



(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)



1-0052606

(51)^{2020.01} G10L 19/005; G10L 19/008

(13) B

(21) 1-2022-00188

(22) 05/06/2020

(86) PCT/EP2020/065631 05/06/2020

(87) WO2020/249480 17/12/2020

(30) 19179750.5 12/06/2019 EP

(45) 27/10/2025 451

(43) 25/05/2022 410A

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

(72) FUCHS, Guillaume (FR); MULTRUS, Markus (DE); DÖHLA, Stefan (DE);
EICHENSEER, Andrea (DE).

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

(54) PHƯƠNG PHÁP CHE GIẤU TỔN HAO CÁC THAM SỐ ÂM THANH KHÔNG
GIAN VÀ PHƯƠNG PHÁP GIẢI MÃ NGỮ CẢNH ÂM THANH ĐƯỢC MÃ HÓA
ÂM THANH ĐỊNH HƯỚNG, THIẾT BỊ CHE GIẤU TỔN HAO, VÀ BỘ GIẢI MÃ

(21) 1-2022-00188

(57) Sáng chế đề cập đến phương pháp che giấu tổn hao các tham số âm thanh không gian và phương pháp giải mã ngữ cảnh âm thanh được mã hóa âm thanh định hướng, thiết bị che giấu tổn hao, và bộ giải mã. Phương pháp che giấu tổn hao các tham số âm thanh không gian, các tham số âm thanh không gian bao gồm ít nhất thông tin hướng tới, phương pháp bao gồm các bước sau: nhận tập hợp thứ nhất gồm các tham số không gian bao gồm ít nhất thông tin hướng tới thứ nhất; nhận tập hợp thứ hai gồm các tham số không gian bao gồm ít nhất thông tin hướng tới thứ hai; thay thế hướng thông tin tới thứ hai của tập hợp thứ hai bằng thông tin hướng tới thay thế được suy ra từ thông tin hướng tới thứ nhất, nếu ít nhất thông tin hướng tới thứ hai hoặc một phần thông tin hướng tới thứ hai bị mất hoặc hư hại.

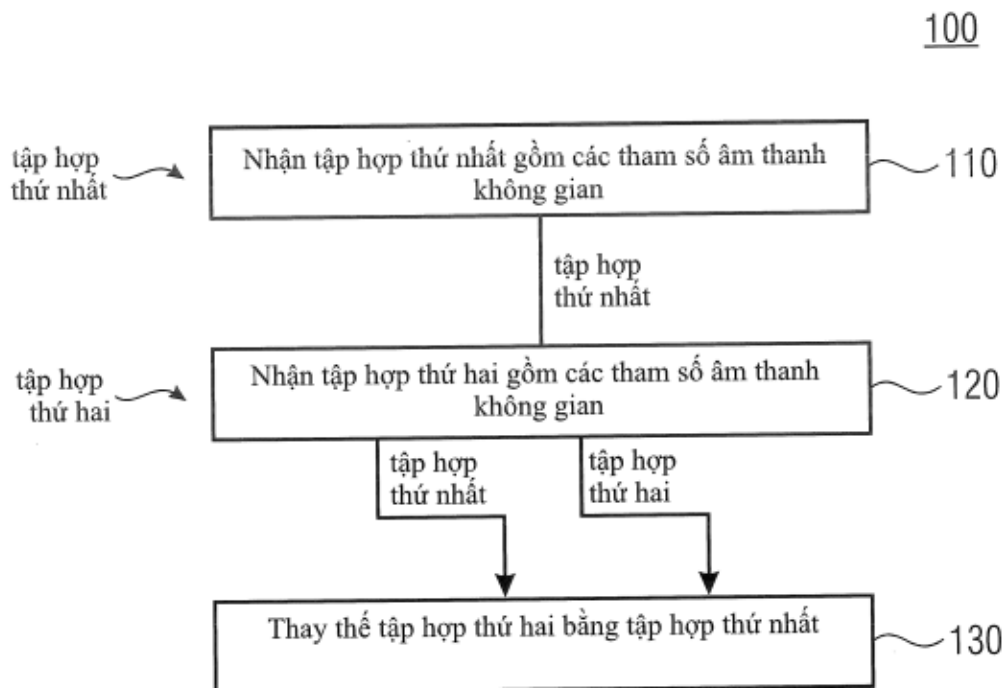


Fig. 3a

Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến phương pháp che giấu tổn hao các tham số âm thanh không gian, phương pháp giải mã ngữ cảnh âm thanh được mã hóa DirAC và các chương trình máy tính tương ứng. Các phương án của sáng chế đề cập đến thiết bị che giấu tổn hao cho sự che giấu lỗi của các tham số âm thanh không gian và đến bộ giải mã bao gồm thiết bị che giấu lỗi gói. Các phương án được ưu tiên mô tả khái niệm/phương pháp giảm cấp chất lượng bù do các khung hoặc các gói tổn hao và sai lệch xảy ra khi truyền dẫn ngữ cảnh âm thanh cho ảnh không gian được mã hóa theo tham số bởi hệ biến hóa mã hóa âm thanh định hướng (directional audio coding - DirAC).

Tình trạng kỹ thuật của sáng chế

Sự truyền thông tiếng nói và âm thanh có thể gặp các vấn đề chất lượng khác nhau do sự tổn hao gói trong khi truyền dẫn. Các điều kiện thực sự xấu trong mạng, như các lỗi bit và chập chờn, có thể dẫn đến sự tổn hao một số gói. Các tổn hao này dẫn đến một số thành phần lạ nghiêm trọng, như các tiếng lách cách, tiếng rơi tõm hoặc các sự im lặng không mong muốn mà làm suy giảm đáng kể chất lượng được nhận biết của tiếng nói được khôi phục hoặc tín hiệu âm thanh tại phía bộ nhận. Để giải quyết sự ảnh hưởng bất lợi của tổn hao gói, các thuật toán che giấu lỗi (packet loss concealment - PLC) được đề xuất trong các sơ đồ mã hóa tiếng nói và âm thanh thông thường. Các thuật toán như vậy thường vận hành tại phía bộ nhận bằng cách tạo ra tín hiệu âm thanh tổng hợp để che giấu dữ liệu bị mất trong dòng bit được nhận.

