phần tâm của hình cầu, mà được đặt tại vị trí nhất định so với hình cầu. Trong trường hợp này, chỉ các phần nhất định của hình cầu là có thể thấy, nhưng không cần thiết cho tất cả các phần của các mục thông tin tín hiệu điều khiển nhất định hình cầu là có sẵn. Cần thiết cho một số phần (được yêu cầu), các mục thông tin tín hiệu điều khiển nhất định mà tốt hơn được tính toán trước như được thảo luận sau đây hoặc tức là, cách khác, thu được bằng cách đo là có sẵn.

Cách khác, sơ đồ phân giản lược có thể được xem như phạm vi tối đa hai chiều, trong đó nguồn âm được mở rộng theo không gian có thể được định vị. Trong tình huống như vậy, khoảng cách ngang mở rộng giữa 0% và 100% và khoảng cách dọc mở rộng giữa 0% và 100%. Khoảng cách dọc hoặc sự mở rộng thực tế và khoảng cách ngang hoặc sự mở rộng thực tế có thể được ánh xạ, qua thừa số định tỉ lệ tuyệt đối nhất định đến các khoảng cách hoặc sự mở rộng tuyệt đối. Ví dụ, khi thừa số định tỉ lệ là 10 mét, 25% có thể tương ứng với 2,5 mét theo hướng ngang. Trong hướng dọc, các thừa số định tỉ lệ có thể là giống hoặc khác với thừa số định tỉ lệ trong hướng ngang. Do đó, đối với ví dụ về khoảng cách ngang/dọc/sự mở rộng, phần S5 có thể mở rộng, so với kích thước ngang, giữa 33% và 42% của thừa số định tỉ lệ (tối đa) và phần S5 sẽ mở rộng, trong phạm vi dọc, giữa 33% và 50% của thừa số định tỉ lệ dọc. Do đó, phạm vi phổ tối đa hình cầu hoặc không phải hình cầu có thể được chia nhỏ thành các phạm vi không gian giới hạn hoặc ví dụ, các phần S1 đến S24.

Để thích ứng lưới tọa độ theo cách hiệu quả cho cảm quan nghe của người, tốt hơn là có độ phân giải thấp trong hướng dọc hoặc hướng góc nâng và có độ phân giải cao hơn theo hướng ngang hoặc hướng góc phương vị. Một cách minh họa, một người có thể chỉ sử dụng các phần của hình cầu mà bao trùm toàn bộ phạm vi góc nâng mà có nghĩa rằng chỉ đường đơn của các phần mở rộng từ, ví dụ, S1 đến S12 có sẵn như các phần khác nhau hoặc các phạm vi không gian giới hạn trong đó kích thước ngang được cho bởi các giá trị góc nhất định và kích thước dọc kéo dài từ -90° đến $+90^\circ$ đối với mỗi phần. Một cách tự nhiên, các kỹ thuật chia vùng có sẵn, ví dụ, có trên ví dụ Fig.6, 24 phần trong đó các phần từ S1 đến S12 bao trùm, đối với mỗi phần,, toàn bộ phạm vi góc nâng hoặc dọc giữa -90° và 0° hoặc giữa 0% và 50%, trong đó các phần khác từ S13 đến S24 bao phủ bán cầu phía trên giữa các góc nâng từ 0° đến 90° hoặc bao phủ nửa phía trên của “đường ngang” mở rộng giữa 50% và 100%.

Fig.7 minh họa sự thực thi được ưu tiên khác của giao diện thông tin không gian 10 của Fig.1a. Cụ thể, giao diện thông tin không gian bao gồm giao diện tiếp nhận (người dùng) thực tế để nhận sự biểu thị phạm vi không gian. Sự biểu thị phạm vi không gian có thể được nhập vào bởi bản thân người dùng hoặc có thể được suy ra từ thông tin truy vết phân đầu trong trường hợp thực tế ảo hoặc bộ khớp được bổ sung 30 khớp phạm vi không gian được giới hạn được nhận thực tế với các phạm vi không gian ứng viên có

sẵn mà đã biết từ bộ cung cấp thông tin 200 để tìm ra phạm vi không gian ứng viên mà gần nhất với phạm vi không gian giới hạn đầu vào thực tế. Dựa trên phạm vi không gian ứng viên được khớp, bộ cung cấp thông tin tín hiệu điều khiển 200 từ Fig.1a phân phối một hoặc nhiều mục thông tin tín hiệu điều khiển như dữ liệu liên kênh hoặc các hàm bộ lọc. Phạm vi không gian ứng viên được khớp hoặc phạm vi không gian giới hạn có thể bao gồm cặp các góc phương vị hoặc cặp các góc nâng hoặc cả hai như được minh họa, ví dụ, trên Fig.1b, thể hiện phạm vi góc phương vị và phạm vi góc nâng cho phần.

Cách khác, như được minh họa trên Fig.6, phạm vi không gian giới hạn có thể được giới hạn bởi thông tin trên khoảng cách ngang, thông tin về phạm vi dọc hoặc thông tin về khoảng cách dọc và thông tin về khoảng cách ngang. Khi phạm vi không gian tối đa được quét trong hai chiều, không chỉ khoảng cách dọc hoặc ngang đơn là đủ mà còn cặp khoảng cách dọc và khoảng cách ngang như được minh họa với phần S5 là cần thiết. Cách khác, thông tin phạm vi không gian giới hạn có thể bao gồm mã xác định phạm vi không gian giới hạn như phần cụ thể của phạm vi không gian tối đa trong đó phạm vi không gian tối đa bao gồm nhiều phần khác nhau. Mã này, ví dụ, được cho bởi các chỉ báo S1 đến S24, vì mỗi mã được kết hợp duy nhất với phần hai chiều hoặc ba chiều theo hình học nhất định tại bản đồ phân giản lược 600.

Fig.8 minh họa sự thực thi khác của giao diện thông tin không gian bao gồm, một lần nữa, giao diện tiếp nhận người dùng 100 nhưng giờ còn bao gồm, thêm vào đó, bộ tính toán sự chiếu 120 và bộ xác định phạm vi không gian được nối tiếp theo 140. Giao diện tiếp nhận người dùng 100 thu một cách minh họa vị trí người nghe trong đó vị trí người nghe bao gồm vị trí thực tế của người dùng trong môi trường nhất định và/hoặc định hướng của người dùng tại vị trí nhất định. Do đó, vị trí người nghe có thể liên quan đến vị trí thực tế hoặc định hướng thực tế hoặc cả hai, vị trí của người nghe thực tế và định hướng của người nghe thực tế. Dựa trên dữ liệu này, bộ tính toán phép chiếu 120 tính toán, sử dụng thông tin về nguồn âm mở rộng theo không gian, được gọi là dữ liệu phép chiếu viên. Thông tin SESS có thể bao gồm hình học của nguồn âm mở rộng theo không gian và/hoặc vị trí của nguồn âm được mở rộng theo không gian và/hoặc định hướng của nguồn âm mở rộng theo không gian, v.v.. Dựa trên dữ liệu chiếu viên, bộ xác định phạm vi không gian 140 xác định phạm vi không gian giới hạn trong một trong số các lựa chọn thay thế được minh họa trên Fig.6, hoặc như được thảo luận trên Fig.10,

Fig.11 hoặc Fig.12 đến Fig.18, trong đó phạm vi không gian được cho bởi hai hoặc nhiều hơn hai điểm đặc tính được minh họa trên các ví dụ từ Fig.12 đến Fig.18, trong đó tập hợp các đặc tính luôn luôn định nghĩa phạm vi không gian giới hạn nhất định từ phạm vi không gian đầy đủ.

Fig.9a và Fig.9b minh họa các cách khác nhau để tính toán dữ liệu chiếu viên được xuất ra bởi khối 120 của Fig.8. Trong phương án của Fig.9a, giao diện thông tin không gian được tạo cấu hình để tính toán viên của nguồn âm được mở rộng theo không gian sử dụng, như thông tin về nguồn âm mở rộng theo không gian, hình học của nguồn âm mở rộng theo không gian như được biểu thị bởi khối 121. Viên của nguồn âm mở rộng theo không gian được chiếu 122 về phía người nghe sử dụng vị trí người nghe để thu được phép chiếu của viên hai chiều và ba chiều lên mặt phẳng chiếu. Cách khác, như được minh họa trên Fig.9b, nguồn âm mở rộng theo không gian và, cụ thể, hình học của nguồn âm mở rộng theo không gian như được định nghĩa bởi thông tin về hình học của nguồn âm mở rộng theo không gian được chiếu theo hướng về phía vị trí người nghe được minh họa tại khối 123, và viên của hình học được chiếu được tính toán như được biểu thị trong khối 124 để thu được phép chiếu của viên hai chiều hoặc ba chiều lên mặt phẳng chiếu. Phạm vi không gian giới hạn biểu diễn sự mở rộng dọc/ngang hoặc góc phương vị/góc nâng của viên được chiếu trên phương án của Fig.9a hoặc của viên của hình học được chiếu như thu được bởi sự thực thi Fig.9b.

Fig.10 minh họa sự thực thi được ưu tiên của giao diện thông tin không gian 10. Nó bao gồm giao diện vị trí người nghe 100 mà cũng được minh họa trên Fig.8 như giao diện tiếp nhận người dùng. Ngoài ra, vị trí và hình học của nguồn âm mở rộng theo không gian được nhập vào như cũng được minh họa trên Fig.8. Bộ chiếu 120 được cung cấp và bộ tính toán 140 để tính toán phạm vi không gian giới hạn.

Fig.11 minh họa sự thực thi được ưu tiên của giao diện thông tin không gian bao gồm giao diện 100, bộ chiếu 120, và bộ tính toán vị trí phạm vi không gian giới hạn 140. Giao diện 100 được tạo cấu hình để nhận vị trí người nghe. Bộ chiếu 120 được tạo cấu hình để tính toán phép chiếu của viên hai chiều hoặc ba chiều được kết hợp với nguồn âm mở rộng theo không gian lên mặt phẳng chiếu sử dụng vị trí người nghe như được nhận bởi giao diện 100 và sử dụng, thêm vào đó, thông tin về hình học của nguồn âm

mở rộng theo không gian và, ngoài ra, sử dụng thông tin về vị trí của nguồn âm mở rộng theo không gian trong không gian. Tốt hơn là, vị trí được định nghĩa của nguồn âm mở rộng theo không gian trong không gian và, ngoài ra, hình học của nguồn âm mở rộng theo không gian trong không gian được thu để tái tạo nguồn âm mở rộng theo không gian qua dòng bit đến tại bộ tách kênh dòng bit hoặc bộ phân tách cảnh 180. Bộ tách kênh dòng bit 180 trích, từ dòng bit, thông tin về hình học của nguồn âm mở rộng theo không gian và cung cấp thông tin này đến máy chiếu. Bộ dồn kênh dòng bit cũng trích vị trí của nguồn âm mở rộng theo không gian từ dòng bit và chuyển tiếp thông tin này đến bộ chiếu.

Tốt hơn là, dòng bit cũng bao gồm tín hiệu âm thanh cho SESS có một hoặc hai tín hiệu âm thanh khác nhau và, tốt hơn là, bộ dồn kênh dòng bit cũng trích, từ dòng bit, sự biểu diễn được nén của một hoặc nhiều tín hiệu âm thanh, và (các) tín hiệu được giải nén/được mã hóa bởi bộ giải mã như bộ giải mã âm thanh 190. Một hoặc nhiều tín hiệu được giải mã cuối cùng được chuyển tiếp đến bộ xử lý âm thanh 300 của Fig.1a, ví dụ, và bộ xử lý kết xuất ít nhất hai nguồn âm trong đường với các mục tín hiệu điều khiển được cung cấp bởi bộ cung cấp thông tin tín hiệu điều khiển 200 của Fig.1a.

Mặc dù Fig.11 minh họa thiết bị tái tạo liên quan đến dòng bit có bộ dồn kênh dòng bit 190, sự tái tạo có thể cũng xảy ra trong tình huống khác với ngữ cảnh bộ mã hóa/bộ giải mã. Ví dụ, vị trí và hình học được định nghĩa trong không gian có thể tồn tại tại thiết bị tái tạo như cảnh thực tế ảo hoặc thực tế được bổ sung, trong đó dữ liệu được tạo ra trên địa điểm và được giả định trên cùng một địa điểm. Bộ dồn kênh dòng bit 180 và bộ giải mã âm thanh 190 không thực sự cần thiết, và thông tin về hình học của nguồn âm mở rộng theo không gian và vị trí của nguồn âm mở rộng theo không gian có sẵn mà không có bất kỳ sự trích từ dòng bit nào.

Các phương án được ưu tiên của sáng chế được thảo luận. Các phương án đề cập đến kết xuất các nguồn âm mở rộng theo không gian trong 6DoF VR/AR (thực tế ảo/thực tế bổ sung).

Các phương án được ưu tiên của sáng chế được dẫn đến phương pháp, thiết bị hoặc chương trình máy tính được thiết kế để nâng cao sự tái tạo của các nguồn âm mở rộng theo không gian (Spatially Extended Sound Sources - SESS). Cụ thể, các phương án của

phương pháp theo sáng chế hoặc thiết bị xem xét vị trí liên quan biến đổi theo thời gian giữa nguồn âm mở rộng theo không gian và vị trí người nghe thực tế. Nói cách khác, các phương án của phương pháp hoặc thiết bị theo sáng chế cho phép chiều rộng nguồn nghe khớp với mức không gian của đối tượng âm được biểu diễn tại bất kỳ vị trí tương ứng với người nghe. Như vậy, phương án theo phương pháp hoặc thiết bị của sáng chế áp dụng cụ thể các ứng dụng thực tế ảo, hỗn hợp và bổ sung 6 mức tự do (6-degrees-of-freedom -6DoF) trong đó nguồn âm mở rộng theo không gian bổ sung các nguồn điểm được sử dụng một cách truyền thống.

Phương án theo phương pháp hoặc thiết bị của sáng chế kết xuất nguồn âm mở rộng theo không gian bằng cách sử dụng phạm vi không gian giới hạn. Phạm vi không gian giới hạn phụ thuộc vào vị trí của người nghe so với nguồn âm mở rộng theo không gian.

Fig. 1a thể hiện sơ đồ khối toàn cảnh của bộ kết xuất nguồn âm mở rộng theo không gian theo phương án của phương pháp hoặc thiết bị theo sáng chế. Các thành phần chính của sơ đồ khối là:

1. Vị trí người nghe: Khối này cung cấp vị trí tức thời của người nghe, như, ví dụ, được đo bởi hệ thống truy vết thực tế ảo. Khối có thể được thực thi như bộ phát hiện 100 để phát hiện hoặc như giao diện 100 để nhận vị trí người nghe.
2. Vị trí và hình học của nguồn âm mở rộng theo không gian: Khối cung cấp dữ liệu vị trí và hình học của nguồn âm mở rộng theo không gian sẽ được kết xuất, ví dụ, như một phần của sự biểu diễn cảnh thực tế ảo.
3. Phép chiếu và sự tính toán viên lồi: Khối 120 tính toán viên lồi của hình học nguồn âm mở rộng theo không gian và khi đó chiếu theo hướng về phía vị trí người nghe (ví dụ, “mặt phẳng ảnh”, xem dưới đây). Cách khác, hàm tương tự có thể đạt được đầu tiên bằng cách chiếu hình học về phía vị trí người nghe: và sau đó tính toán viên lồi của nó.
4. Vị trí của sự xác định phạm vi không gian giới hạn: Khối 140 tính toán vị trí của phạm vi không gian giới hạn từ dữ liệu chiếu viên lồi được tính

toán bởi khối trước đó. Trong sự tính toán này, có thể cũng xem xét vị trí người nghe và sau đó sự lân cận/khoảng cách của người nghe (xem dưới đây). Đầu ra là, ví dụ, các vị trí điểm định nghĩa chung phạm vi không gian giới hạn.

Fig.10 minh họa toàn cảnh của sơ đồ khối của phương án của phương pháp hoặc thiết bị theo sáng chế. Đường nét đứt biểu thị sự truyền dẫn siêu dữ liệu như hình học và vị trí.

Vị trí của các điểm định nghĩa chung phạm vi không gian giới hạn phụ thuộc vào hình học, cụ thể, mức không gian cụ thể, của nguồn âm mở rộng theo không gian và vị trí tương ứng của người nghe so với nguồn âm mở rộng theo không gian. Cụ thể, các điểm định nghĩa phạm vi không gian giới hạn có thể được định vị trên sự chiếu viền lồi của nguồn âm mở rộng theo không gian lên mặt phẳng chiếu. Mặt phẳng chiếu có thể là mặt phẳng ảnh, tức là, mặt phẳng vuông góc với đường ngắm từ người nghe đến nguồn âm mở rộng theo không gian hoặc bề mặt hình cầu quanh đầu của người nghe. Mặt phẳng chiếu được định vị tại khoảng cách nhỏ tùy ý từ tâm đến đầu của người nghe. Cách khác, viền lồi chiếu của nguồn âm mở rộng theo không gian có thể được tính toán từ các góc phương vị và góc nâng mà tập con của các tọa độ hình cầu mà là tập con của các tọa độ hình cầu liên quan từ phối cảnh của đầu người nghe. Trong các ví dụ minh họa dưới đây, mặt phẳng chiếu được ưu tiên do đặc điểm trực giác. Trong sự thực thi tính toán của viền lồi được nêu ra, sự biểu diễn góc được ưu tiên do hình thức đơn giản hơn và độ phức tạp tính toán thấp hơn. Cả phép chiếu của viền lồi của nguồn âm mở rộng theo không gian giống với viền lồi của hình học nguồn âm mở rộng theo không gian được chiếu, tức là, sự tính toán viền lồi và phép chiếu lên mặt phẳng ảnh có thể được sử dụng theo thứ tự.