DirAC là công nghệ xử lý âm thanh không gian được thúc đẩy về cảm quan mà biểu diễn mạch lạc và hiệu quả trường âm bởi tập hợp các tham số không gian và tín hiệu trộn giảm. Tín hiệu trộn giảm có thể là các tín hiệu đơn âm, âm lập thể, hoặc đa kênh trong định dạng âm thanh như định dạng A, định dạng B, cũng được biết đến như Ambisonics thứ tự thứ nhất (first order Ambisonics - FAO). Tín hiệu trộn giảm được hoàn thiện bởi các tham số DirAC không gian mà mô tả ngữ cảnh âm thanh về hướng

tới (direction-of-arrival - DOA) và độ khuếch tán mỗi đơn vị thời gian/tần số. Trong bộ nhớ, các ứng dụng truyền theo dòng hoặc truyền thông, tín hiệu trộn giảm được mã hóa bởi bộ mã hóa lỗi thông thường (ví dụ, EVS hoặc sự mở rộng âm lập thể/đa kênh của EVS hoặc bất kỳ bộ mã hóa-giải mã đơn/lập thể/đa kênh), nhằm mục đích bảo toàn dạng sóng âm lập thể của mỗi kênh. Bộ mã hóa lỗi có thể được dựng quanh sơ đồ mã hóa trên cơ sở biến đổi hoặc sơ đồ mã hóa tiếng nói vận hành trong miền thời gian, như CELP. Bộ mã hóa lỗi có thể còn tích hợp các công cụ phục hồi lỗi có sẵn như các thuật toán che giấu tổn hao gói (packet loss concealment - PLC).

Mặt khác, không có giải pháp có sẵn để bảo vệ các tham số không gian DirAC. Do đó, có nhu cầu cho phương pháp được cải thiện.

Tài liệu tham khảo

- [1] V. Pulkki, M-V. Laitinen, J. Vilkkamo, J. Ahonen, T. Lokki, and T. Pihlajamäki, "Directional audio coding - perception-based reproduction of spatial sound", International Workshop on the Principles and Application on Spatial Hearing, Nov. 2009, Zao; Miyagi, Japan.
- [2] V. Pulkki, "Virtual source positioning using vector base amplitude panning", J. Audio Eng. Soc., 45(6):456-466, June 1997.
- [3] J. Ahonen and V. Pulkki, "Diffuseness estimation using temporal variation of intensity vectors", in Workshop on Applications of Signal Processing to Audio and Acoustics WASPAA, Mohonk Mountain House, New Paltz, 2009.
- [4] T. Hirvonen, J. Ahonen, and V. Pulkki, "Perceptual compression methods for metadata in Directional Audio Coding applied to audiovisual teleconference", AES 126th Convention 2009, May 7–10, Munich, Germany.
- [5] A. Politis, J. Vilkkamo and V. Pulkki, "Sector-Based Parametric Sound Field Reproduction in the Spherical Harmonic Domain," in *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 5, pp. 852-866, Aug. 2015.

Bản chất kỹ thuật của sáng chế

Mục đích của sáng chế là đề xuất nguyên lý cho sự che giấu lỗi trong ngữ cảnh DirAC.

Mục đích này đạt được bởi đối tượng của các điểm yêu cầu bảo hộ độc lập kèm theo.

Các phương án của sáng chế đề xuất phương pháp che giấu lỗi của các tham số âm thanh không gian, các tham số âm thanh không gian bao gồm ít nhất thông tin hướng tới. Phương pháp bao gồm các bước sau:

- nhận tập hợp thứ nhất gồm các tham số âm thanh không gian bao gồm thông tin hướng tới thứ nhất và thông tin độ khuếch tán thứ nhất;
- nhận tập hợp thứ hai gồm các tham số âm thanh không gian bao gồm thông tin hướng tới thứ hai và thông tin độ khuếch tán thứ hai;
- thay thế hướng thông tin tới thứ hai của tập hợp thứ hai bằng thông tin hướng tới thay thế được suy ra từ thông tin hướng tới thứ nhất nếu ít nhất thông tin hướng tới thứ hai hoặc một phần thông tin hướng hướng tới thứ hai bị mất.

Các phương án của sáng chế được dựa trên việc tìm kiếm rằng trường hợp tổn hao hoặc hư hại của thông tin tới, thông tin tới bị mất/bị hư hại có thể được thay thế bởi thông tin tới được suy ra từ thông tin tới có sẵn khác. Ví dụ, thông tin tới thứ hai bị mất, có thể thay thế bởi thông tin tới thứ nhất. Nói cách khác, điều này có nghĩa rằng phương án cung cấp sự che giấu tổn hao gói cho âm thanh tham số không gian mà thông tin định hướng là trường hợp tổn hao truyền dẫn được khôi phục bằng cách sử dụng thông tin định hướng được nhận tốt và sự hòa sắc. Do đó, các phương án cho phép giải quyết sự tổn hao gói trong sự truyền dẫn âm thanh không gian được mã hóa với các tham số trực tiếp.

Các phương án khác đề xuất phương pháp, trong đó các tập hợp thứ nhất và thứ hai gồm các tham số âm thanh không gian lần lượt bao gồm thông tin độ khuếch tán thứ nhất và thứ hai. Trong trường hợp này, chiến lược có thể như sau: theo các phương án, thông tin độ khuếch tán thứ nhất hoặc thứ hai được suy ra từ ít nhất một tỉ lệ năng lượng tương ứng với ít nhất một thông tin hướng tới. Theo các phương án, phương pháp còn bao gồm thay thế thông tin độ khuếch tán thứ hai của tập hợp thứ hai bằng thông tin độ khuếch tán thay thế được suy ra từ thông tin độ khuếch tán thứ nhất. Đây là một phần của chiến lược được gọi là giữ dựa trên giả định rằng độ khuếch tán không thay đổi

nhiều giữa hai khung. Vì lý do này, phương pháp đơn giản, nhưng hiệu quả để giữ các tham số của khung được nhận tốt cuối cùng cho các khung bị mất trong khi truyền dẫn. Một phần khác của toàn bộ chiến lược này là để thay thế thông tin tới thứ hai bằng thông tin tới thứ nhất, trong đó đã được thảo luận trong ngữ cảnh của phương án cơ bản. Nhìn chung an toàn để xem xét rằng ảnh không gian phải ổn định tương đối theo thời gian, mà có thể được tịnh tiến cho các tham số DirAC, tức là, hướng tới mà giả định cũng không thay đổi nhiều giữa các khung.