Khi vị trí người nghe tương ứng với các thay đổi nguồn âm mở rộng theo không gian, khi đó phép chiếu của nguồn âm mở rộng theo không gian lên mặt phẳng chiếu thay đổi theo đó. Lần lượt, các vị trí của các điểm định nghĩa sự thay đổi phạm vi không gian giới hạn theo đó. Các điểm sẽ được chọn sao cho chúng thay đổi một cách mượt mà cho sự di chuyển liên tục của nguồn âm mở rộng theo không gian và người nghe. Viền lồi được chiếu được thay đổi khi hình học của nguồn âm mở rộng theo không gian được thay đổi. Điều này bao gồm quay hình học nguồn âm mở rộng theo không gian

trong không gian 3D mà thay thế viền lồi được chiếu. Việc quay hình học bằng với sự thay thế góc của vị trí người nghe liên quan đến nguồn âm mở rộng theo không gian và được đề cập đến theo cách bao hàm như vị trí tương ứng của người nghe và nguồn âm mở rộng theo không gian. Ví dụ, chuyển động tròn của người nghe quanh nguồn âm mở rộng theo không gian được biểu diễn bằng cách quay các điểm định nghĩa sự thay đổi phạm vi không gian giới hạn quanh trọng tâm. Một cách cân bằng, việc quay nguồn âm mở rộng theo không gian với các kết quả người nghe tĩnh theo cách giống với các điểm định nghĩa phạm vi không gian giới hạn.

Mức không gian như được tạo ra bởi phương án của phương pháp hoặc thiết bị theo sáng chế được tái tạo kế thừa một cách chính xác cho khoảng cách bất kỳ giữa nguồn âm mở rộng theo không gian và người nghe. Một cách tự nhiên, khi người dùng tiếp cận nguồn âm mở rộng theo không gian, góc mở giữa các điểm định nghĩa sự thay đổi phạm vi không gian giới hạn như phù hợp để tạo mô hình thực tế vật lý.

Do đó, sự đặt góc của các điểm định nghĩa phạm vi không gian giới hạn được xác định duy nhất bởi vị trí của viền lồi được chiếu trên mặt phẳng chiếu.

Để chi rõ hình dạng hình học/viền lồi của nguồn âm mở rộng theo không gian, phép tính gần đúng được sử dụng (và, có thể, được truyền dẫn đến bộ kết xuất hoặc lõi kết xuất) bao gồm 1D đơn giản hóa, ví dụ, đường thẳng, đường cong; 2D, ví dụ, hình ê-lip, hình chữ nhật, hình đa giác; hoặc hình 3D, ví dụ, thể bầu dục, hình hộp phẳng, hình đa giác. Hình học của nguồn âm mở rộng theo không gian hoặc hình gần đúng tương ứng, lần lượt có thể được mô tả theo các cách khác nhau, bao gồm:

- Sự mô tả tham số, tức là, hình thức hóa của hình học qua biểu thức toán học mà chấp nhận tham số bổ sung. Ví dụ, thể bầu dục ở dạng 3D có thể được mô tả bởi hàm ẩn trên hệ tọa độ Đề-các và các tham số bổ sung là sự mở rộng các trục cơ bản theo ba hướng. Các tham số khác có thể bao gồm việc quay 3D, các hàm biến dạng của bề mặt thể bầu dục.
- Phân mô tả đa giác, tức là, tập hợp các dạng hình học ban đầu như các đường thẳng, các hình tam giác, hình vuông, hình tứ diện và hình hộp phẳng. Các hình đa giác và hình đa diện ban đầu có thể được nối chuỗi đến các hình học phức lớn hơn.

Trong các cảnh ứng dụng nhất định, sự tập trung nằm trên sự lưu trữ/truyền dẫn gọn nhẹ và có thể tương tác của nội dung 6DoF VR/AR. Trong trường hợp này, toàn bộ chuỗi bao gồm ba bước sau:

1. Viết/mã hóa các nguồn âm mở rộng theo không gian mong muốn thành dòng bit
2. Truyền dẫn/lưu trữ dòng bit được tạo ra. Theo sáng chế này, dòng bit chứa, ngoài các phần tử khác, sự mô tả của các hình học nguồn âm mở rộng theo không gian (theo tham số hoặc hình đa diện) và (các) tín hiệu cơ bản nguồn được kết hợp, như bản ghi tiếng piano âm đơn hoặc âm lập thể. Dạng sóng có thể được nén sử dụng các thuật toán mã hóa âm thanh cảm quan, như mp3 hoặc mã hóa âm thanh tiên tiến (Advanced Audio Coding - AAC) MPEG-2/4.
3. Giải mã/kết xuất các nguồn âm mở rộng theo không gian dựa trên dòng bit được truyền dẫn như được mô tả trước đó.

Tiếp theo, các ví dụ thực thi thực tế khác nhau sẽ được biểu diễn. Điều này bao gồm nguồn âm mở rộng theo không gian hình cầu, nguồn âm mở rộng theo không gian thể bầu dục, nguồn âm mở rộng theo không gian đường thẳng, nguồn âm mở rộng theo không gian hình hộp phẳng, các phạm vi không gian giới hạn phụ thuộc vào khoảng cách, và/hoặc nguồn âm mở rộng theo không gian dạng piano hoặc dạng nguồn âm mở rộng theo không gian như nhạc cụ bất kỳ khác.

Như được mô tả trong các phương án của phương pháp hoặc thiết bị theo sáng chế, các phương pháp khác nhau trên đây để xác định vị trí của các điểm định nghĩa phạm vi không gian giới hạn có thể được áp dụng. Các ví dụ thực tế sau đây minh họa một số phương pháp riêng biệt trong các trường hợp cụ thể. Trong sự thực thi hoàn thiện của phương pháp hoặc thiết bị theo sáng chế, các phương pháp khác nhau có thể được tổ hợp như cân nhắc phù hợp về độ phức tạp tính toán, mục đích ứng dụng, chất lượng âm thanh và việc thực hiện dễ dàng.

Hình học nguồn âm mở rộng theo không gian được biểu thị như lưới bề mặt. Lưu ý rằng sự hiển thị hóa lưới không biểu thị rằng hình học nguồn âm mở rộng theo không gian được mô tả bởi phương pháp đa giác vì trong thực tế, hình học nguồn âm mở rộng

theo không gian có thể được tạo ra từ thông số kỹ thuật theo tham số. Vị trí người nghe được biểu thị bởi tam giác màu xanh. Trong các ví dụ sau đây, mặt phẳng ảnh được chọn như mặt phẳng chiếu và được biểu thị như mặt phẳng màu xám trong suốt mà biểu thị tập con hữu hạn của mặt phẳng chiếu. Hình học được chiếu của nguồn âm mở rộng theo không gian lên mặt phẳng chiếu được thể hiện với lưới bề mặt tương tự. Các điểm định nghĩa phạm vi không gian giới hạn trên viền lồi được nhô ra được biểu thị như các đường chéo trên mặt phẳng chiếu. Các điểm được chiếu phía sau định nghĩa phạm vi không gian giới hạn lên hình học nguồn âm mở rộng theo không gian được thể hiện như các chấm. Các điểm tương ứng định nghĩa phạm vi không gian giới hạn lên viền lồi được chiếu và các điểm được chiếu phía sau định nghĩa phạm vi không gian giới hạn lên hình học nguồn âm mở rộng theo không gian được nối bởi các đường để hỗ trợ nhận biết sự tương ứng trực quan. Các vị trí của tất cả các đối tượng liên quan được thể hiện trong hệ tọa độ Đề-các với các đơn vị theo mét. Lựa chọn của hệ tọa độ được biểu thị không ám chỉ rằng các tính toán liên quan được thực hiện với tọa độ Đề-các.

Ví dụ thứ nhất trên Fig.12 cân nhắc nguồn âm mở rộng theo không gian hình cầu. Nguồn âm mở rộng theo không gian hình cầu có kích thước cố định và vị trí cố định so với người nghe. Ba tập hợp khác nhau gồm ba, năm, và tám điểm định nghĩa phạm vi không gian giới hạn được chọn trên viền lồi được chiếu. Tất cả ba tập hợp các điểm định nghĩa phạm vi không gian giới hạn được chọn với khoảng cách đồng đều trên đường cong viền lồi. Các vị trí độ lệch của các điểm định nghĩa phạm vi không gian trên đường cong viền lồi được chọn một cách thận trọng sao cho mức ngang của hình học nguồn âm mở rộng theo không gian được biểu diễn tốt. Fig.12 minh họa nguồn âm mở rộng theo không gian hình cầu với các số khác nhau (Tức là, 3 (phía trên), 5 (ở giữa), và 8 (phía dưới)) các điểm định nghĩa phạm vi không gian giới hạn được phân bố đồng đều trên viền lồi.

Ví dụ tiếp theo trên Fig.13 cân nhắc nguồn âm được mở rộng theo không gian thể bầu dục. Nguồn âm được mở rộng theo không gian thể bầu dục có hình dạng, vị trí và việc quay cố định trong không gian 3D. Bốn điểm định nghĩa phạm vi không gian giới hạn được chọn trong ví dụ sau. Ba phương pháp khác nhau xác định vị trí của các điểm định nghĩa phạm vi không gian giới hạn được lấy ví dụ:

a) hai điểm định nghĩa phạm vi không gian giới hạn được đặt tại hai điểm cực trị ngang và hai điểm định nghĩa phạm vi không gian giới hạn được đặt tại hai điểm cực trị dọc. Trong đó, sự định vị điểm cực trị là đơn giản và thường phù hợp. Ví dụ này thể hiện rằng phương pháp có thể tạo ra các vị trí điểm mà tương đối gần với nhau.

b) Tất cả bốn điểm định nghĩa phạm vi không gian giới hạn được phân bố đồng đều trên viên lõi được chiếu. Độ lệch của các điểm định nghĩa vị trí phạm vi không gian giới hạn được chọn sao cho vị trí điểm trên cùng trùng với vị trí điểm trên cùng trong a).

c) Tất cả bốn điểm định nghĩa phạm vi không gian giới hạn được phân bố đồng đều trên viên lõi được chiếu bị co lại. Vị trí độ lệch của các vị trí điểm bằng với vị trí độ lệch được chọn trong b). Thuật toán co của viên lõi được chiếu được thực hiện về phía trọng tâm của viên lõi được chiếu với thừa số căng độc lập hướng.

Do đó, Fig.13 minh họa nguồn âm mở rộng theo không gian thể bầu dục với bốn điểm định nghĩa phạm vi không gian giới hạn dưới ba phương pháp khác nhau xác định vị trí của các điểm định nghĩa phạm vi không gian giới hạn: a/phía trên) các điểm cực trị ngang và dọc, b/giữa) các điểm được phân bố đồng đều trên viên lõi, c/phía dưới) các điểm được phân bố đồng đều trên viên lõi co.

Ví dụ tiếp theo trên Fig.14 cân nhắc nguồn âm mở rộng theo không gian đường thẳng. Trong đó, các ví dụ trước đó xem xét hình học nguồn âm mở rộng theo không gian về thể tích, ví dụ này minh họa hình học nguồn âm mở rộng theo không gian có thể cũng được chọn như đối tượng đơn chiều trong không gian 3D. Hình nhỏ a) biểu thị hai điểm định nghĩa phạm vi không gian giới hạn được đặt trên các điểm cực trị của hình học nguồn âm mở rộng theo không gian đường thẳng hữu hạn. b) Hai điểm định nghĩa phạm vi không gian giới hạn được đặt tại các điểm cực trị của hình học nguồn âm mở rộng theo không gian đường thẳng hữu hạn và một điểm bổ sung được đặt giữa đường. Như được mô tả trong các phương án của phương pháp hoặc thiết bị theo sáng chế, việc đặt các điểm bổ sung trong hình học nguồn âm mở rộng theo không gian có thể giúp điền đầy các khoảng trống lớn trong các hình học nguồn âm mở rộng theo không gian. c) Hình học nguồn âm mở rộng theo không gian đường thẳng tương tự như trong a) và b) được xem xét, tuy nhiên, góc tương ứng về phía người nghe được thay thế sao cho

chiều dài được chiếu của hình học đường thẳng là nhỏ hơn đáng kể. Như được mô tả trong các phương án của phương pháp hoặc thiết bị theo sáng chế nêu trên, kích thước giảm của viên lõi được chiếu có thể được biểu diễn bởi số lượng nhỏ các điểm định nghĩa phạm vi không gian giới hạn, trong ví dụ cụ thể này, bởi điểm đơn được đặt trong tâm của hình học đường thẳng.

Do đó, Fig.14 minh họa nguồn âm mở rộng theo không gian đường thẳng với ba phương pháp khác nhau để phân bố vị trí của các điểm định nghĩa phạm vi không gian giới hạn: a/phía trên) hai điểm cực trị trên viên lõi được chiếu; b/giữa) hai điểm cực trị trên viên lõi được chiếu với điểm bổ sung tại trung đoạn; c/phía dưới) một hoặc hai điểm định nghĩa phạm vi không gian giới hạn trong tâm của đường viên lõi như viên lõi được chiếu của đường thẳng được quay là quá nhỏ để cho phép nhiều hơn một hoặc hai điểm.

Ví dụ tiếp theo trên Fig.15 cân nhắc nguồn âm mở rộng theo không gian hình hộp phẳng. Nguồn âm được mở rộng theo không gian hình hộp phẳng có kích thước cố định và vị trí cố định, tuy nhiên vị trí tương ứng của người nghe thay đổi. Các hình con a) và b) biểu thị các phương pháp khác nhau đặt bốn điểm định nghĩa phạm vi không gian giới hạn trên viên lõi được chiếu. Các vị trí điểm được chiếu được xác định duy nhất bằng lựa chọn về viên lõi được chiếu. c) biểu thị bốn điểm định nghĩa phạm vi không gian giới hạn mà không có các vị trí chiếu phía sau được phân tách tốt. Thay vào đó, các khoảng cách của các vị trí điểm được chọn bằng với khoảng cách của trọng tâm của hình học nguồn âm mở rộng theo không gian.

Do đó, Fig.15 minh họa nguồn âm mở rộng theo không gian hình hộp phẳng với ba phương pháp khác nhau để phân bố các điểm định nghĩa phạm vi không gian giới hạn: a/ phía trên) hai điểm định nghĩa phạm vi không gian giới hạn trên trục dọc; b/giữa) hai điểm định nghĩa phạm vi không gian giới hạn trên các điểm cực trị ngang của viên lõi được chiếu và hai điểm định nghĩa phạm vi không gian giới hạn trên các điểm cực trị dọc của viên lõi được chiếu; c/phía dưới) các khoảng cách điểm được chiếu phía sau được chọn để bằng với khoảng cách của trọng tâm của hình học nguồn âm mở rộng theo không gian.

Ví dụ tiếp theo trên Fig.16 xem xét nguồn âm mở rộng theo không gian hình cầu có kích thước cố định và hình dạng, nhưng tại ba khoảng cách khác nhau so với vị trí

người nghe. b) Các điểm định nghĩa phạm vi không gian giới hạn được phân bố đồng đều trên đường cong viền lõi. Số lượng các điểm định nghĩa phạm vi không gian giới hạn được xác định động từ chiều dài của đường cong viền lõi và khoảng cách tối thiểu giữa các vị trí điểm khả thi. a) nguồn âm mở rộng theo không gian hình cầu tại khoảng cách gần sao cho bốn điểm định nghĩa phạm vi không gian giới hạn được chọn trên viền lõi được chiếu. b) Nguồn âm mở rộng theo không gian hình cầu tại khoảng cách trung bình sao cho ba điểm định nghĩa phạm vi không gian giới hạn được chọn trên viền lõi được chiếu. a) Nguồn âm mở rộng theo không gian hình cầu là tại khoảng cách xa sao cho chỉ hai điểm định nghĩa phạm vi không gian giới hạn được chọn trên viền lõi được chiếu. Như được mô tả trên các phương án của phương pháp hoặc thiết bị theo sáng chế trên đây, số lượng các điểm định nghĩa phạm vi không gian giới hạn có thể được xác định từ mức được biểu diễn trong các tọa độ góc hình cầu.

Do đó, Fig.16 minh họa nguồn âm mở rộng theo không gian hình cầu có kích thước bằng nhau tại các khoảng cách khác nhau: a/phía trên) khoảng cách gần với bốn điểm định nghĩa phạm vi không gian giới hạn được phân bố đồng đều trên viền lõi được chiếu; b/giữa) khoảng cách giữa với ba điểm định nghĩa phạm vi không gian giới hạn được phân bố đồng đều trên viền lõi được chiếu; c/phía dưới) khoảng cách xa với hai điểm định nghĩa phạm vi không gian giới hạn được phân bố đồng đều trên viền lõi được chiếu.

Ví dụ cuối cùng trên Fig.17 và Fig.18 xem xét nguồn âm mở rộng theo không gian dạng đàn pianô được đặt trong thế giới thực. Người dùng đeo màn hiển thị được gắn trên đầu (head-mounted display - HMD) và các ống nghe. Cảnh thực tế ảo được biểu diễn đến người dùng bao gồm các canvas thế giới mở và mô hình đàn pianô đứng trên sàn trong khu vực di chuyển tự do (xem Fig.17). Canvas thế giới mở là ảnh tĩnh hình cầu được chiếu lên hình cầu bao quanh người nghe. Trong trường hợp cụ thể này, canvas thế giới mở biểu thị trời xanh cùng với các đám mây trắng. Người dùng có thể đi quanh và nhìn và nghe đàn pianô từ các góc khác nhau. Trong trường hợp này, đàn pianô được kết xuất sử dụng tín hiệu điều khiển biểu diễn nguồn điểm đơn được đặt ở trọng tâm hoặc biểu diễn nguồn âm mở rộng theo không gian với ba điểm định nghĩa phạm vi không gian giới hạn trên viền lõi được chiếu (xem Fig.18).

Để đơn giản hóa sự tính toán điểm, hình học đàn pianô được tưởng tượng có dạng thể bầu dục với các kích thước giống nhau (xem Fig.17). Hai điểm thay thế được đặt trên các điểm cực trị bên trái và bên phải trên đường xích đạo, trong đó điểm thay thế thứ ba duy trì tại cực nam, xem Fig.18. Sự sắp xếp đảm bảo rằng chiều dài nguồn ngang phù hợp từ tất cả các góc tại chi phí tính toán giảm cao.

Do đó, Fig.17 minh họa nguồn âm mở rộng theo không gian dạng đàn piano với thể bầu dục theo tham số gần đúng; và Fig.18 minh họa nguồn âm mở rộng theo không gian dạng đàn piano với ba điểm định nghĩa phạm vi không gian giới hạn được phân bố trên các điểm cực trị dọc của viền lõi được chiếu và vị trí đỉnh dọc của viền lõi được chiếu. Lưu ý rằng đối với hình ảnh hóa tốt hơn, các điểm định nghĩa phạm vi không gian giới hạn được đặt trên viền lõi được chiếu kéo giãn.

Ứng dụng của công nghệ được mô tả có thể là một phần của tiêu chuẩn VR/AR 6DoF âm thanh. Trong ngữ cảnh này, có ngữ cảnh mã hóa/dòng bit/bộ giải mã (+bộ kết xuất) cổ điển:

- Trong bộ mã hóa, hình dạng của nguồn âm mở rộng theo không gian sẽ được mã hóa như thông tin phụ cùng với các dạng sóng của nguồn âm được mở rộng theo không gian mà có thể là

- tín hiệu đơn, hoặc
- tín hiệu âm lập thể (tốt hơn là được giải tương quan đầy đủ), hoặc
- thậm chí nhiều tín hiệu được ghi (cũng được ưu tiên giải tương quan đầy đủ)

đặc trưng nguồn âm được mở rộng theo không gian. Các dạng sóng có thể được mã hóa tốc độ bit thấp.

- Trong bộ giải mã/bộ kết xuất, hình dạng nguồn âm được mở rộng theo không gian và các dạng sóng tương ứng được truy hồi từ dòng bit và được sử dụng để kết xuất nguồn âm được mở rộng theo không gian như được mô tả trước đó.

Phụ thuộc vào các phương án được sử dụng và các thay thế cho các phương án được mô tả, sẽ được lưu ý rằng giao diện có thể được thực thi như bộ truy vết hoặc bộ phát hiện thực tế để phát hiện vị trí của người nghe. Tuy nhiên, vị trí người nghe sẽ được nhận diện hình từ thiết bị truy vết bên ngoài và được nạp vào thiết bị tái tạo qua giao

diện. Tuy nhiên, giao diện có thể biểu diễn chỉ đầu vào dữ liệu cho dữ liệu đầu ra từ bộ truy vết bên ngoài hoặc có thể cũng biểu diễn bản thân bộ truy vết.

Như được liệt kê, bộ tạo dòng bit có thể được thực thi để tạo ra dòng bit với chỉ một tín hiệu âm cho nguồn âm được mở rộng theo không gian và, các tín hiệu âm còn lại được tạo ra trên phía bộ giải mã hoặc phía tái tạo bằng cách giải tương quan. Khi chỉ tín hiệu đơn tồn tại, và khi toàn bộ không gian được điền đầy bằng nhau với tín hiệu đơn này, bất kỳ thông tin vị trí nào đều không cần thiết. Tuy nhiên, có thể có ích khi có, trong tình huống như vậy, ít nhất thông tin về hình học của nguồn âm mở rộng theo không gian.

Phụ thuộc vào sự thực thi, tốt hơn là sử dụng, trong bộ cung cấp thông tin tín hiệu điều khiển 200 của Fig.1a, Fig.1b, Fig.4, Fig.5, một số loại dữ liệu được tính trước để có các mục thông tin tín hiệu điều khiển chính xác cho môi trường nhất định. Dữ liệu tính toán trước này, tức là, tập hợp các giá trị cho mỗi phần như từ sơ đồ phần 600 của Fig.6 có thể được đo và lưu trữ để dữ liệu trong, ví dụ, bảng tìm kiếm 210 và các khối HRTF 220 được xác định theo kinh nghiệm. Trong các phương án khác, dữ liệu này có thể được tính toán trước hoặc dữ liệu có thể được suy ra theo quy trình tính toán trước theo kinh nghiệm hỗn hợp. Tiếp theo, phương án được ưu tiên để tính toán dữ liệu này được đưa ra.

Trong khi tạo ra bảng tìm kiếm, các giá trị IACC, IAPD và IALD cần cho sự tổng hợp SESS, như được mô tả trước đó, và được tính toán trước cho số lượng các phạm vi mức nguồn.

Như được đề cập trước đó, như mô hình phía dưới, SESS được mô tả bởi số lượng vô hạn gồm các nguồn điểm được giải tương quan được phân bố qua toàn bộ phạm vi mức nguồn. Mô hình này được làm gần đúng ở đây bằng cách đặt một nguồn điểm giải tương quan tại mỗi vị trí tập hợp dữ liệu HRTF trong phạm vi mức nguồn mong muốn. Bằng cách cuộn các tín hiệu này với HRTF tương ứng, tín hiệu tại bên trái và bên phải thu được, $Y_l(\omega)$ tương ứng $Y_r(\omega)$, có thể được xác định. Từ những điều này, giá trị IACC, IAPD và IALD có thể được suy ra. Sau đây, đạo hàm các biểu thức tương ứng được đưa ra.

Cho N tín hiệu giải tương quan $S_n(\omega)$ với mật độ phổ công suất bằng nhau:

$$S_n(\omega) = P(\omega) \cdot e^{j\phi_n(\omega)}, \quad n = 1, \dots, N, \quad (14)$$

với

$$\mathbf{E}\{S_n(\omega) \cdot S_m^*(\omega)\} = \begin{cases} 0, & n \neq m \\ P(\omega)^2, & n = m, \end{cases} \quad (15)$$

trong đó N bằng số điểm tập hợp dữ liệu HRTF trong phạm vi mức nguồn mong muốn. N tín hiệu đầu vào do đó được đặt tại vị trí tập hợp dữ liệu HRTF khác nhau, với

$$\text{HRTF}_l(\omega, n) = A_{l,n} \cdot e^{j\phi_{l,n}}, \quad n = 1, \dots, N, \quad (16)$$

$$\text{HRTF}_r(\omega, n) = A_{r,n} \cdot e^{j\phi_{r,n}}, \quad n = 1, \dots, N. \quad (17)$$

Lưu ý: $A_{l,n}$, $A_{r,n}$, $\phi_{l,n}$, và $\phi_{r,n}$, nhìn chung phụ thuộc vào ω . Tuy nhiên, sự phụ thuộc này được loại bỏ ở đây cho sự đơn giản hóa về ký hiệu. Sử dụng phương trình (16), (17), các tín hiệu tai trái và phải, $Y_l(\omega)$ tương ứng $Y_r(\omega)$, có thể được thể hiện như sau:

$$Y_l(\omega) = \sum_{n=1}^N A_{l,n} \cdot e^{j\phi_{l,n}} \cdot S_n(\omega), \quad (18)$$

$$Y_r(\omega) = \sum_{n=1}^N A_{r,n} \cdot e^{j\phi_{r,n}} \cdot S_n(\omega). \quad (19)$$

Để xác định IACC, IALD và IAPD, các biểu thức đầu tiên $\mathbf{E}\{Y_l(\omega) \cdot Y_r^*(\omega)\}$, $\mathbf{E}\{|Y_l(\omega)|^2\}$ và $\mathbf{E}\{|Y_r(\omega)|^2\}$ được đạo hàm:

$$\begin{aligned}
\mathbf{E}\{Y_l(\omega) \cdot Y_r^*(\omega)\} &= \mathbf{E}\left\{\sum_{n=1}^N A_{l,n} \cdot e^{j\phi_{l,n}} \cdot S_n(\omega) \cdot \sum_{m=1}^N A_{r,m} \cdot e^{-j\phi_{r,m}} \cdot S_m^*(\omega)\right\} \\
&= \mathbf{E}\left\{\sum_{n=1}^N \sum_{m=1}^N A_{l,n} \cdot A_{r,m} \cdot e^{j(\phi_{l,n}-\phi_{r,m})} \cdot S_n(\omega) \cdot S_m^*(\omega)\right\} \\
&= \sum_{n=1}^N \sum_{m=1}^N A_{l,n} \cdot A_{r,m} \cdot e^{j(\phi_{l,n}-\phi_{r,m})} \cdot \mathbf{E}\{S_n(\omega) \cdot S_m^*(\omega)\} \\
&= P(\omega)^2 \cdot \sum_{n=1}^N A_{l,n} \cdot A_{r,n} \cdot e^{j(\phi_{l,n}-\phi_{r,n})},
\end{aligned} \tag{20}$$

$$\begin{aligned}
\mathbf{E}\{|Y_l(\omega)|^2\} &= \mathbf{E}\left\{\sum_{n=1}^N A_{l,n} \cdot e^{j\phi_{l,n}} \cdot S_n(\omega) \cdot \sum_{m=1}^N A_{l,m} \cdot e^{-j\phi_{l,m}} \cdot S_m^*(\omega)\right\} \\
&= \mathbf{E}\left\{\sum_{n=1}^N \sum_{m=1}^N A_{l,n} \cdot A_{l,m} \cdot e^{j(\phi_{l,n}-\phi_{l,m})} \cdot S_n(\omega) \cdot S_m^*(\omega)\right\} \\
&= \sum_{n=1}^N \sum_{m=1}^N A_{l,n} \cdot A_{l,m} \cdot e^{j(\phi_{l,n}-\phi_{l,m})} \cdot \mathbf{E}\{S_n(\omega) \cdot S_m^*(\omega)\} \\
&= \sum_{n=1}^N A_{l,n}^2
\end{aligned} \tag{21}$$

$$\mathbf{E}\{|Y_r(\omega)|^2\} = P(\omega)^2 \cdot \sum_{n=1}^N A_{r,n}^2. \tag{22}$$

Sử dụng phương trình (20) đến (22), các biểu thức sau đây cho IACC(ω), IALD(ω) và IAPD(ω) có thể được xác định:

$$\begin{aligned}
\text{IACC}(\omega) &= \frac{\mathbf{E}\{Y_l(\omega) \cdot Y_r^*(\omega)\}}{\sqrt{\mathbf{E}\{|Y_l(\omega)|^2\} \cdot \mathbf{E}\{|Y_r(\omega)|^2\}}} \\
&= \frac{\sum_{n=1}^N A_{l,n} \cdot A_{r,n} \cdot e^{j(\phi_{l,n}-\phi_{r,n})}}{\sqrt{\sum_{n=1}^N A_{l,n}^2 \cdot \sum_{m=1}^N A_{r,m}^2}},
\end{aligned} \tag{23}$$

$$\begin{aligned}
\text{IALD}(\omega) &= 10 \log_{10} \frac{\mathbf{E}\{|Y_l(\omega)|^2\}}{\mathbf{E}\{|Y_r(\omega)|^2\}} \\
&= 10 \log_{10} \frac{\sum_{n=1}^N A_{l,n}^2}{\sum_{n=1}^N A_{r,n}^2}, \\
(24)
\end{aligned}$$

$$\begin{aligned}
\text{IAPD}(\omega) &= \angle(\mathbf{E}\{Y_l(\omega) \cdot Y_r^*(\omega)\}) \\
&= \angle(\text{IACC}(\omega)) \\
&= \angle\left(\sum_{n=1}^N A_{l,n} \cdot A_{r,n} \cdot e^{j(\phi_{l,n} - \phi_{r,n})}\right). \\
(25)
\end{aligned}$$

Độ khuếch đại tai trái và phải, $G_l(\omega)$ tương ứng $G_r(\omega)$, được xác định bằng cách chuẩn hóa $E\{|Y_l(\omega)|^2\}$ tương ứng $E\{|Y_r(\omega)|^2\}$ bởi số lượng nguồn cũng như công suất nguồn:

$$G_l(\omega) = \sqrt{\frac{\mathbf{E}\{|Y_l(\omega)|^2\}}{N \cdot P(\omega)^2}} = \sqrt{\frac{\sum_{n=1}^N A_{l,n}^2}{N}}$$

(26)

$$G_r(\omega) = \sqrt{\frac{\mathbf{E}\{|Y_r(\omega)|^2\}}{N \cdot P(\omega)^2}} = \sqrt{\frac{\sum_{n=1}^N A_{r,n}^2}{N}}$$

(27)

Như có thể thấy, tất cả các biểu thức phụ thuộc chỉ vào tập hợp dữ liệu HRTF được chọn và không phụ thuộc vào các tín hiệu đầu vào nữa.

Để giảm độ phức tạp tính toán trong khi tạo ra bảng tìm kiếm, một khả năng là không cân nhắc từng vị trí tập hợp dữ liệu HRTF có sẵn. Trong trường hợp này, khoảng cách mong muốn được định nghĩa. Trong khi quy trình này giảm độ phức tạp tính toán trong khi

tính toán trước, đến một mức độ, điều này sẽ dẫn đến sự suy giảm của giải pháp.

Các phương án được ưu tiên của sáng chế cung cấp các lợi ích đáng kể so với tình trạng kỹ thuật.

Từ thực tế rằng phương pháp được đề xuất yêu cầu hai tín hiệu đầu vào được giải tương quan, nhiều lợi ích phát sinh so với tình trạng kỹ thuật hiện thời mà yêu cầu số lượng tín hiệu đầu vào được giải tương quan lớn hơn:

- Phương pháp được đề xuất trình bày độ phức tạp tính toán thấp hơn, vì chỉ một bộ giải tương quan phải được áp dụng. Ngoài ra, chỉ hai tín hiệu đầu vào phải được lọc.
- Sự giải tương quan theo cặp thường cao hơn khi tạo ra ít tín hiệu được giải tương quan hơn (và đồng thời cho phép lượng tương tự của sự suy giảm tín hiệu), sự tái tạo chính xác hơn của các tín hiệu điều khiển thính giác được mong đợi.
- Tương tự, nhiều sự suy giảm tín hiệu hơn được mong đợi để đạt đến lượng giống với sự giải tương quan theo cặp và do đó độ chính xác giống nhau của các tín hiệu điều khiển thính giác được tái tạo.

Tiếp theo, một vài đặc điểm thú vị của các phương án của sáng chế được tóm tắt.

1. Chỉ hai tín hiệu đầu vào được giải tương quan (hoặc một tín hiệu đầu vào cộng thêm bộ giải tương quan) được cần đến.

2. Sự điều chỉnh [lựa chọn tần số] của các tín hiệu điều khiển lập thể của các tín hiệu đầu vào này để đạt được một cách hiệu quả các tín hiệu đầu ra lập thể cho nguồn âm mở rộng theo không gian (thay vì tạo mô hình nhiều nguồn điểm đơn mà bao phủ diện tích/thể tích của SESS)

(a) Các ICC đầu vào luôn luôn được điều chỉnh.

(b) Các ICPD/ICTD và ICLD có thể được điều chỉnh theo bước xử lý chuyên dụng hoặc có thể được đưa vào các tín hiệu sử dụng sự xử lý HRIR/HRTF với các đặc điểm này.

3. Các tín hiệu điều khiển lập thể đích [lựa chọn tần số] được xác định từ lưu trữ được tính toán trước (bảng tìm kiếm hoặc các phương tiện khác lưu trữ dữ liệu đa chiều như bảng mã véctor hoặc khớp hàm đa chiều, GMM, SVM) như hàm của phạm vi không

gian sẽ được điền đầy (ví dụ cụ thể: phạm vi góc phương vị, phạm vi góc nâng)

(a) Các IACC đích luôn luôn được lưu trữ và thu hồi/được sử dụng để tổng hợp.

(b) Các IAPD/IATD và IALD đích có thể được lưu trữ và được thu hồi/sử dụng để tổng hợp hoặc được thay thế bằng sự xử lý HRIR/HRTF.