Theo các phương án khác, thông tin hướng tới thay thế phù hợp với thông tin hướng tới thứ nhất. Trong trường hợp như vậy, chiến lược được gọi là hòa sắc của hướng có thể được sử dụng. Ở đây bước thay thế có thể, theo các phương án, bao gồm bước hòa sắc thông tin hướng tới thay thế. Ngoài ra hoặc cách khác, các bước thay thế có thể bao gồm đưa khi nhiều âm là thông tin hướng tới để thu được thông tin hướng tới thay thế. Sự hòa sắc có thể giúp trường âm được kết xuất tự nhiên hơn và hài hòa hơn bằng cách đưa vào nhiều âm ngẫu nhiên cho hướng trước đó trước khi sử dụng chúng cho cùng khung. Theo các phương án, bước đưa vào được thực hiện nếu thông tin độ khuếch tán thứ nhất hoặc thứ hai biểu thị độ khuếch tán cao. Cách khác, có thể thực hiện nếu thông tin độ khuếch tán thứ nhất hoặc thứ hai là trên ngưỡng được định trước cho thông tin độ khuếch tán biểu thị độ khuếch tán cao. Theo các phương án khác, thông tin độ khuếch tán bao gồm nhiều không gian trên tỉ lệ giữa các thành phần định hướng và không định hướng của ngữ cảnh âm thanh được mô tả bởi tập hợp thứ nhất và/hoặc tập hợp thứ hai gồm các tham số âm thanh không gian. Theo các phương án, nhiều âm ngẫu nhiên sẽ được đưa vào phụ thuộc vào thông tin độ khuếch tán thứ nhất và thứ hai. Cách khác, nhiều âm ngẫu nhiên sẽ được đưa vào được định tỉ lệ bởi thừa số phụ thuộc vào thông tin độ khuếch tán thứ nhất và/hoặc thứ hai. Do đó, theo phương án, phương pháp còn bao gồm bước phân tích âm điệu của ngữ cảnh âm thanh được mô tả bởi tập hợp thứ nhất và/hoặc thứ hai gồm các tham số âm thanh không gian hoặc phân tích âm điệu của sự trộn giảm được truyền dẫn thuộc về tập hợp thứ nhất và/hoặc thứ hai gồm tham số âm thanh không gian để thu được giá trị âm điệu mô tả âm điệu. Nhiều âm ngẫu nhiên sẽ được đưa vào khi đó phụ thuộc vào giá trị âm điệu. Theo các phương án, sự định tỉ lệ giảm được thực hiện bởi thừa số giảm cùng nhau với sự nghịch đảo giá trị âm điệu hoặc nếu âm điệu tăng.

Theo chiến lược khác, phương pháp bao gồm bước ngoại suy thông tin hướng tới thứ nhất để thu được hướng thay thế của thông tin tới có thể được sử dụng. Theo phương pháp này, có thể hình dung để ước lượng hướng của các sự kiện âm trong ngữ cảnh âm thanh để ngoại suy hướng được ước lượng. Điều này liên quan cụ thể nếu sự kiện âm được cục bộ hóa tốt trong không gian và nguồn điểm (mô hình trực tiếp có độ khuếch tán thấp). Theo các phương án, sự ngoại suy dựa trên một hoặc nhiều thông tin hướng tới bổ sung thuộc về một hoặc nhiều tập hợp gồm các tham số âm thanh không gian. Theo các phương án, phép ngoại suy được thực hiện, nếu thông tin độ khuếch tán thứ nhất hoặc thứ hai biểu thị độ khuếch tán thấp hoặc nếu thông tin độ khuếch tán thứ nhất và/hoặc thứ hai ở dưới ngưỡng được định trước cho thông tin độ khuếch tán.

Theo các phương án, tập hợp thứ nhất gồm các tham số âm thanh không gian thuộc về điểm thứ nhất theo thời gian và/hoặc khung thứ nhất, cả tập hợp thứ hai gồm các tham số âm thanh không gian thuộc về điểm thứ hai theo thời gian và/hoặc khung thứ hai. Cách khác, thời điểm thứ hai là tiếp theo với thời điểm thứ nhất hoặc khung thứ hai tiếp theo khung thứ nhất. Khi trở lại phương án trong đó hầu hết các tập hợp tham số âm thanh không gian được sử dụng cho phép ngoại suy, rõ ràng rằng có thể có nhiều tập hợp các tham số âm thanh không gian thuộc về nhiều thời điểm/khung ví dụ, tiếp theo lẫn nhau, được sử dụng.

Theo phương án khác, tập hợp thứ nhất gồm các tham số âm thanh không gian trước đó, trong đó tập con thứ nhất gồm các tham số âm thanh không gian cho băng tần thứ nhất và tập con thứ hai gồm các tham số âm thanh không gian cho băng tần thứ hai. Tập hợp thứ hai gồm các tham số âm thanh không gian trước đó, trong đó tập con thứ nhất khác gồm các tham số âm thanh không gian cho băng tần thứ nhất và tập con thứ hai khác gồm các tham số âm thanh không gian cho băng tần thứ hai.

Phương án khác đề xuất phương pháp giải mã ngữ cảnh âm thanh được mã hóa DirAC bao gồm sự trộn giảm, tập hợp thứ nhất gồm các tham số âm thanh không gian và tập hợp thứ hai gồm các tham số âm thanh không gian. Phương pháp này bao gồm các bước của phương pháp cho sự tổn hao khung như được thảo luận trên đây.

Theo các phương án, các phương pháp được thảo luận trên đây có thể được thực thi bằng máy tính. Do đó, phương án được đề cập đến vật ghi có thể đọc được bằng máy

tính chứa trên đó chương trình máy tính có mã chương trình để thực hiện, khi vận hành trên máy tính, có phương pháp theo một trong số các điểm nêu trên.

Phương án khác đề cập đến thiết bị che giấu tổn hao cho sự che giấu tổn hao các tham số âm thanh không gian (tham số tương tự bao gồm ít nhất một thông tin hướng tới). Thiết bị bao gồm bộ nhận và bộ xử lý. Bộ nhận được tạo cấu hình để nhận tập hợp thứ nhất gồm các tham số âm thanh không gian và tập hợp thứ hai gồm các tham số âm thanh không gian (nêu trên). Bộ xử lý được tạo cấu hình để thay thế hướng thông tin tới thứ hai của tập hợp thứ hai bằng thông tin hướng tới thay thế được suy ra từ thông tin hướng tới thứ nhất trong trường hợp thông tin hướng tới thứ hai bị mất hoặc hư hại. Phương án khác đề xuất bộ giải mã cho sơ đồ âm thanh được mã hóa DirAC bao gồm thiết bị che giấu tổn hao.