Sự thực thi được ưu tiên của sáng chế có thể là một phần của VR/AR DoF âm thanh MPEG-I (tiêu chuẩn thực tế ảo/ thực tế bổ sung). Trong ngữ cảnh này, có ngữ cảnh ứng dụng mã hóa/dòng bit/bộ giải mã (và bộ kết xuất). Trong bộ mã hóa, hình dạng của nguồn âm mở rộng theo không gian hoặc một số nguồn âm mở rộng theo không gian sẽ được mã hoá như thông tin phụ cùng với (một hoặc nhiều) dạng sóng “không gian” của nguồn âm mở rộng theo không gian. Các dạng sóng này biểu diễn tín hiệu được nhập vào khối 300, tức là, tín hiệu âm thanh cho nguồn âm mở rộng theo không gian có thể là tốc độ bit thấp được mã hóa bằng AAC, EVS hoặc bất kỳ bộ mã hóa nào khác. Trong bộ giải mã/bộ kết xuất, trong đó ứng dụng là, ví dụ, được minh họa trên Fig.11 như bao gồm bộ tách kênh dòng bit (bộ phân tách 180 và bộ giải mã âm thanh 190), hình dạng SESS và các dạng sóng tương ứng được truy hồi từ dòng bit và được sử dụng để kết xuất SESS. Các quy trình được minh họa so với sáng chế hiện thời cung cấp bộ giải mã/bộ kết xuất chất lượng cao, nhưng độ phức tạp.

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

Phụ thuộc vào các yêu cầu thực thi nhất định, các phương án của sáng chế có thể được thực thi 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 có các tín hiệu điều khiển có thể đọc được bằng điện tử, mà có khả năng kết hợp với hệ thống máy tính khả trình, sao cho một trong số các phương pháp được mô tả ở đây được thực hiện.

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

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

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

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

Phương án khác của các phương pháp theo sáng chế là, dòng dữ liệu hoặc chuỗi tín hiệu thể hiện chương trình máy tính để thực hiện một trong các phương pháp đã được mô tả ở đây. Ví dụ, có thể 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 Liên mạng.

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

Phương án khác bao gồm máy tính có chương trình máy tính được cài đặt trên 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

pháp được mô tả ở đây. Thông thường, các phương pháp tốt hơn là được thực hiện bởi thiết bị phần cứng bất kỳ.

Các phương án được mô tả ở trên chỉ mang tính minh họa cho các nguyên lý của sáng chế. Cần hiểu rằng các cải biên và biến đổi của các phương án và các chi tiết được mô tả ở đây sẽ là rõ ràng đối với người có trình độ trong lĩnh vực kỹ thuật tương ứng. Do đó, mục đích của sáng chế là chỉ được 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 bản mô tả và giải thích của các phương án ở đây.

YÊU CẦU BẢO HỘ

1. Thiết bị tổng hợp nguồn âm mở rộng theo không gian, thiết bị bao gồm:

giao diện thông tin không gian (100) để nhận chỉ thị phạm vi không gian biểu thị phạm vi không gian giới hạn cho nguồn âm mở rộng theo không gian trong phạm vi không gian tối đa (600);

bộ cung cấp thông tin tín hiệu điều khiển (200) để cung cấp một hoặc nhiều mục thông tin tín hiệu điều khiển phản hồi phạm vi không gian giới hạn, trong đó một hoặc nhiều mục thông tin tín hiệu điều khiển bao gồm giá trị tương quan liên kênh được cung cấp phản hồi lại phạm vi không gian giới hạn; và

bộ xử lý âm thanh (300) để xử lý tín hiệu âm thanh biểu diễn nguồn âm mở rộng theo không gian sử dụng một hoặc nhiều mục thông tin tín hiệu điều khiển,

trong đó tín hiệu âm thanh bao gồm kênh âm thanh thứ nhất cho nguồn âm được mở rộng theo không gian và kênh âm thanh thứ hai cho nguồn âm được mở rộng theo không gian, hoặc trong đó tín hiệu âm thanh bao gồm kênh âm thanh thứ nhất cho nguồn âm được mở rộng theo không gian và kênh âm thanh thứ hai cho nguồn âm được mở rộng theo không gian được suy ra từ kênh âm thanh thứ nhất bởi bộ xử lý kênh thứ hai (310), và

trong đó bộ xử lý âm thanh (300) được tạo cấu hình để thực hiện sự xử lý tương quan cho kênh âm thanh thứ nhất cho nguồn âm được mở rộng theo không gian và kênh âm thanh thứ hai cho nguồn âm được mở rộng theo không gian để áp đặt (320) sự tương quan giữa kênh âm thanh thứ nhất cho nguồn âm được mở rộng theo không gian và kênh âm thanh thứ hai cho nguồn âm được mở rộng theo không gian sử dụng giá trị tương quan liên kênh được cung cấp phản hồi phạm vi không gian giới hạn.

2. Thiết bị theo điểm 1,

trong đó bộ cung cấp thông tin tín hiệu điều khiển (200) được tạo cấu hình để cung cấp, như mục thông tin tín hiệu điều khiển khác, ít nhất một trong số mục chênh lệch pha liên kênh, mục chênh lệch thời gian liên kênh, sự chênh lệch mức liên kênh và mục độ khuếch đại, và mục thông tin độ khuếch đại thứ nhất và độ khuếch đại thứ hai, và

trong đó bộ xử lý âm thanh (300) được tạo cấu hình để áp đặt sự chênh lệch pha liên kênh, sự chênh lệch thời gian liên kênh hoặc sự chênh lệch mức liên kênh hoặc các mức tuyệt đối của kênh âm thanh thứ nhất và kênh âm thanh thứ hai sử dụng ít nhất một trong số mục chênh lệch pha liên kênh, mục chênh lệch thời gian liên kênh, sự chênh lệch mức liên kênh và mục độ khuếch đại, và mục độ khuếch đại thứ nhất và thứ hai.

3. Thiết bị theo điểm 1,

trong đó bộ xử lý âm thanh (300) được tạo cấu hình để áp đặt, tiếp theo sự xác định (320) của sự tương quan, sự chênh lệch pha liên kênh (330), sự chênh lệch thời gian liên kênh hoặc sự chênh lệch mức liên kênh (340) hoặc các mức tuyệt đối của kênh thứ nhất và kênh thứ hai, hoặc

trong đó bộ xử lý kênh thứ hai (310) bao gồm bộ lọc giải tương quan hoặc bộ xử lý mạng nơron để suy ra, từ kênh âm thanh thứ nhất, kênh âm thanh thứ hai để kênh âm thanh thứ hai được giải tương quan từ kênh âm thanh thứ nhất.

4. Thiết bị theo điểm 1,

trong đó bộ cung cấp thông tin tín hiệu điều khiển (200) bao gồm bộ cung cấp hàm bộ lọc (220) để cung cấp các hàm bộ lọc âm thanh như một hoặc nhiều mục thông tin phản hồi phạm vi không gian giới hạn, và

trong đó bộ xử lý âm thanh (300) bao gồm bộ ứng dụng bộ lọc (350) để áp dụng các hàm bộ lọc âm thanh đến kênh âm thanh thứ nhất và kênh âm thanh thứ hai.

5. Thiết bị theo điểm 4,

trong đó các hàm bộ lọc âm thanh bao gồm, đối với mỗi kênh âm thanh thứ nhất và thứ hai, hàm biến đổi liên quan phần đầu, phản hồi xung liên quan phần đầu, phản hồi xung phòng âm lập thể hoặc phản hồi xung phòng, hoặc

trong đó bộ xử lý kênh thứ hai (310) bao gồm bộ lọc giải tương quan hoặc bộ xử lý mạng nơron để suy ra, từ kênh âm thanh thứ nhất, kênh âm thanh thứ hai để kênh âm thanh thứ hai được giải tương quan từ kênh âm thanh thứ nhất.

6. Thiết bị theo điểm 4 hoặc 5,

trong đó bộ ứng dụng bộ lọc (350) được tạo cấu hình để áp dụng các hàm bộ lọc âm thanh cho kết quả của sự xác định giải tương quan (320) được thực hiện bởi bộ xử lý âm thanh (300) phản hồi giá trị tương quan liên kênh.

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

trong đó bộ cung cấp thông tin tín hiệu điều khiển (200) bao gồm bộ nhớ (210) để lưu trữ thông tin trên các mục thông tin tín hiệu điều khiển khác nhau liên quan đến các phạm vi không gian giới hạn khác nhau, và

giao diện đầu ra để truy hồi, sử dụng bộ nhớ (210), một hoặc nhiều mục thông tin tín hiệu điều khiển được kết hợp với phạm vi không gian giới hạn.

8. Thiết bị theo điểm 7, trong đó bộ nhớ (210) bao gồm ít nhất một trong số bảng tìm kiếm, bảng mã vectơ, khớp hàm đa chiều, mô hình hỗn hợp Gausơ (Gaussian Mixture Model - GMM), và máy vectơ hỗ trợ (Support Vector Machine - SVM), và

trong đó giao diện đầu ra được tạo cấu hình để truy hồi một hoặc nhiều mục thông tin tín hiệu điều khiển bằng cách tìm kiếm bảng tìm kiếm hoặc bằng cách sử dụng bảng mã vectơ, hoặc bằng cách sử dụng khớp hàm đa chiều, hoặc bằng cách sử dụng GMM hoặc SVM.

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

trong đó bộ cung cấp thông tin tín hiệu điều khiển (200) được tạo cấu hình để lưu trữ thông tin trên một hoặc nhiều mục thông tin tín hiệu điều khiển được kết hợp với tập hợp gồm các phạm vi không gian ứng viên được đặt cách, tập hợp các phạm vi không gian giới hạn được đặt cách bao phủ phạm vi không gian tối đa (600), trong đó bộ cung cấp thông tin tín hiệu điều khiển (200) được tạo cấu hình để khớp (30) phạm vi không gian giới hạn với phạm vi không gian giới hạn ứng viên định nghĩa phạm vi không gian giới hạn và cung cấp một hoặc nhiều mục thông tin tín hiệu điều khiển ứng viên được kết hợp với phạm vi không gian giới hạn ứng viên được khớp, hoặc

trong đó phạm vi không gian giới hạn bao gồm ít nhất một trong số cặp góc phương vị, cặp góc nâng, thông tin về khoảng cách ngang, thông tin về khoảng cách dọc, thông tin về khoảng cách chung, và cặp góc phương vị và cặp góc nâng, hoặc

trong đó sự biểu thị phạm vi không gian bao gồm mã (S3, S5) nhận biết phạm vi không gian giới hạn như phần cụ thể của phạm vi không gian tối đa (600), trong đó phạm vi không gian tối đa (600) bao gồm nhiều phần khác nhau.

10. Thiết bị theo điểm 9, trong đó một phần trong số nhiều phần khác nhau có sự mở rộng thứ nhất theo hướng góc phương vị hoặc hướng ngang và sự mở rộng thứ hai theo hướng góc nâng hoặc hướng dọc, trong đó sự mở rộng thứ hai theo hướng góc nâng hoặc hướng dọc của phần là lớn hơn sự mở rộng thứ nhất, hoặc trong đó sự mở rộng thứ hai bao phủ phạm vi hướng góc nâng hoặc hướng dọc tối đa.

11. Thiết bị theo điểm 9 hoặc 10, trong đó nhiều phần khác nhau được định nghĩa theo cách mà khoảng cách giữa các tâm của các phần liền kề theo hướng góc phương vị hoặc hướng ngang là lớn hơn 5 độ hoặc thậm chí lớn hơn hoặc bằng 10 độ.

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

trong đó bộ xử lý âm thanh (300) được tạo cấu hình để tạo ra, từ tín hiệu âm thanh, kênh thứ nhất được xử lý và kênh thứ hai được xử lý cho sự kết xuất lập thể hoặc sự kết xuất loa phát thanh hoặc sự kết xuất loa phát thanh giảm nhiễu xuyên âm hoạt động.

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

trong đó bộ cung cấp thông tin tín hiệu điều khiển (200) được tạo cấu hình để cung cấp một hoặc nhiều giá trị tín hiệu điều khiển liên kênh như một hoặc nhiều mục thông tin,

trong đó bộ xử lý âm thanh (300) được tạo cấu hình để tạo ra (320, 330, 340, 350), từ tín hiệu âm thanh, kênh thứ nhất được xử lý và kênh thứ hai được xử lý sao cho kênh thứ nhất được xử lý và kênh thứ hai được xử lý có một hoặc nhiều tín hiệu điều khiển liên kênh như được điều khiển bởi một hoặc nhiều giá trị tín hiệu điều khiển liên kênh.

14. Thiết bị thiết bị theo một điểm trong số các điểm nêu trên, trong đó bộ cung cấp thông tin tín hiệu điều khiển (200) được tạo cấu hình để cung cấp một hoặc nhiều mục thông tin tín hiệu điều khiển cho nhiều băng tần phản hồi phạm vi không gian giới hạn giống nhau cho nhiều băng tần, trong đó các mục thông tin tín hiệu điều khiển cho các băng khác nhau là khác nhau.

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

trong đó bộ cung cấp thông tin tín hiệu điều khiển (200) được tạo cấu hình để cung cấp một hoặc nhiều mục thông tin tín hiệu điều khiển cho nhiều băng tần khác nhau, và

trong đó bộ xử lý âm thanh (300) được tạo cấu hình để xử lý tín hiệu âm thanh trong miền phổ, trong đó mục thông tin tín hiệu điều khiển được áp dụng cho nhiều giá trị phổ của tín hiệu âm thanh trong băng.

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

trong đó kênh âm thanh thứ nhất và kênh âm thanh thứ hai được giải tương quan với nhau bởi độ giải tương quan nhất định,

trong đó bộ cung cấp thông tin tín hiệu điều khiển (200) được tạo cấu hình để cung cấp giá trị tương quan liên kênh như một hoặc nhiều mục thông tin tín hiệu điều khiển, và

trong đó bộ xử lý âm thanh (300) được tạo cấu hình để giảm (320) mức độ tương quan nhất định giữa kênh thứ nhất và kênh thứ hai đến giá trị được biểu thị bởi giá trị tương quan liên kênh được cung cấp bởi bộ cung cấp thông tin tín hiệu điều khiển (200).

17. Thiết bị theo một điểm trong số các điểm nêu trên, thiết bị còn bao gồm giao diện tín hiệu âm thanh (305) để nhận tín hiệu âm thanh biểu diễn nguồn âm mở rộng theo không gian, trong đó tín hiệu âm thanh chỉ bao gồm kênh âm thanh thứ nhất hoặc trong đó tín hiệu âm thanh chỉ bao gồm kênh âm thanh thứ nhất và kênh âm thanh thứ hai, hoặc trong đó tín hiệu âm thanh không bao gồm nhiều kênh âm thanh hơn kênh âm thanh thứ nhất và kênh âm thanh thứ hai.

18. Thiết bị theo một điểm trong số các điểm nêu trên, trong đó giao diện thông tin không gian (100) được tạo cấu hình để

nhận (100) vị trí người nghe như sự chỉ thị phạm vi không gian,

tính toán (120) sự chiếu của viên hai chiều hoặc ba chiều được kết hợp với nguồn âm mở rộng theo không gian lên mặt phẳng chiếu sử dụng, như sự chỉ thị phạm vi không gian, vị trí người nghe và thông tin về nguồn âm mở rộng theo không gian như hình học hoặc vị trí của nguồn âm được mở rộng theo không gian hoặc để tính toán (120) viên hai chiều hoặc ba chiều của sự chiếu của hình học của nguồn âm mở rộng theo không gian lên mặt phẳng chiếu sử dụng, như sự chỉ thị phạm vi không gian, vị trí người nghe

và thông tin về nguồn âm mở rộng theo không gian như hình học hoặc vị trí của nguồn âm mở rộng theo không gian, và

xác định (140) phạm vi không gian được giới hạn từ dữ liệu chiếu viên.

19. Thiết bị theo điểm 18, trong đó giao diện thông tin không gian (100) được tạo cấu hình để tính toán (121) viên của nguồn âm mở rộng theo không gian sử dụng như thông tin về nguồn âm mở rộng theo không gian, hình học của nguồn âm mở rộng theo không gian và để chiếu (122) viên theo hướng về phía người nghe sử dụng vị trí người nghe để thu được sự chiếu của viên hai chiều hoặc ba chiều lên mặt phẳng chiếu, hoặc để chiếu (123) hình học của nguồn âm mở rộng theo không gian như được định nghĩa bởi thông tin về hình học của nguồn âm mở rộng theo không gian theo hướng về phía vị trí người nghe và để tính toán (124) viên của hình học được chiếu để thu được sự chiếu của viên hai chiều hoặc ba chiều lên mặt phẳng chiếu.

20. Thiết bị theo điểm 18 hoặc 19, trong đó giao diện thông tin không gian (100) được tạo cấu hình để xác định phạm vi không gian giới hạn để đường biên của phần được định nghĩa bởi phạm vi không gian giới hạn được đặt trên phía bên phải của mặt phẳng chiếu so với người nghe và/hoặc bên trái của mặt phẳng chiếu so với người nghe và/hoặc phía trên mặt phẳng chiếu so với người nghe và phía dưới mặt phẳng chiếu so với người nghe hoặc trùng, ví dụ, trong dung sai $\pm 10\%$ với một trong số đường biên bên phải, đường biên bên trái, đường biên phía trên và đường biên phía dưới của mặt phẳng chiếu so với người nghe.