Mô tả vắn tắt các hình vẽ

Các phương án của sáng chế sẽ được thảo luận sau đây viện dẫn đến các hình vẽ đi kèm, trong đó

Các hình vẽ Fig.1a, Fig.1b thể hiện sơ đồ khối dạng giản lược minh họa sự phân tích và tổng hợp DirAC;

Fig.2 thể hiện sơ đồ khối chi tiết dạng giản lược của sự phân tích và tổng hợp DirAC trong bộ mã hóa âm thanh 3D tốc độ bit thấp hơn;

Fig.3a thể hiện lưu đồ giản lược của phương pháp che giấu lỗi theo phương án cơ bản;

Fig.3b thể hiện thiết bị che giấu lỗi dạng giản lược theo phương án cơ bản;

Các hình vẽ Fig.4a, Fig.4b thể hiện các sơ đồ giản lược của các hàm độ khuếch tán được đo của DDR (Fig.4a kích thước cửa sổ $W = 16$, Fig.4b kích thước cửa sổ $W = 512$) để minh họa phương án;

Fig.5 thể hiện sơ đồ giản lược của hướng được đo (phương vị và nâng) trong hàm của độ khuếch tán để minh họa các phương án;

Fig.6a thể hiện lưu đồ giản lược của phương pháp giải mã ngữ cảnh âm thanh được mã hóa DirAC theo các phương án; và

Fig.6b thể hiện sơ đồ khối dạng giản lược của bộ giải mã cho ngữ cảnh âm thanh được mã hóa DirAC theo phương án.

Sau đây, các phương án của sáng chế sẽ được thảo luận sau đây viện dẫn đến các hình vẽ đi kèm, trong đó các số tham chiếu giống nhau được cung cấp cho các vật thể/phần tử có chức năng giống nhau hoặc tương tự, để phần mô tả của nó có thể sử dụng và thay đổi lẫn nhau. Trước khi thảo luận các phương án của sáng chế cụ thể, đưa ra sự giới thiệu DirAC.

Mô tả chi tiết sáng chế

Giới thiệu về DirAC: DirAC là sự tái tạo âm không gian được thúc đẩy cảm quan. Giả định rằng tại một nấc thời gian và đối với một băng tới hạn, độ phân giải không gian của hệ thống nghe bị hạn chế để giải mã một tín hiệu điều khiển cho hướng và một tín hiệu điều khiển khác cho sự nhất quán liên thính.

Dựa trên các giả định này, DirAC biểu diễn âm không gian trong một băng tần bằng cách giao chéo hai dòng: dòng khuếch tán không định hướng và dòng không khuếch tán định hướng. Sự xử lý DirAC được thực hiện trong hai pha:

Pha thứ nhất là sự phân tích như được minh họa bởi Fig.1a và pha thứ hai là sự tổng hợp như được minh họa bởi Fig.1b.

Fig.1a thể hiện giai đoạn phân tích 10 bao gồm một hoặc nhiều bộ lọc băng thông 12a-n nhận các tín hiệu micro W, X, Y, và Z, giai đoạn phân tích 14e cho năng lượng và 14i cho mật độ. Bằng việc sử dụng sự sắp xếp theo thời gian độ khuếch tán Ψ (số tham chiếu 16d) có thể được xác định. Độ khuếch tán Ψ được xác định dựa trên sự phân tích năng lượng 14c và mật độ 14i. Dựa trên mật độ và sự phân tích 14i, hướng 16e có thể được xác định. Kết quả của sự xác định hướng là góc phương vị và góc nâng. Ψ , azi và ele được xuất ra như siêu dữ liệu. Các siêu dữ liệu này được sử dụng bởi thực thể tổng hợp 20 được thể hiện bởi Fig.1b.

Thực thể tổng hợp 20 như được thể hiện bởi Fig.1b bao gồm dòng thứ nhất 22a và dòng thứ hai 22b. Dòng thứ nhất bao gồm nhiều bộ lọc băng thông 12a-n và thực thể tính toán cho các micro ảo 24. Dòng thứ hai 22b bao gồm các giá trị trung bình để xử lý siêu dữ liệu, cụ thể, 26 cho tham số độ khuếch tán và 27 cho tham số hướng. Hơn nữa,

bộ giải tương quan 28 được sử dụng trong giai đoạn tổng hợp 20, trong đó thực thể giải tương quan 28 nhận dữ liệu của hai dòng 22a, 22b. Đầu ra của bộ giải tương quan 28 có thể được cấp cho các loa phát thanh 29.

Trong giai đoạn phân tích DirAC, micro trùng nhau thứ tự thứ nhất ở định dạng B được xem xét như đầu vào và độ khuếch tán và hướng tới của âm được phân tích trong miền tần số.

Trong giai đoạn tổng hợp DirAC, âm được chia thành hai dòng, dòng không khuếch tán và dòng khuếch tán. Dòng không khuếch tán được tái tạo như các nguồn điểm sử dụng sự quét biên độ, mà có thể được thực hiện bằng cách sử dụng sự quét biên độ góc vectơ (vector base amplitude panning - VBAP) [2]. Dòng khuếch tán phụ trách sự cảm biến của sự phát triển và được tạo ra bằng cách truyền tải đến các loa phát thanh các tín hiệu được giải tương quan lẫn nhau.

Các tham số DirAC, cũng được gọi là siêu dữ liệu không gian hoặc siêu dữ liệu DirAC sau đây, bao gồm các bộ dữ liệu độ khuếch tán và hướng. Các hướng có thể được biểu diễn trong tọa độ hình cầu bởi hai góc, góc phương vị và góc nâng, trong khi độ khuếch tán là thừa số vô hướng giữa 0 và 1.

Sau đây, hệ thống của sự mã hóa âm thanh không gian DirAC sẽ được thảo luận viện dẫn đến Fig.2. Fig.2 thể hiện sự phân tích DirAC hai giai đoạn 10' và sự phân tích DirAC 20'. Sau đây, sự phân tích DirAC bao gồm sự phân tích giàn lọc 12, bộ ước lượng hướng 16i và bộ ước lượng độ khuếch tán 16d. Cả hai, 16i và 16d xuất ra dữ liệu độ khuếch tán/hướng như siêu dữ liệu không gian. Dữ liệu này có thể được mã hóa sử dụng bộ mã hóa 17. Sự phân tích trực tiếp 20' bao gồm bộ giải mã siêu dữ liệu không gian 21, sự tổng hợp đầu ra 23, sự tổng hợp giàn lọc 12 cho phép xuất ra tín hiệu đến loa phát thanh FOA/HOA.

Song song với giai đoạn phân tích trực tiếp 10' và giai đoạn tổng hợp trực tiếp 20' được thảo luận, mà xử lý siêu dữ liệu không gian, bộ mã hóa/bộ giải mã EVS được sử dụng. Trên phía bộ phân tích, việc tạo chùm/lựa chọn tín hiệu được thực hiện dựa trên định dạng B tín hiệu đầu vào (xem thực thể tạo chùm/lựa chọn tín hiệu 15). Tín hiệu khi đó được mã hóa EVS (xem số tham chiếu 17). Tín hiệu khi đó được mã hóa EVS. Trên phía bộ tổng hợp (xem số tham chiếu 20'), bộ giải mã EVS 25 được sử dụng. Bộ giải

mã EVS này xuất ra tín hiệu đến sự phân tích giàn lọc 12, mà xuất ra các tín hiệu cho sự tổng hợp đầu ra 23.

Vì kết cấu của sự phân tích trực tiếp/sự tổng hợp trực tiếp 10'/20' đã được thảo luận, tính năng sẽ được thảo luận chi tiết.

Bộ mã hóa thường phân tích 10' ngữ cảnh âm thanh không gian trong định dạng B. Cách khác, sự phân tích DirAC có thể được điều chỉnh để phân tích các định dạng âm thanh khác như các đối tượng âm thanh hoặc các tín hiệu đa kênh hoặc tổ hợp của bất kỳ định dạng âm thanh không gian. Sự phân tích DirAC trích sự biểu diễn tham số từ ngữ cảnh âm thanh đầu vào. Hướng tới (direction of arrival - DOA) và độ khuếch tán được đo mỗi đơn vị thời gian - tần số tạo ra các tham số. Sự phân tích DirAC được theo sau bởi bộ mã hóa siêu dữ liệu không gian, mà lượng tử hóa và mã hóa các tham số DirAC để thu được sự biểu diễn tham số tốc độ bit thấp.

Cùng với các tham số, tín hiệu trộn giảm được suy ra từ các nguồn hoặc các tín hiệu đầu vào âm thanh khác nhau được mã hóa để truyền dẫn bởi bộ mã hóa lõi âm thanh truyền thống. Trong phương án được ưu tiên, bộ mã hóa âm thanh EVS được ưu tiên để mã hóa tín hiệu trộn giảm, nhưng sáng chế không chỉ giới hạn ở bộ mã hóa lõi này và có thể được áp dụng cho bất kỳ bộ mã hóa lõi âm thanh. Tín hiệu trộn giảm bao gồm các kênh khác nhau, được gọi là các kênh vận chuyển: tín hiệu có thể, ví dụ, là bốn tín hiệu hệ số hợp thành tín hiệu định dạng B, cặp âm lập thể hoặc sự trộn giảm đơn âm lập thể phụ thuộc vào tốc độ bit đích. Các tham số không gian được mã hóa và dòng bit âm thanh được mã hóa được dồn kênh trước khi được truyền dẫn qua kênh truyền thông.

Trong bộ giải mã, các kênh vận chuyển được giải mã bởi bộ giải mã lõi, trong khi siêu dữ liệu DirAC được giải mã đầu tiên trước khi được vận chuyển với các kênh vận chuyển được giải mã đến sự tổng hợp DirAC. Sự tổng hợp DirAC sử dụng siêu dữ liệu được giải mã để điều khiển sự tái tạo dòng âm trực tiếp và hỗn hợp của chúng với dòng âm khuếch tán. Trường âm được tái tạo có thể được tái tạo trên lớp loa phát thanh tùy ý hoặc có thể được tạo ra trong định dạng Ambisonics (HOA/FOA) với thứ tự tùy ý.

Sự ước lượng tham số DirAC: Trong mỗi băng tần, hướng tới của âm cùng với độ khuếch tán của âm được ước lượng. Từ sự phân tích thời gian - tần số của các thành

phần định dạng B đầu vào áp suất $w^i(n), x^i(n), y^i(n), z^i(n)$, và véctor vận tốc có thể được xác định là:

$$P^i(n, k) = W^i(n, k)$$

$$U^i(n, k) = X^i(n, k)e_x + Y^i(n, k)e_y + Z^i(n, k)e_z$$

trong đó i là chỉ số của đầu vào và, k và n là các chỉ số thời gian và tần số của ô thời gian - tần số, và e_x, e_y, e_z biểu diễn các véctor đơn vị Đề-các. $P(n, k)$ và $U(n, k)$ được sử dụng để tính toán các tham số DirAC, cụ thể, DOA và độ khuếch tán qua sự tính toán của véctor mật độ:

$$I(k, n) = \frac{1}{2} \Re\{P(k, n) \cdot \overline{U(n, k)}\},$$

trong đó $\overline{(\cdot)}$ biểu thị phép liên hợp phức. Độ khuếch tán của trường âm được tổ hợp được cho bởi:

$$\psi(k, n) = 1 - \frac{\|E\{I(k, n)\}\|}{cE\{E(k, n)\}}$$

trong đó $E\{\cdot\}$ biểu thị toán tử tính trung bình theo thời gian, c là tốc độ âm và $E(k, n)$ là năng lượng trường âm được cho bởi:

$$E(n, k) = \frac{\rho_0}{4} \|U(n, k)\|^2 + \frac{1}{\rho_0 c^2} |P(n, k)|^2$$

Độ khuếch tán của trường âm được xác định là tỉ lệ giữa mật độ âm và mật độ năng lượng có các giá trị giữa 0 và 1.

Hướng tới (direction of arrival - DOA) được biểu diễn bằng véctor đơn vị $direction(n, k)$, được xác định như

$$direction(n, k) = - \frac{I(n, k)}{\|I(n, k)\|}$$

Hướng tới được xác định bởi sự phân tích năng lượng của đầu vào định dạng B và có thể được xác định như hướng trái ngược của véctor mật độ. Hướng được xác định trong hệ tọa độ Đề Các nhưng có thể dễ dàng được biến đổi trong hệ tọa độ hình cầu được xác định bởi bán kính thống nhất, góc phương vị và góc nâng.