21. Phương pháp tổng hợp nguồn âm mở rộng theo không gian, phương pháp bao gồm các bước:

nhận chỉ thị phạm vi không gian biểu thị phạm vi không gian giới hạn cho nguồn âm mở rộng theo không gian trong phạm vi không gian tối đa (600);

cung cấp một hoặc nhiều mục thông tin tín hiệu điều khiển phản hồi phạm vi không gian giới hạn trong đó một hoặc nhiều mục thông tin tín hiệu điều khiển bao gồm giá trị tương quan liên kênh được cung cấp phản hồi phạm vi không gian giới hạn; và

xử lý tín hiệu âm thanh biểu diễn nguồn âm mở rộng theo không gian sử dụng một hoặc nhiều mục thông tin tín hiệu điều khiển,

trong đó tín hiệu âm thanh bao gồm kênh âm thanh thứ nhất cho nguồn âm được mở rộng theo không gian và kênh âm thanh thứ hai cho nguồn âm được mở rộng theo không gian, hoặc trong đó tín hiệu âm thanh bao gồm kênh âm thanh thứ nhất cho nguồn âm được mở rộng theo không gian và kênh âm thanh thứ hai cho nguồn âm được mở rộng theo không gian được suy ra từ kênh âm thanh thứ nhất bởi bộ xử lý kênh thứ hai (310), và

trong đó bước xử lý bao gồm thực hiện sự xử lý tương quan cho kênh âm thanh thứ nhất cho nguồn âm được mở rộng theo không gian và kênh âm thanh thứ hai cho nguồn âm được mở rộng theo không gian để áp đặt (320) sự tương quan giữa kênh âm thanh thứ nhất cho nguồn âm được mở rộng theo không gian và kênh âm thanh thứ hai cho nguồn âm được mở rộng theo không gian sử dụng giá trị tương quan liên kênh được cung cấp phản hồi phạm vi không gian giới hạn.

22. Vật ghi đọc được bằng máy tính bao gồm chương trình máy tính để thực hiện, khi chạy trên máy tính hoặc bộ xử lý, phương pháp theo điểm 21.