Trong trường hợp truyền dẫn, các tham số cần để được truyền dẫn đến phía bộ nhận qua dòng bit. Đối với sự truyền dẫn nhanh qua mạng với dung lượng có hạn, dòng bit tốc độ bit thấp được ưu tiên mà có thể đạt được bằng cách thiết kế sơ đồ mã hóa hiệu quả cho các tham số DirAC. Có thể sử dụng cho các kỹ thuật minh họa như tạo nhóm băng tần bằng cách tính trung bình các tham số theo các băng tần khác nhau và/hoặc các đơn vị thời gian, dự báo, sự lượng tử hóa và sự mã hóa entropy khác nhau. Tại bộ giải mã, các tham số có thể được giải mã cho mỗi đơn vị thời gian/tần số (k,n) trong trường hợp không có lỗi xảy ra trong mạng. Tuy nhiên, nếu các điều kiện mạng không đủ tốt để đảm bảo sự truyền dẫn gói phù hợp, gói có thể bị mất trong khi truyền dẫn. Sáng chế nhằm mục đích cung cấp giải pháp trong trường hợp sau.

Ban đầu, DirAC không được dự định xử lý các tín hiệu ghi định dạng B, cũng được biết đến như các tín hiệu Ambisonics thứ tự thứ nhất. Tuy nhiên, sự phân tích có thể dễ dàng được mở rộng đến bất kỳ mảng micro nào tổ hợp các micro đẳng hướng hoặc định hướng. Trong trường hợp này, sáng chế vẫn liên quan vì bản chất các tham số DirAC không thay đổi.

Ngoài ra, các tham số DirAC, cũng được biết đến như siêu dữ liệu, có thể được tính toán trực tiếp trong khi xử lý tín hiệu micro trước khi được chuyển đến bộ mã hóa âm không gian. Hệ thống mã hóa không gian trên cơ sở DirAC khi đó được nạp trực tiếp bởi các tham số âm thanh không gian tương đương hoặc giống với các tham số DirAC ở dạng siêu dữ liệu và dạng sóng âm thanh của tín hiệu trộn giảm. DoA và độ khuếch tán có thể được suy ra dễ dàng mỗi băng tần từ siêu dữ liệu đầu vào. Định dạng đầu vào như vậy đôi khi được gọi là định dạng âm thanh không gian hỗ trợ siêu dữ liệu (Metadata-assisted spatial audio - MASA). MASA cho phép hệ thống bỏ qua sự đặc hiệu của các mảng micro và các thừa số tạo ra của chúng cần để tính toán các tham số không gian. Điều này sẽ được suy ra bên ngoài hệ thống mã hóa âm thanh không gian sử dụng sự xử lý cụ thể cho thiết bị mà kết hợp các micro.

Các phương án của sáng chế có thể sử dụng hệ thống mã hóa không gian như được minh họa bởi Fig.2, trong đó bộ mã hóa và bộ giải mã âm thanh không gian trên cơ sở DirAC được biểu thị. Các phương án có thể được thảo luận viện dẫn đến Fig.3a và Fig.3b, trong đó các mở rộng đến mô hình DirAC sẽ được thảo luận trước.

Mô hình DirAC có thể theo sáng chế cũng được mở rộng bởi các thành phần định hướng khác nhau với cùng ô thời gian/tần số. Có thể được mở rộng theo hai cách chính:

Sự mở rộng thứ nhất bao gồm gửi hai hoặc nhiều hơn hai DoA mỗi ô T/F. Mỗi DoA khi đó phải được kết hợp với năng lượng, hoặc tỉ lệ năng lượng. Ví dụ DOA thứ l có thể được kết hợp với tỉ lệ năng lượng Γ_l giữa năng lượng của thành phần định hướng và toàn bộ năng lượng ngữ cảnh âm thanh:

$$\Gamma_l(k, n) = \frac{\|E\{I_l(k, n)\}\|}{cE\{E(k, n)\}}$$

trong đó $I_l(k, n)$ là véctor mật độ được kết hợp với hướng thứ l . Nếu L DoA được truyền dẫn cùng với L tỉ lệ năng lượng, độ khuếch tán khi đó có thể được giảm từ L năng lượng như:

$$\Psi(k, n) = 1 - \sum_{l=1}^L \Gamma_l(k, n)$$

Các tham số không gian được truyền dẫn trong dòng bit có thể là L tỉ lệ năng lượng hoặc các tham số mới nhất này có thể cũng được chuyển đổi đến $L-1$ tỉ lệ năng lượng + tham số độ khuếch tán.

$$\Psi = 1 - \sum_{l=1}^L \Gamma_l$$

Sự mở rộng thứ hai chia không gian 2D hoặc 3D thành các phần không chồng lấp và truyền dẫn mỗi phần tập hợp các tham số DirAC (DoA+độ khuếch tán theo phần). Các tác giả sáng chế khi đó nói về DirAC cấp cao như được đưa ra trong tài liệu [5].

Cả hai sự mở rộng thực tế có thể được tổ hợp, và sáng chế liên quan đến cả hai sự mở rộng.

Các hình vẽ Fig.3a và Fig.3b minh họa các phương án của sáng chế, trong đó Fig.3a thể hiện phương pháp với sự tập trung trên nguyên lý cơ bản/phương pháp được sử dụng 100, trong đó thiết bị được sử dụng 50 được thể hiện bởi Fig.3b.

Fig.3a minh họa phương pháp 100 bao gồm các bước cơ bản 110, 120, và 130.

Các bước thứ nhất 110 và 120 là có thể so sánh với nhau, cụ thể đề cập đến việc nhận các tập hợp tham số âm thanh không gian. Trong bước thứ nhất 110, tập hợp thứ nhất được nhận, trong đó trong bước thứ hai 120, tập hợp thứ hai được nhận. Ngoài ra,

các bước nhận khác có thể có (không được thể hiện). Nên lưu ý rằng thiết lập thứ nhất có thể đề cập đến thời điểm thứ nhất/khung thứ nhất, tập hợp thứ hai có thể đề cập đến thời điểm thứ hai (tiếp theo)/khung thứ hai (tiếp theo), v.v.. Như được thảo luận trên đây, tập hợp thứ nhất cũng như tập hợp thứ hai có thể bao gồm thông tin độ khuếch tán (Ψ) và/hoặc thông tin hướng (góc phương vị hoặc góc nâng). Thông tin này có thể được mã hóa bằng cách sử dụng bộ mã hóa siêu dữ liệu không gian. Bây giờ giả định được thực hiện rằng tập hợp thứ hai gồm thông tin bị tổn hao hoặc hư hại trong khi truyền dẫn. Trong trường hợp, tập hợp thứ hai được thay thế bởi tập hợp thứ nhất. Điều này cho phép sự che giấu tổn hao gói cho các tham số âm thanh không gian như các tham số DirAC.

Trong trường hợp tổn hao gói, các tham số DirAC bị xóa của các khung sẽ được hoàn lại để hạn chế ảnh hưởng lên chất lượng. Điều này có thể đạt được bằng cách tạo ra một cách tổng hợp các tham số bị mất bằng cách xem xét các tham số đã nhận trong quá khứ. Ảnh không gian không ổn định có thể được nhận biết như không hài lòng và là thành phần lạ, mặc dù ảnh không gian bất biến nghiêm ngặt có thể được nhận biết là không tự nhiên.

Phương pháp 100 được thảo luận với Fig.3a có thể được thực hiện bởi thực thể 50 được thể hiện bởi Fig.3b. Thiết bị che giấu tổn hao 50 bao gồm giao diện 52 và bộ xử lý 54. Qua giao diện, tập hợp các tham số âm thanh không gian, Ψ_1 , azi_1 , ele_1 , Ψ_2 , azi_2 , Ψ_n , azi_n , ele có thể được nhận. Bộ xử lý 54 phân tích các tập hợp được nhận và, trong trường hợp tập hợp bị mất hoặc tổn hao, tập hợp này thay thế các tập hợp tổn hao hoặc hư hại, ví dụ, bằng tập hợp được nhận trước đó hoặc tập hợp có thể so sánh. Các chiến lược khác có thể được sử dụng, mà được thảo luận sau đây.

Chiến lược giữ: Nhìn chung an toàn để xem xét rằng ảnh không gian phải ổn định tương đối theo thời gian, mà có thể được tịnh tiến cho các tham số DirAC, tức là, hướng tới và độ khuếch tán mà chúng không thay đổi nhiều giữa các khung. Vì lý do này, phương pháp đơn giản, nhưng hiệu quả để giữ các tham số của khung được nhận tốt cuối cùng cho các khung bị mất trong khi truyền dẫn.

Sự ngoại suy của các hướng: Cách khác, có thể hình dung để ước lượng quỹ đạo của các sự kiện âm trong ngữ cảnh âm thanh và sau đó thử ngoại suy quỹ đạo được ước

lượng. Điều này liên quan cụ thể nếu sự kiện âm được cục bộ hóa tốt trong không gian và nguồn điểm, mà được phản ánh trong mô hình DirAC bởi độ khuếch tán thấp. Quỹ đạo được ước lượng có thể được tính toán từ các sự quan sát của các hướng trong quá khứ và khớp với đường con trong số các điểm này, mà có thể khai triển phép nội suy hoặc làm nhẵn. Sự phân tích hồi quy có thể cũng được sử dụng. Phép ngoại suy khi đó được thực hiện bằng cách đánh giá đường cong được khớp vượt qua phạm vi dữ liệu được quan sát.

Trong DirAC, các hướng thường được biểu hiện, được lượng tử hóa và được mã hóa trong các hệ tọa độ cực. Tuy nhiên, thường thuận tiện để xử lý các hướng và khi đó quỹ đạo trong các hệ tọa độ Đề-các để tránh mô đun xử lý các thuật toán 2π .

Hòa sắc của hướng: Khi sự kiện âm khuếch tán hơn, các hướng ít có nghĩa và có thể được xem xét như sự nhận biết của quy trình ngẫu nhiên. Sự hòa sắc có thể giúp trường âm được kết xuất tự nhiên hơn và hài hòa hơn bằng cách đưa vào nhiễu âm ngẫu nhiên cho các hướng trước đó trước khi sử dụng chúng cho các khung tồn hao. Nhiễu âm đưa vào và phương sai của nó có thể đóng vai trò như độ khuếch tán.

Sử dụng sự phân tích ngữ cảnh âm thanh DirAC tiêu chuẩn, chúng ta có thể học sự ảnh hưởng của độ khuếch tán về độ chính xác và nghĩa của hướng của mô hình. Sử dụng tín hiệu định dạng B nhân tạo mà tỉ lệ năng lượng trực tiếp đến khuếch tán (Direct-to-Diffuse energy Ratio - DDR) được cho giữa thành phần sóng phẳng và thành phần trường khuếch tán, chúng ta có thể phân tích các tham số DirAC thu được và độ chính xác của chúng.

Độ khuếch tán theo lý thuyết Ψ là hàm tỉ lệ năng lượng trực tiếp đến khuếch tán (Direct-to-Diffuse energy Ratio - DDR), Γ , và được biểu thị là:

$$\Psi = \frac{P_{diff}}{P_{diff} + P_{pw}} = \frac{1}{1 + \frac{P_{pw}}{P_{diff}}} = \frac{1}{1 + 10^{\Gamma/10}},$$

trong đó P_{pw} và P_{diff} là sóng phẳng và các năng lượng độ khuếch tán, tương ứng, và Γ là DDR được biểu thị theo thang dB.

Đương nhiên, có thể một hoặc tổ hợp trong số ba chiến lược được thảo luận có thể được sử dụng. Chiến lược được sử dụng được lựa chọn bởi bộ xử lý 54 phụ thuộc vào

các tập hợp tham số âm thanh không gian được nhận. Đối với điều này, các tham số âm thanh có thể, theo các phương án, được phân tích để cho phép ứng dụng các chiến lược khác nhau theo các đặc điểm của ngữ cảnh âm thanh và theo độ khuếch tán cụ thể hơn.

Điều này có nghĩa là, theo các phương án, bộ xử lý 54 được tạo cấu hình để cung cấp sự che giấu tổn hao gói cho âm thanh tham số không gian bằng cách sử dụng thông tin định hướng được nhận tốt trước đó và hòa sắc. Theo phương án khác, sự hòa sắc là hàm độ khuếch tán được ước lượng hoặc tỉ lệ năng lượng giữa các thành phần định hướng và không định hướng của ngữ cảnh âm thanh. Theo các phương án, sự hòa sắc là hàm âm điệu được đo của tín hiệu trộn giảm trộn giảm được truyền dẫn. Do đó, bộ phân tích thực hiện sự phân tích của mình dựa trên độ khuếch tán được ước lượng, tỉ lệ năng lượng và/hoặc âm điệu.