1/18

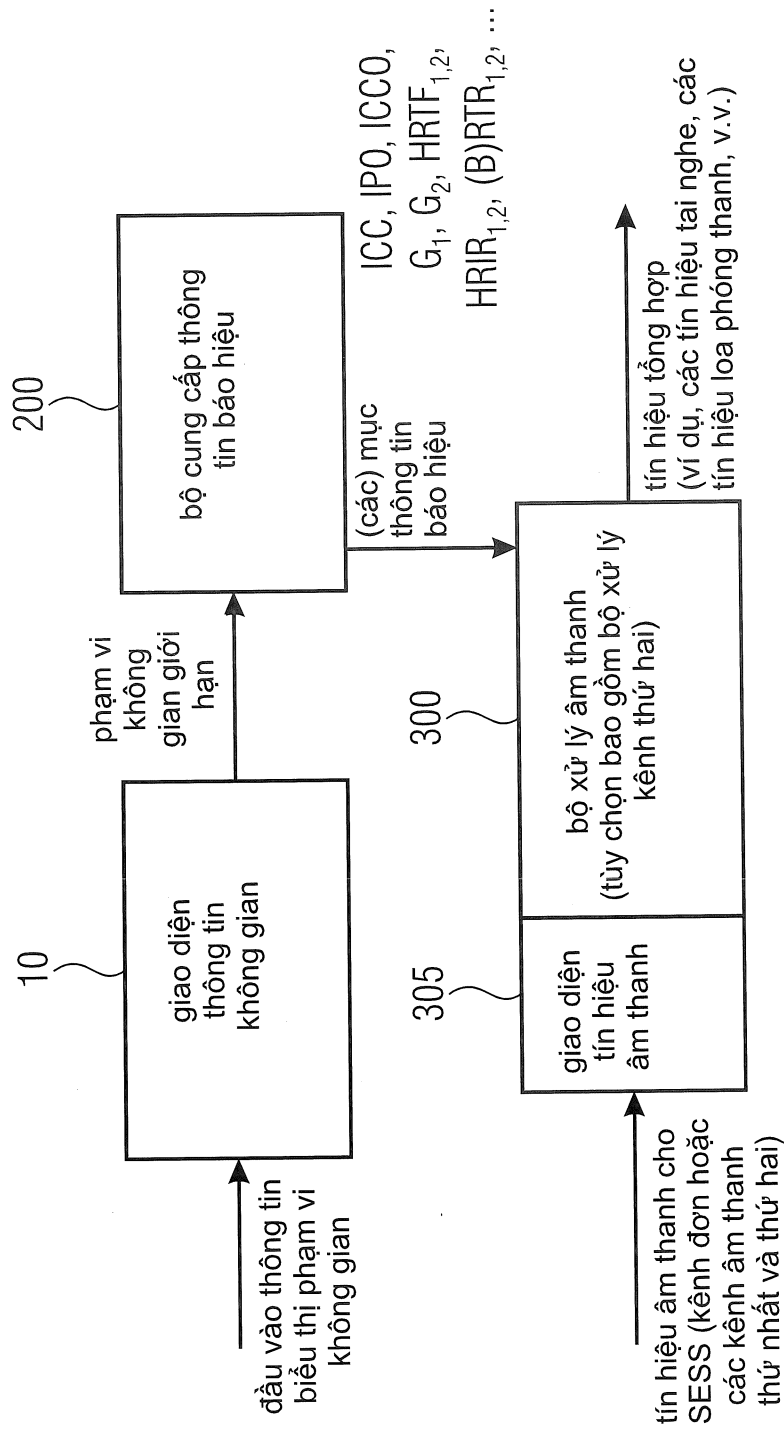


Fig. 1a

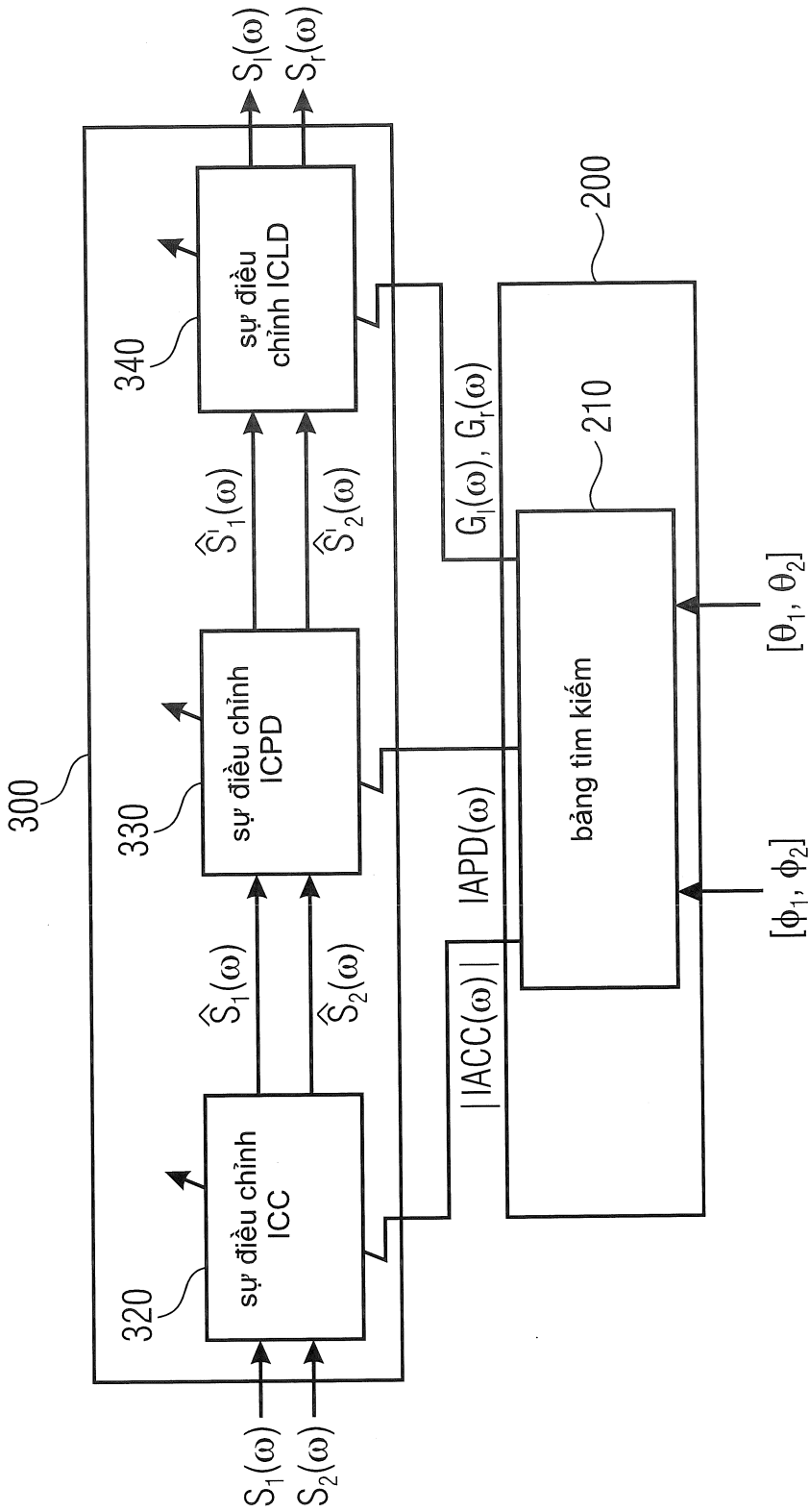


Fig. 1b

3/18

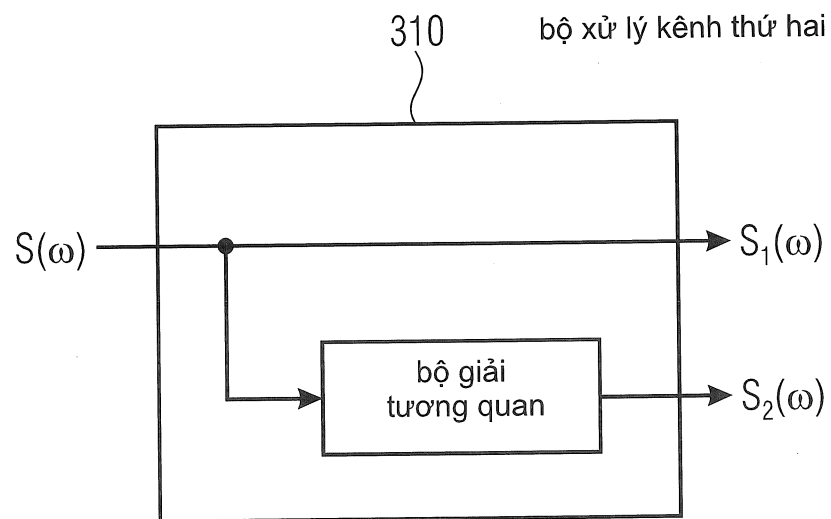


Fig. 2

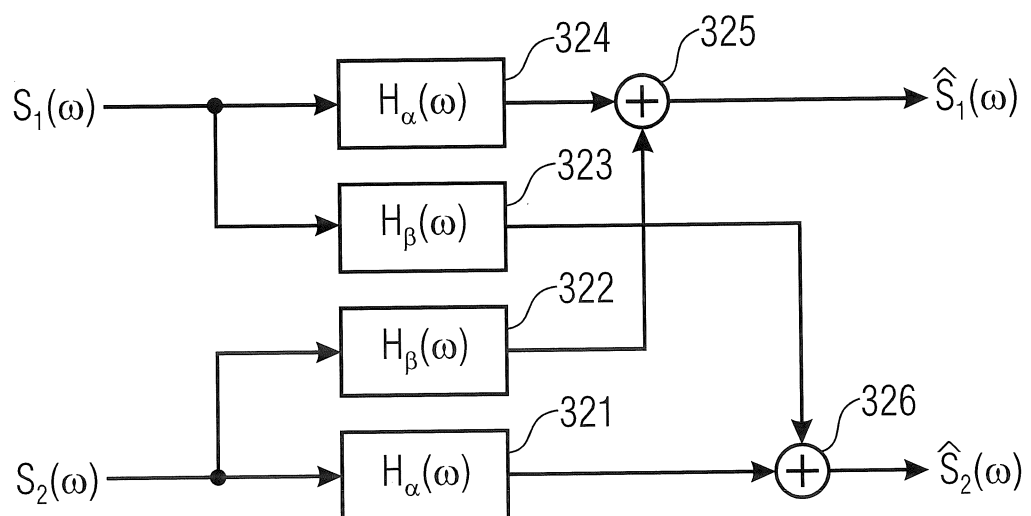


Fig. 3

4/18

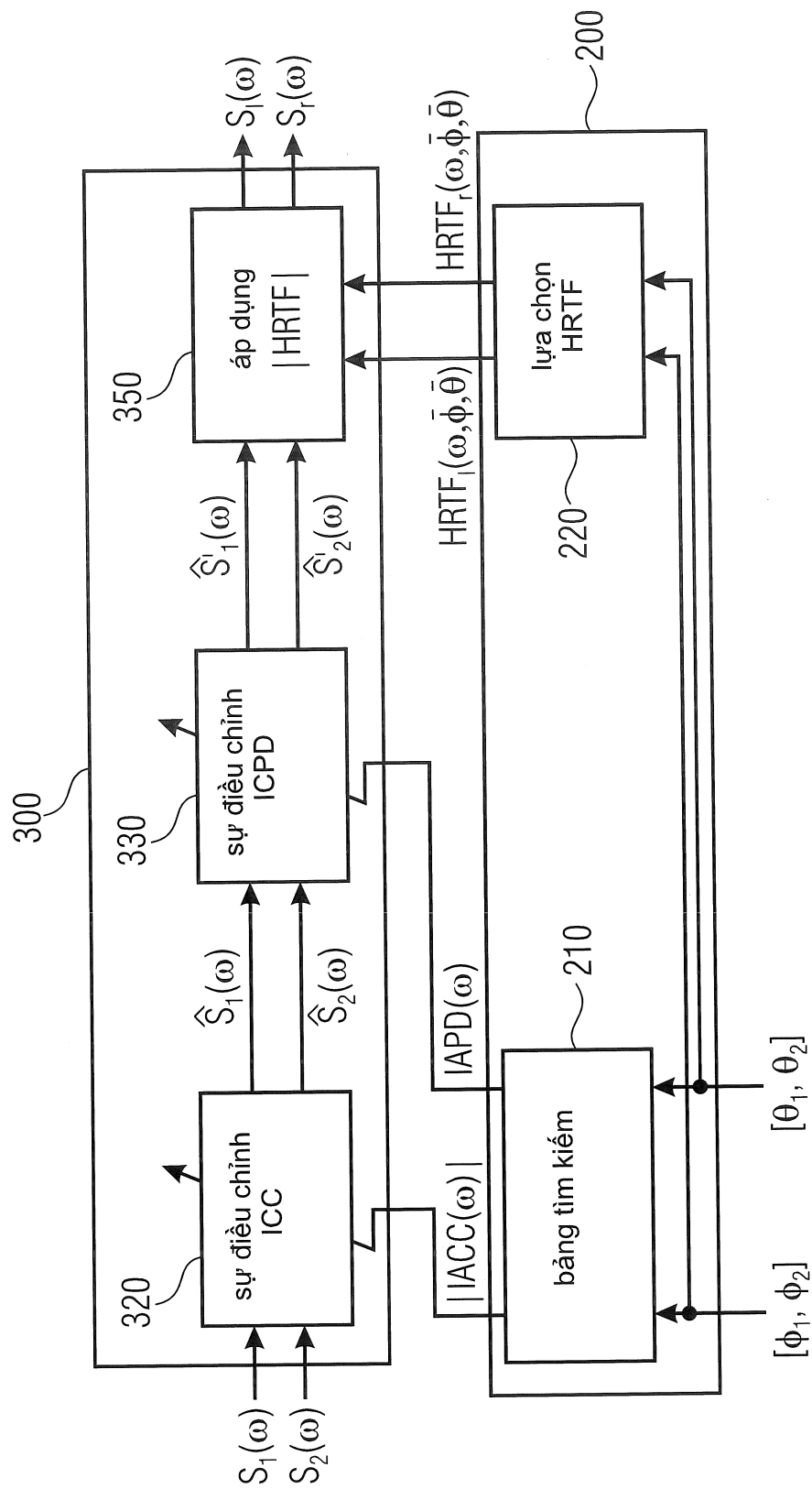


Fig. 4

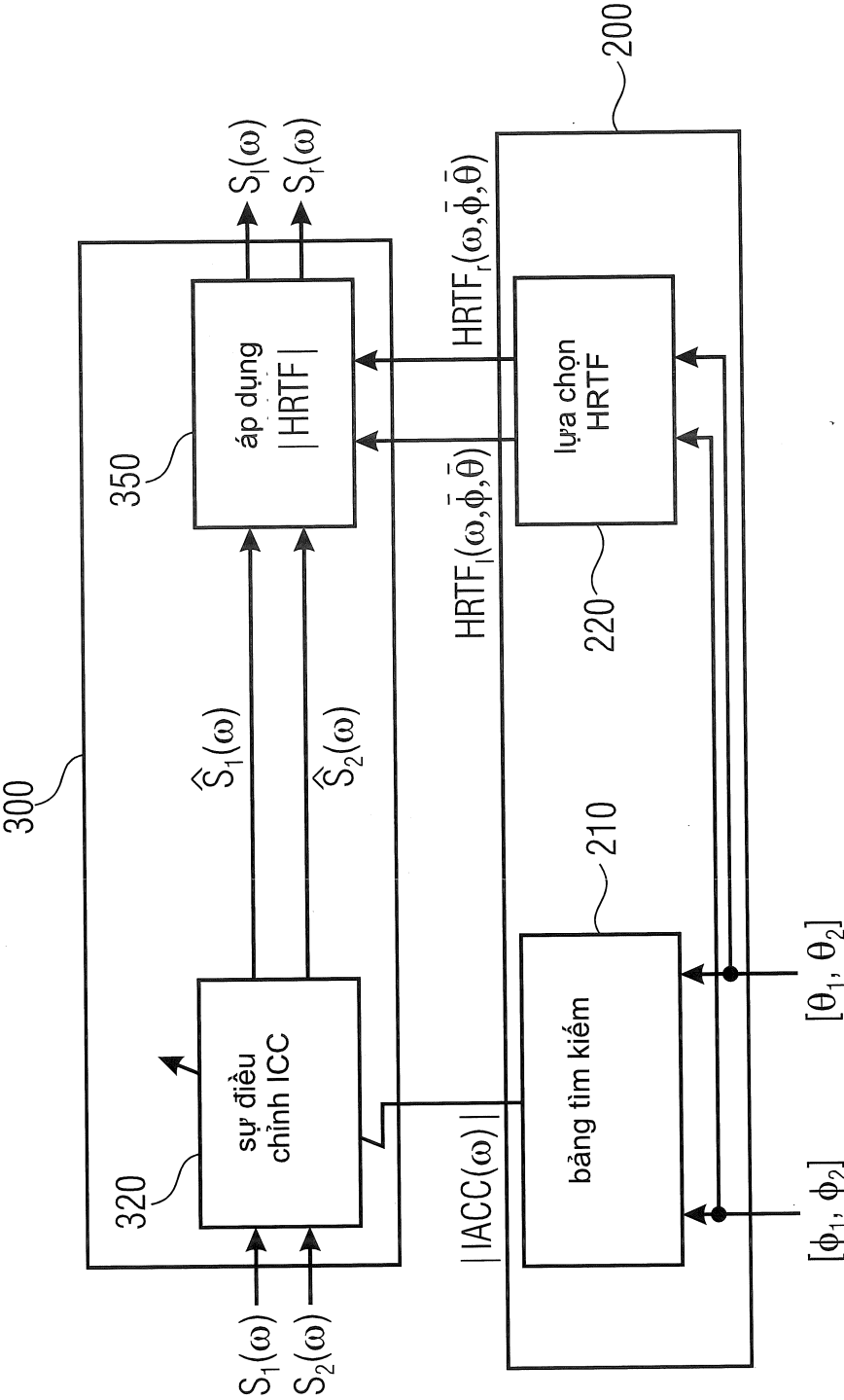
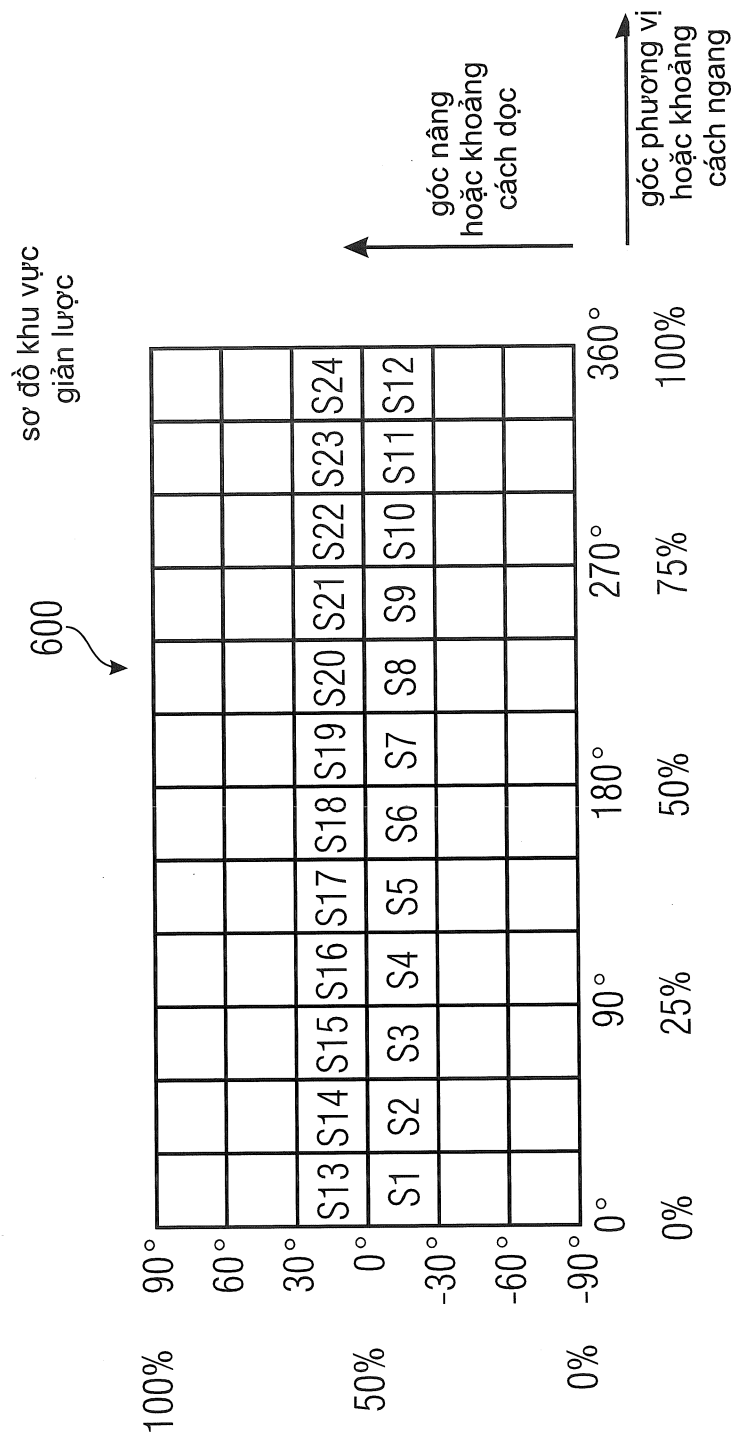


Fig. 5



e.g. S3: $\phi_{\text{range}} = 60^\circ - 90^\circ$
 $\theta_{\text{range}} = -30^\circ - 0$

e.g. S5: phạm vi ngang = 33% - 42%
phạm vi dọc = 33% - 50%

Fig. 6

7/18

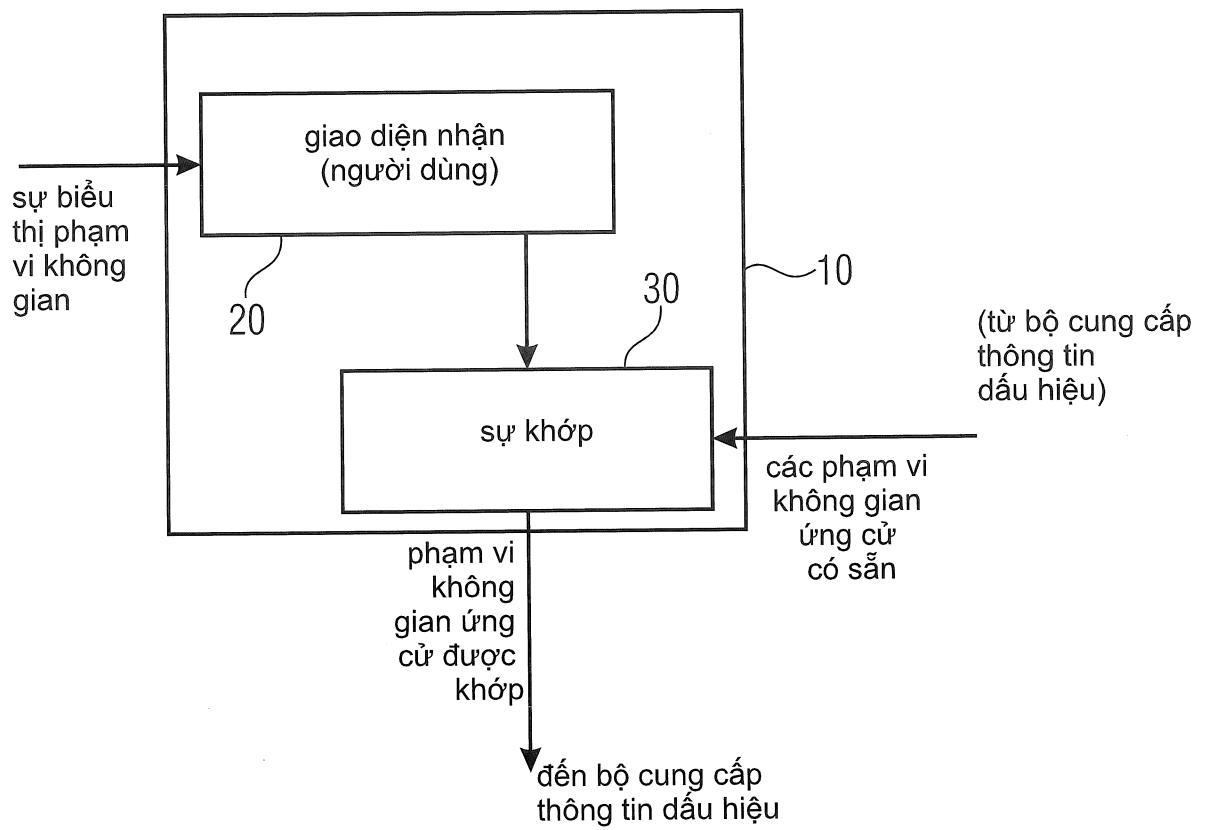


Fig. 7

8/18

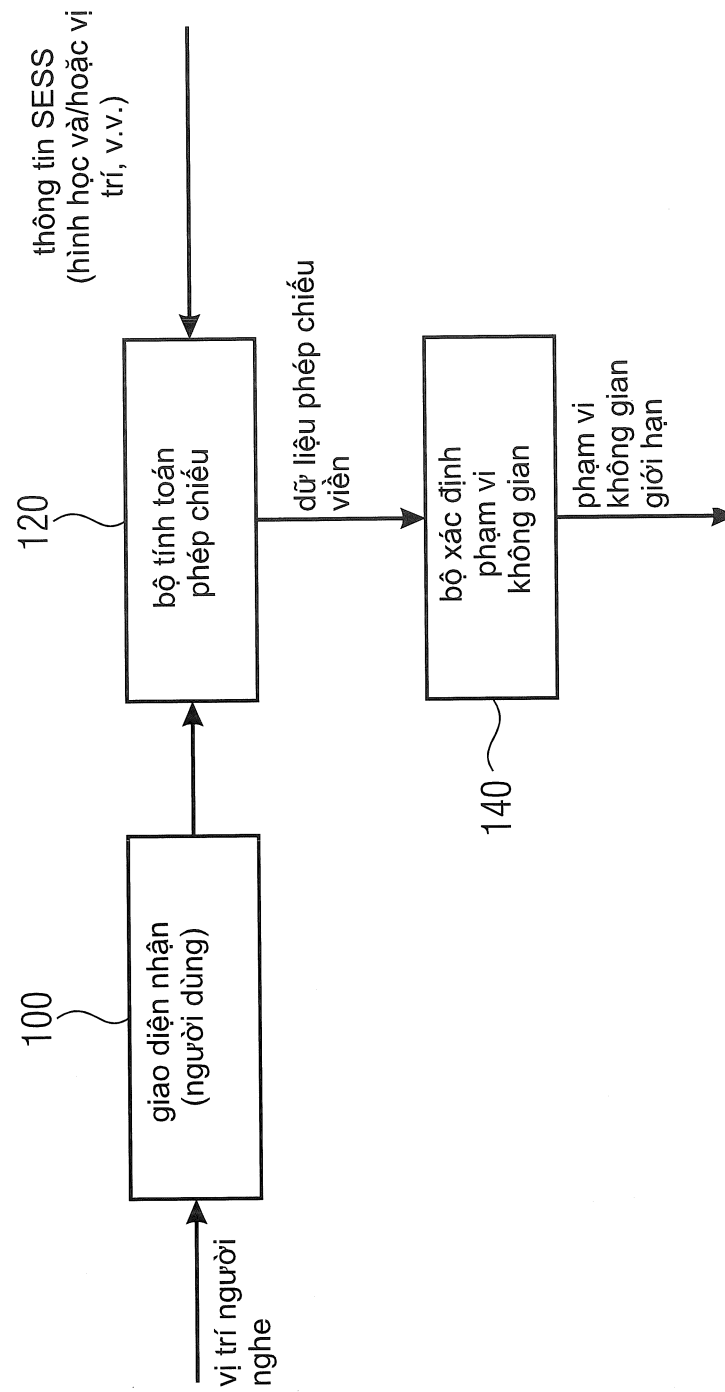


Fig. 8

9/18

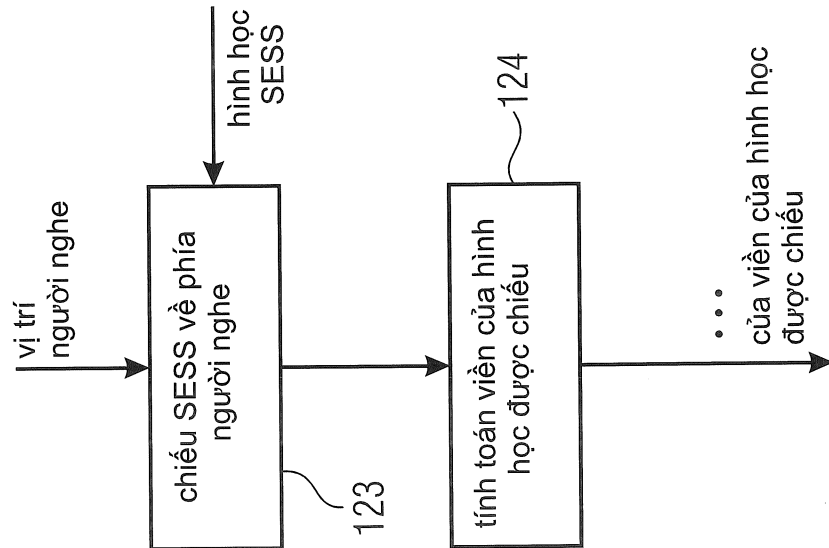


Fig. 9b

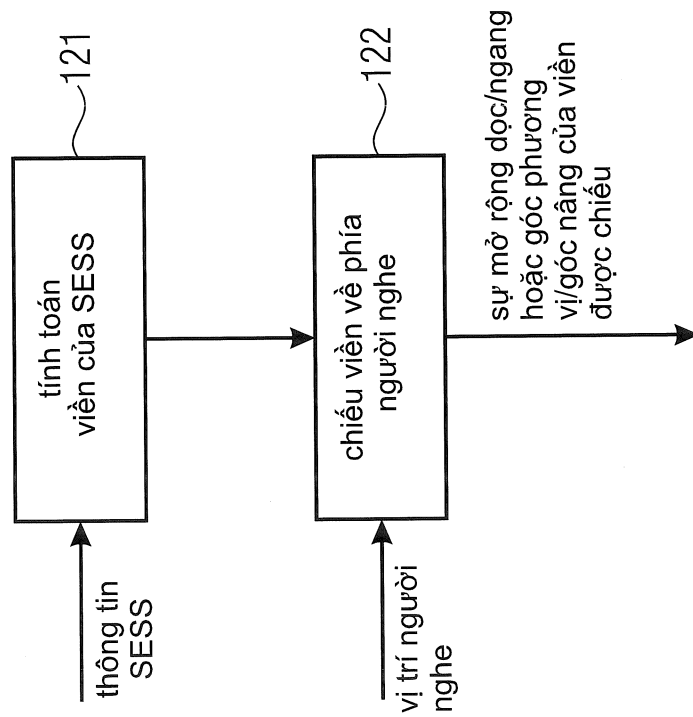


Fig. 9a

10/18

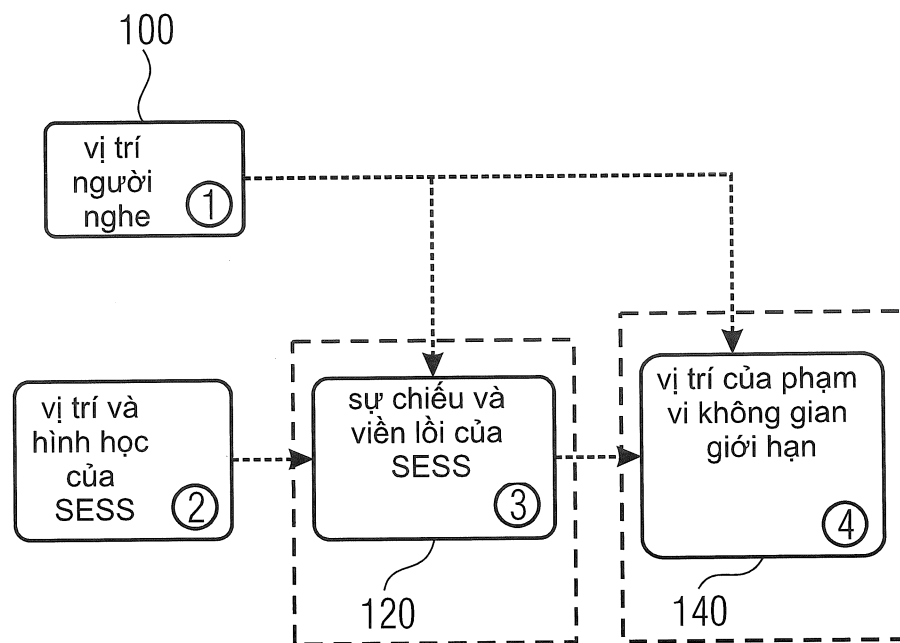


Fig. 10

11/18

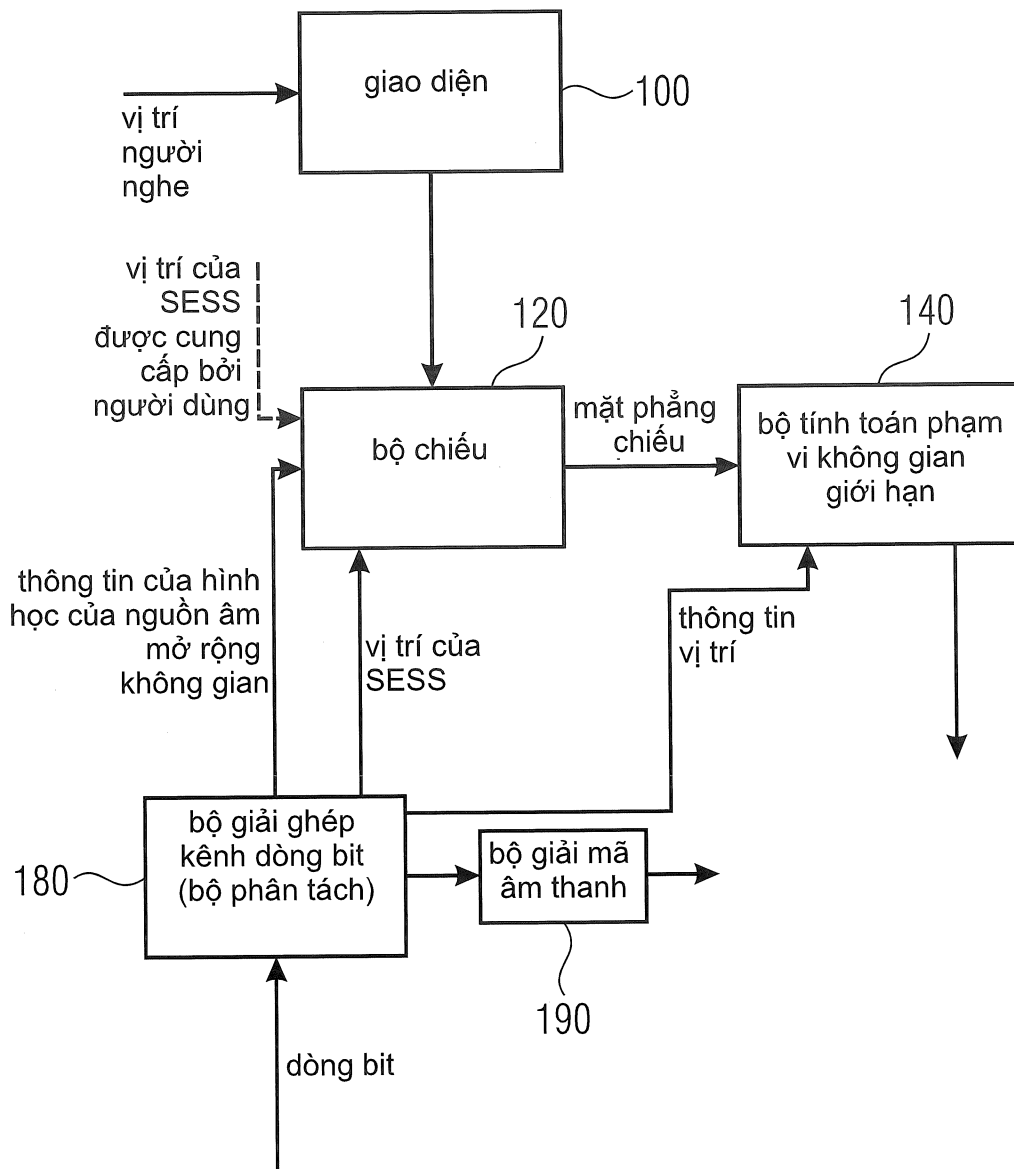
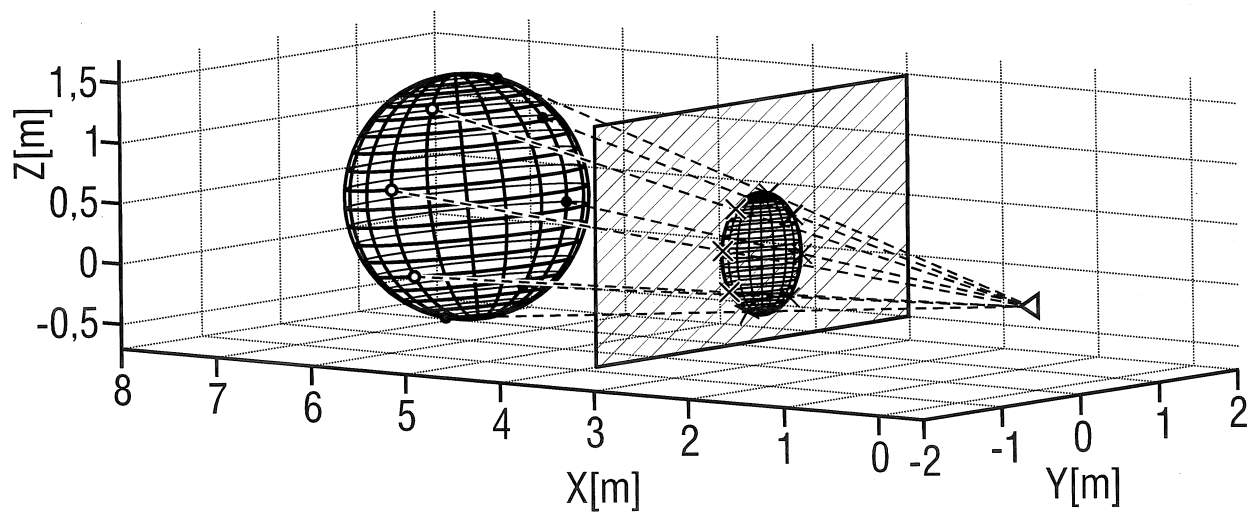
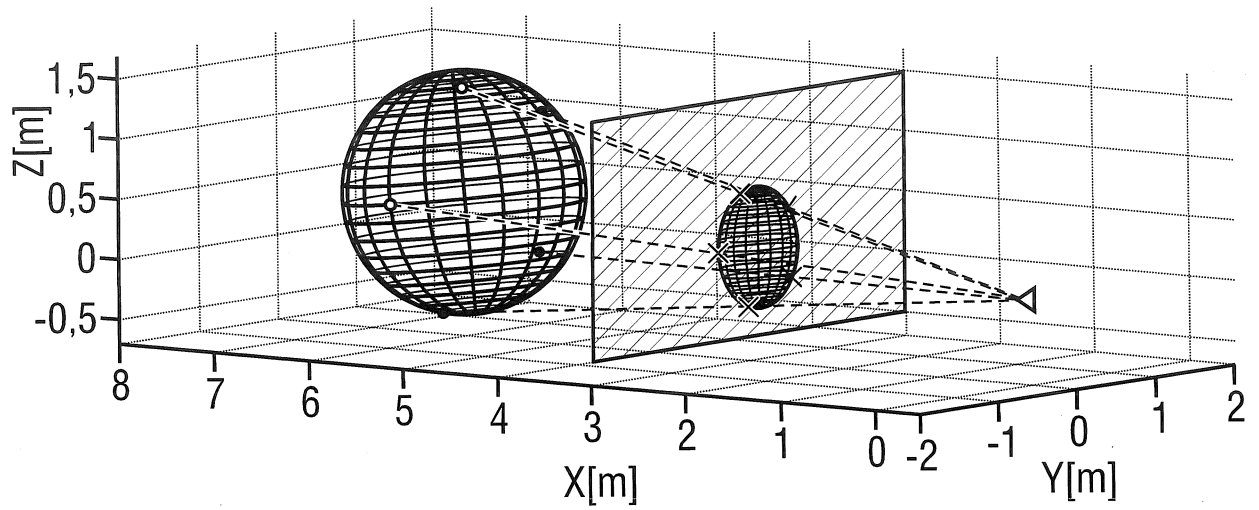
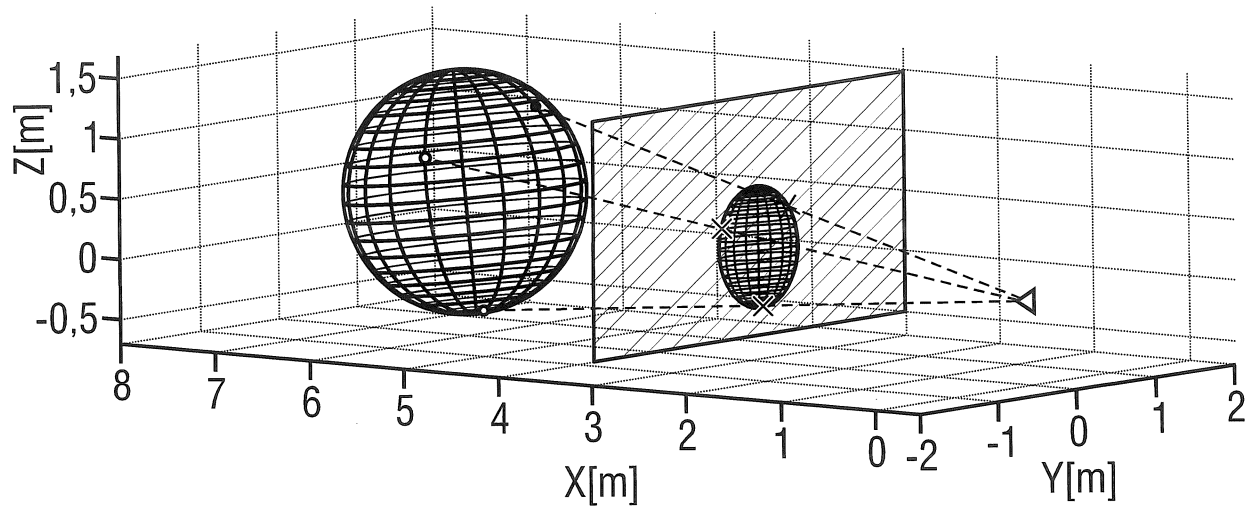


Fig. 11

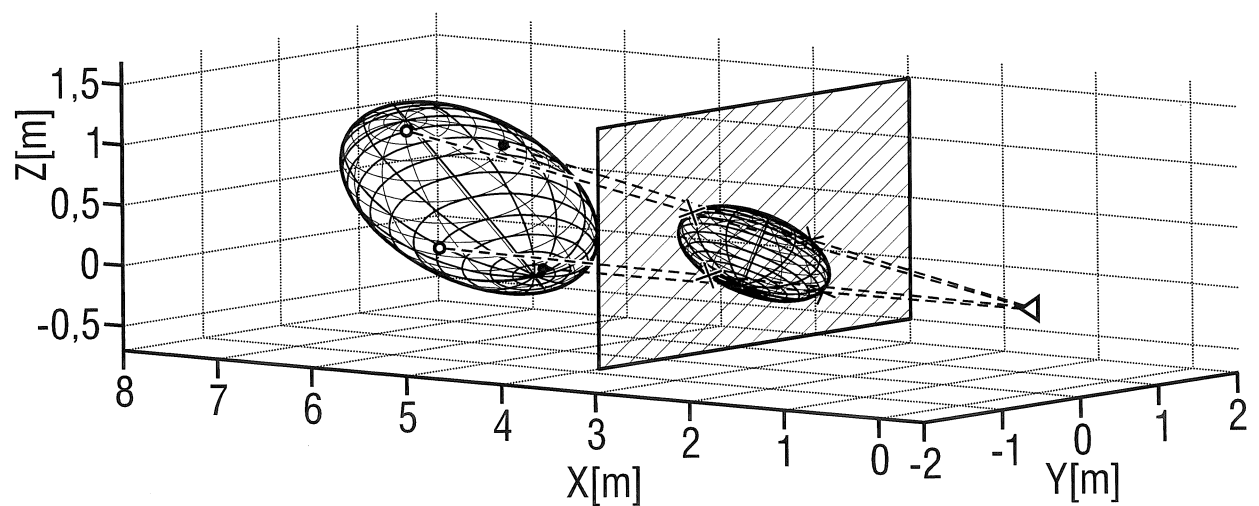
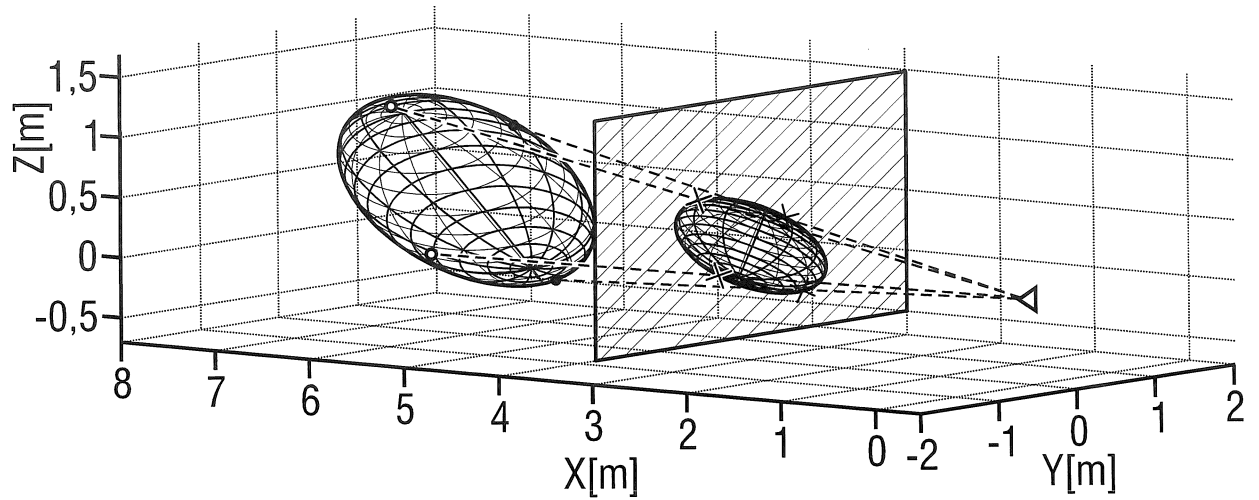
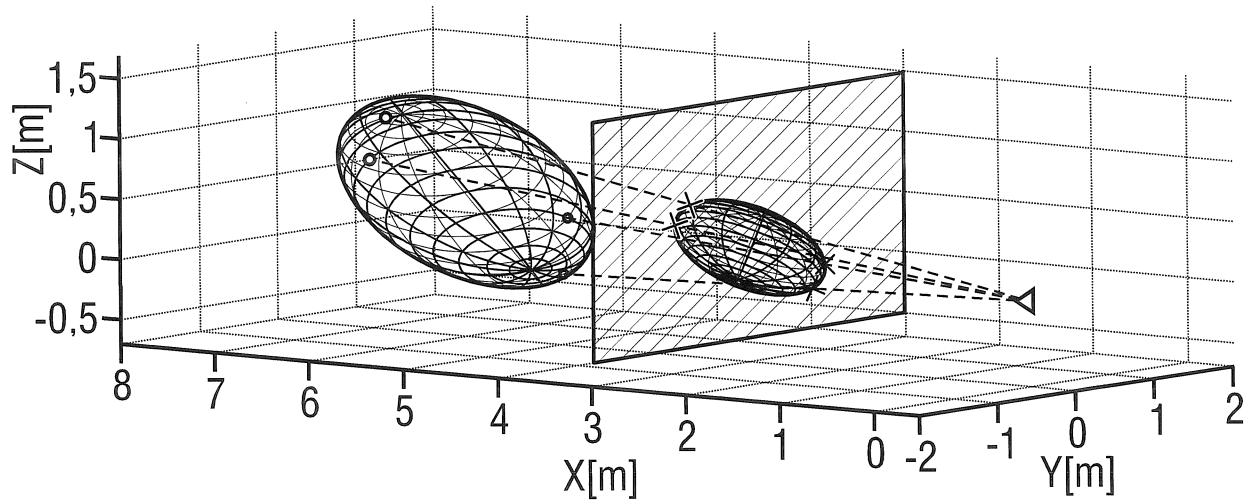
12/18



SESS hình cầu

Fig. 12

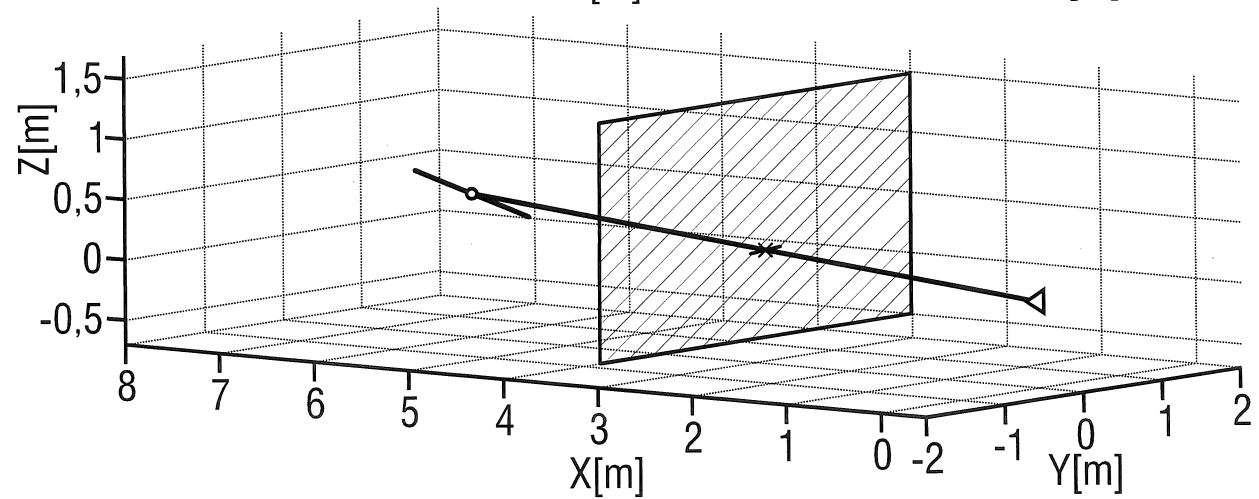
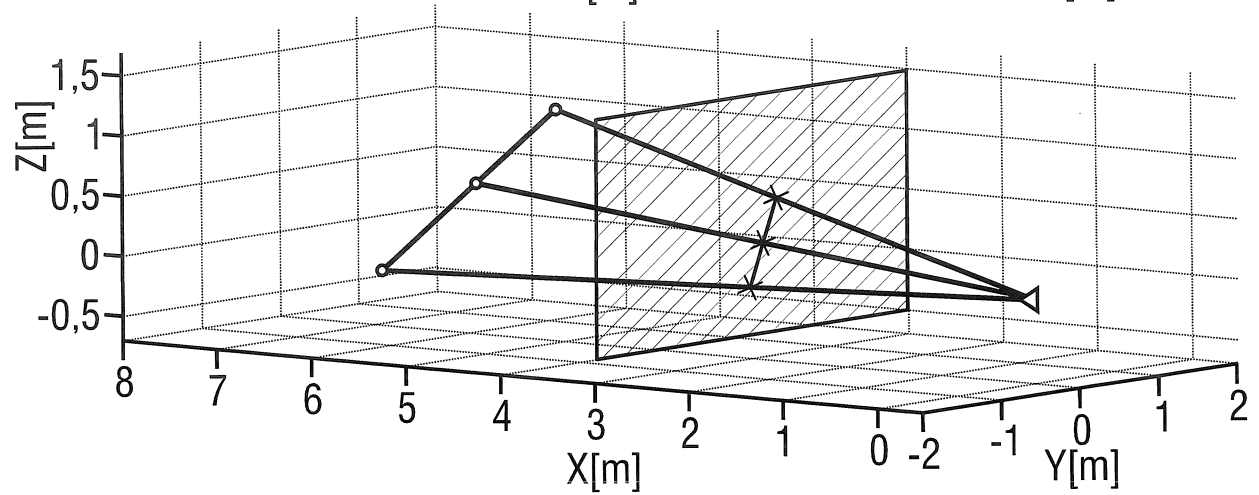
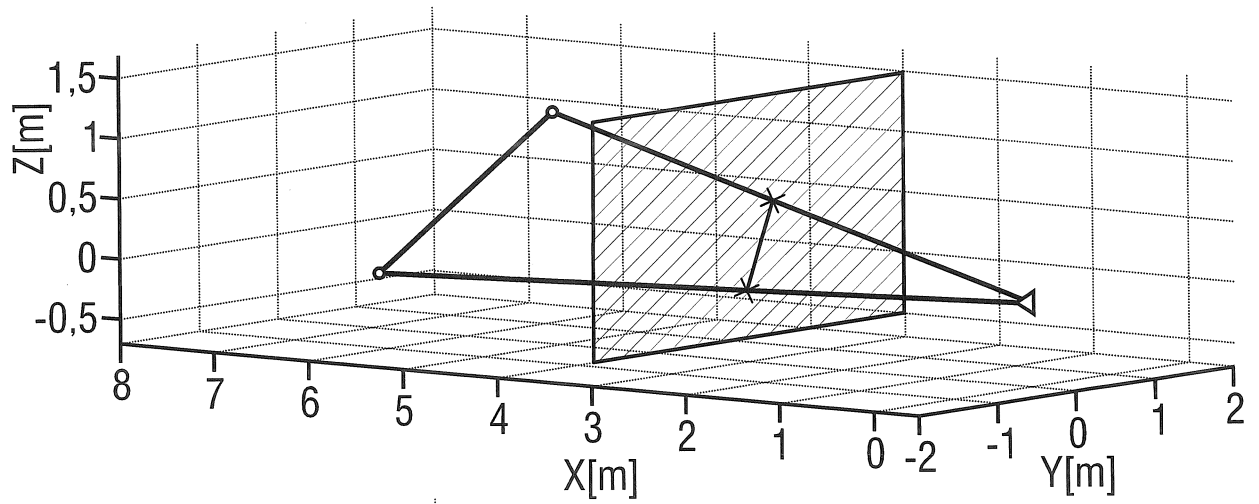
13/18



SESS hình elip

Fig. 13

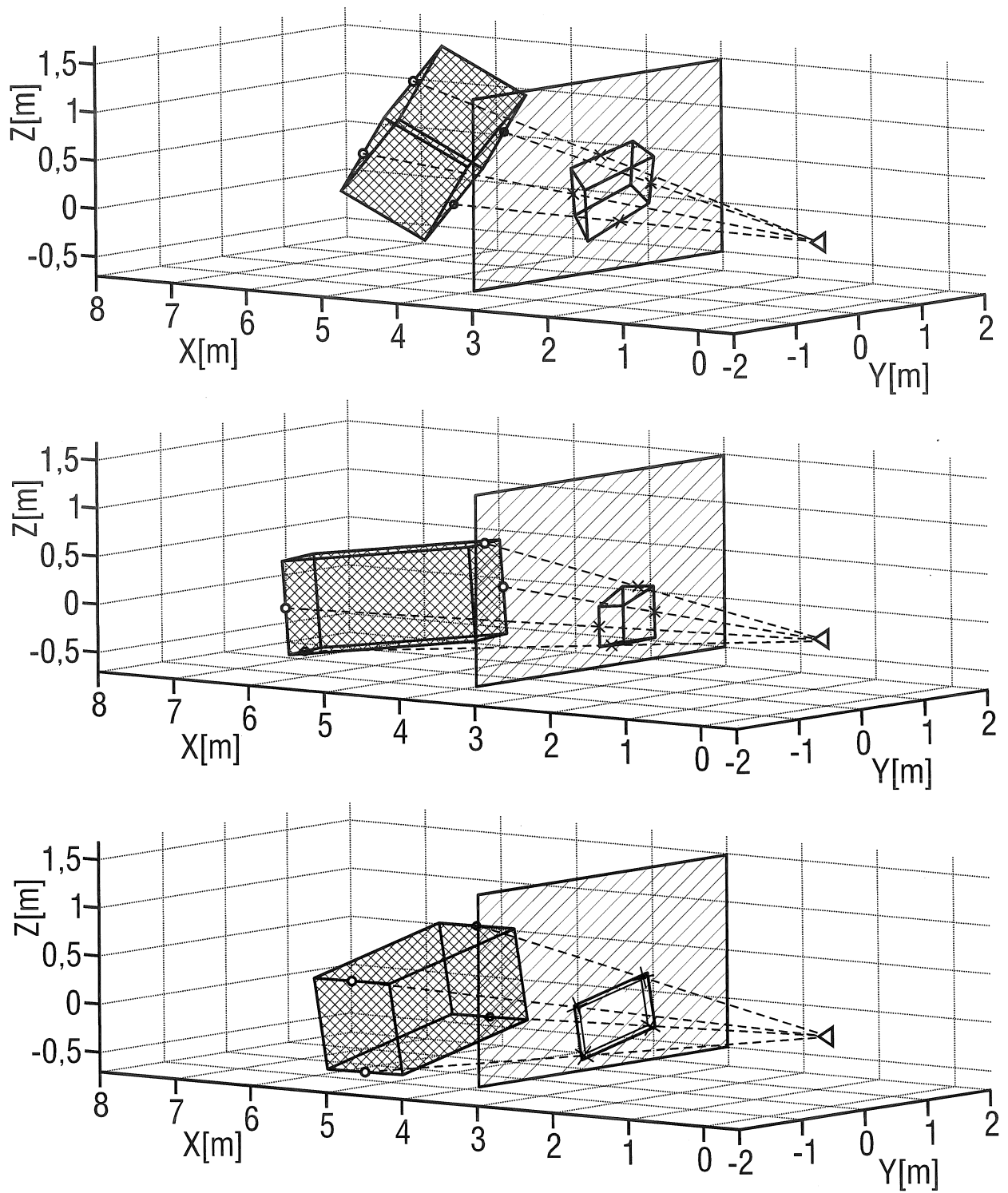
14/18



SESS thẳng

Fig. 14

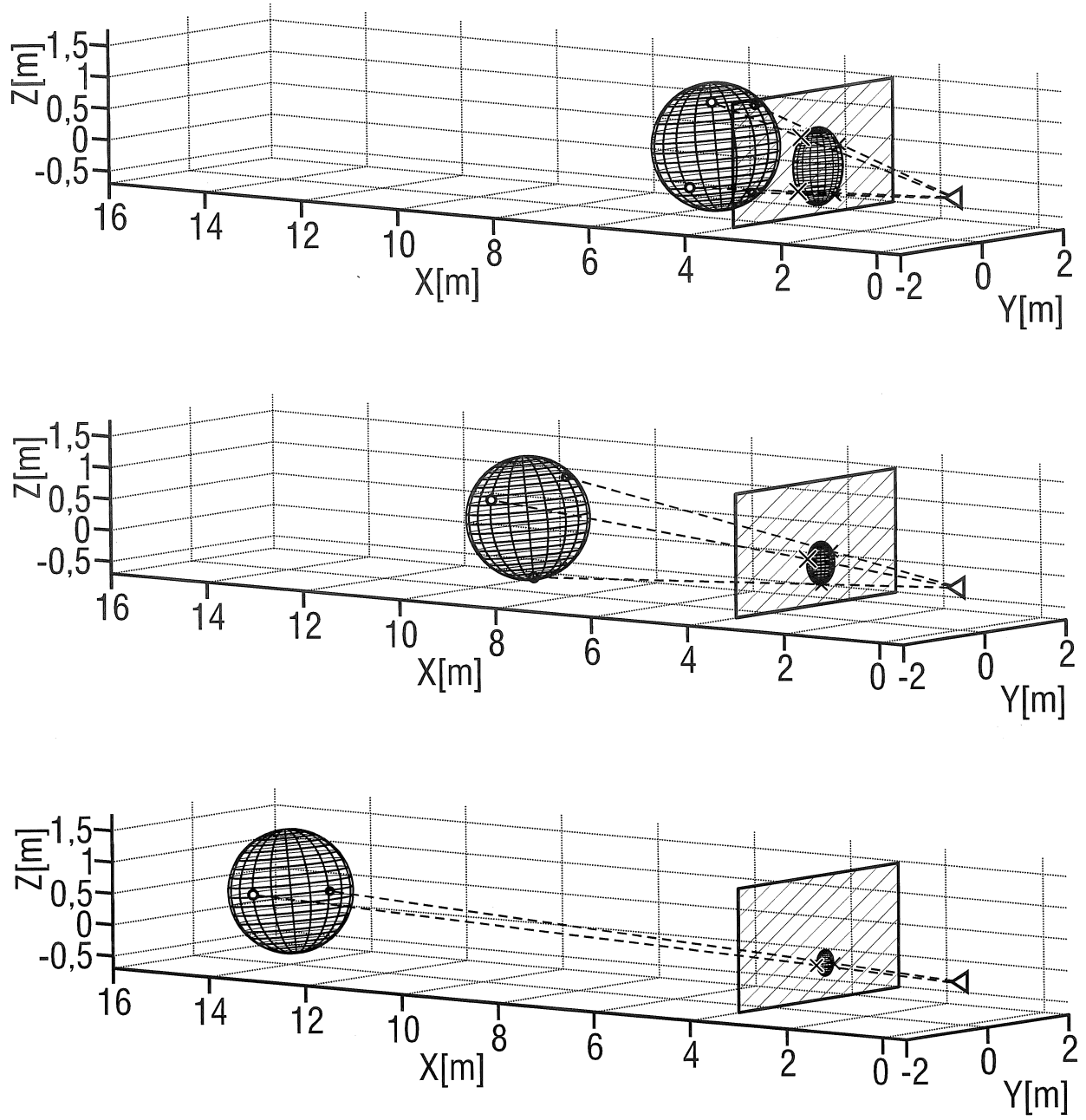
15/18



SESS hình lập
phương phẳng

Fig. 15

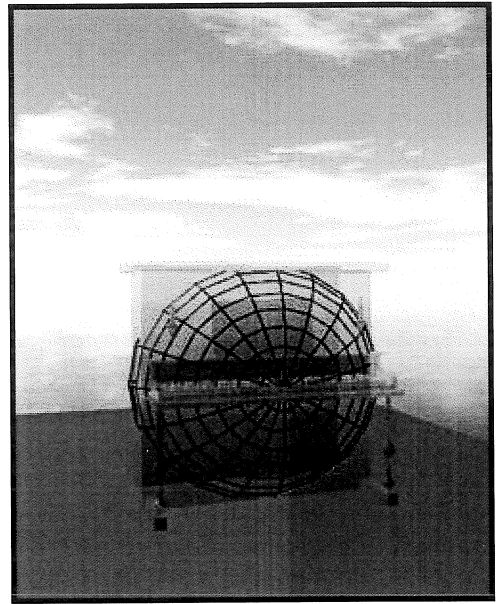
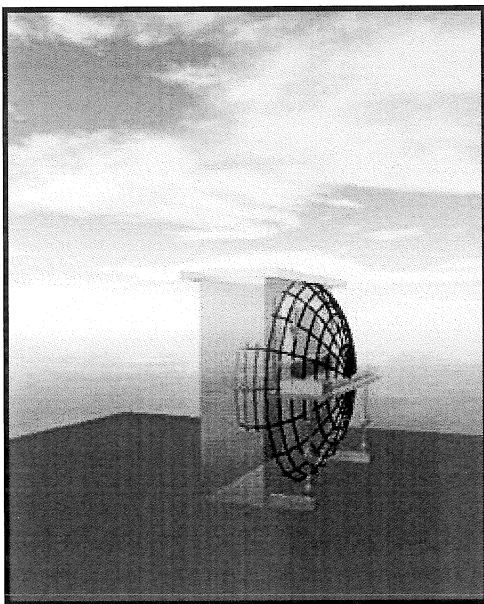
16/18



SESS hình cầu

Fig. 16

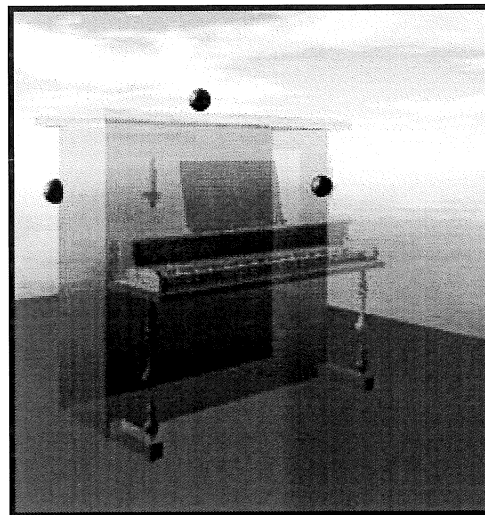
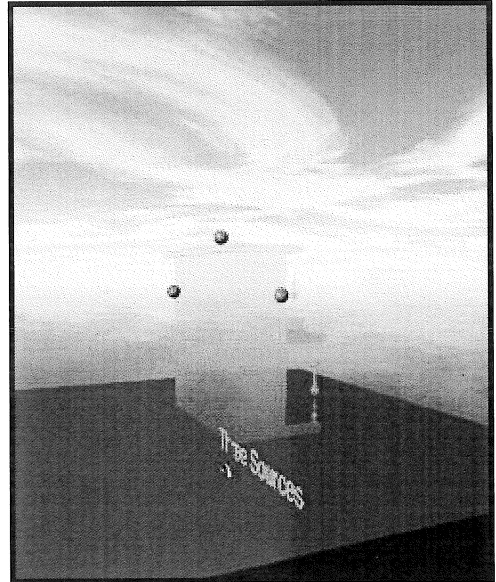
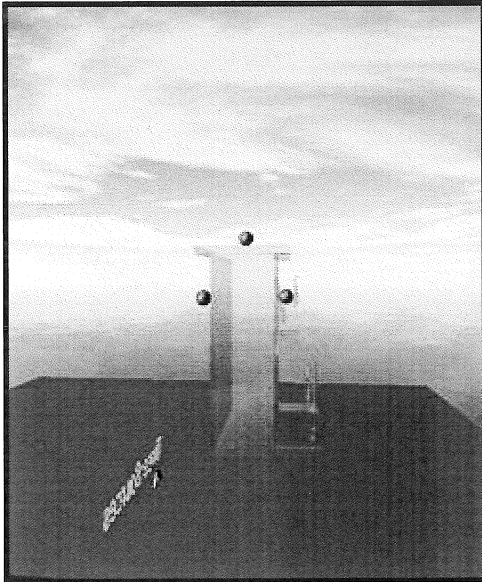
17/18



SESS dạng piano (được biểu thị bằng màu xanh lá cây) với dạng sóng elip tham số gần đúng (được biểu thị dạng lưới đỏ)

Fig. 17

18/18



SESS dạng đàn piano

Fig. 18