Trên Fig.3a và Fig.3b, độ khuếch tán được đo được cho trong hàm DDR bằng cách mô phỏng trường khuếch tán với $N=466$ nhiễu âm màu hồng không tương quan được đặt đều nhau trên hình cầu và sóng phẳng bởi nhiễu âm màu hồng độc lập được đặt tại góc phương vị 0 độ và góc nâng 0 độ. Được xác nhận rằng độ khuếch tán được đo trong sự phân tích DirAC, là ước lượng tốt của độ khuếch tán theo lý thuyết nếu chiều dài cửa sổ quan sát W đủ lớn. Điều này chỉ ra rằng độ khuếch tán có các đặc điểm dài hạn, mà xác nhận rằng tham số có thể trong trường hợp tổn hao gói được dự báo tốt bằng cách đơn giản giữ các giá trị đã được nhận tốt trước đó.

Mặt khác, sự ước lượng các tham số hướng có thể cũng được đánh giá về hàm độ khuếch tán đúng, mà đã được báo cáo trên Fig.4. Được thể hiện rằng góc nâng được ước lượng và góc phương vị của vị trí sóng phẳng lệch từ vị trí đúng nền tảng (góc phương vị 0 độ và góc nâng 0 độ) với độ lệch tiêu chuẩn tăng với độ khuếch tán. Đối với độ khuếch tán là 1, độ lệch tiêu chuẩn là khoảng 90 độ đối với góc phương vị được định nghĩa giữa 0 độ và 360 độ, tương ứng với góc hoàn toàn ngẫu nhiên cho sự phân bố đồng đều. Nói cách khác, góc phương vị khi đó là vô nghĩa. Sự quan sát tương tự có thể được thực hiện cho góc nâng. Nhìn chung, độ chính xác của hướng được ước lượng và độ ý nghĩa giảm với độ khuếch tán. Khi đó được hi vọng rằng hướng trong DirAC sẽ dao động theo thời gian và độ lệch từ giá trị được mong đợi với hàm phương sai của độ khuếch tán. Sự phân tán tự nhiên là một phần của mô hình DirAC, mà là bản chất cho

sự tái tạo đúng của ngữ cảnh âm thanh. Nhìn chung, kết xuất tại hướng bất biến thành phần định hướng của DirAC mặc dù độ khuếch tán cao, sẽ tạo ra nguồn điểm mà nên được nhận biết rộng hơn trong thực tế.

Đối với các lý do nêu trên, các tác giả sáng chế đề xuất áp dụng việc hòa sắc trên hướng trên đỉnh của chiến lược giữ. Biên độ của sự hòa sắc được tạo hàm của độ khuếch tán và có thể, ví dụ, theo sau các mô hình được vẽ trên Fig.4. Hai mô hình cho các góc được đo nâng và phương vị có thể được suy ra cho độ lệch tiêu chuẩn mà được biểu diễn là:

$$\sigma_{azi} = 65\Psi^{3.5} + \sigma_{ele}$$

$$\sigma_{ele} = 33.25\Psi + 1.25$$

Mã giả của sự che giấu tham số DirAC khi đó có thể:

```
for k in frame_start:frame_end
{
    if(bad_frame_indicator[k])
    {
        for band in band_start:band_end
        {
            diff_index = diffuseness_index[k-1][band];
            diffuseness[k][band] = unquantize_diffuseness(diff_index);

            azimuth_index[k][b] = azimuth_index[k-1][b];
            azimuth[k][b] = unquantize_azimuth(azimuth_index[k][b])
            azimuth[k][b] = azimuth[k][b] + random() * dithering_azi_scale[diff_index]

            elevation_index[k][b] = elevation_index[k-1][b];
            elevation[k][b] = unquantize_elevation(elevation_index[k][b])
            elevation[k][b] = elevation[k][b] + random() * dithering_ele_scale[diff_index]
        }
    }
    else
    {
        for band in band_start:band_end
        {
            diffuseness_index[k][b] = read_diffuseness_index()
            azimuth_index[k][b] = read_azimuth_index()
            elevation_index[k][b] = read_elevation_index()

            diffuseness[k][b] = unquantize_diffuseness(diffuseness_index[k][b])
            azimuth[k][b] = unquantize_azimuth(azimuth_index[k][b])
            elevation[k][b] = unquantize_elevation(elevation_index[k][b])
        }
    }
    output_frame[k] = Dirac_synthesis(diffuseness[k][b], azimuth[k][b], elevation[k][b])
}
```

trong đó bad_frame_indicator[k] là cờ hiệu biểu thị liệu khung tại chỉ số k đã được nhận tốt hay không. Trong trường hợp khung tốt, các tham số DirAC được đọc, được

giải mã và không được lượng tử hóa cho mỗi băng tham số tương ứng với phạm vi tần số đã cho. Trong trường hợp khung xấu, độ khuếch tán được giữ trực tiếp từ khung được nhận biết cuối cùng tại băng tham số giống nhau, trong khi góc phương vị và góc nâng được suy ra từ việc không lượng tử hóa các chỉ số được nhận tốt cuối cùng với việc đưa vào giá trị ngẫu nhiên được định tỉ lệ bởi hàm thừa số của chỉ số độ khuếch tán. Hàm `random()` xuất ra giá trị ngẫu nhiên theo sự phân bố đã cho. Quy trình ngẫu nhiên có thể cho phép ví dụ sự phân bố thông thường tiêu chuẩn với trung bình 0 và phương sai đơn vị. Cách khác, có thể theo sau sự phân bố đồng đều giữa -1 và 1 hoặc theo sau mật độ xác suất tam giác sử dụng, ví dụ, mã giả sau đây:

```
random()
{
    rand_val = uniform_random();
    if( rand_val <= 0.0f )
    {
        return 0.5f * sqrt(rand_val + 1.0f) - 0.5f;
    }
    else
    {
        return 0.5f - 0.5f * sqrt(1.0f - rand_val);
    }
}
```

Các tỉ lệ hòa sắc là các hàm của chỉ số độ khuếch tán được kế thừa từ khung được nhận tốt cuối cùng tại băng tham số tương tự và có thể được suy ra từ các mô hình được giảm tạo ra Fig.4. Ví dụ, trong trường hợp này, độ khuếch tán được mã hóa trên 8 chỉ số, chúng có thể tương ứng với các bảng sau:

```
dithering_azi_scale[8] = {
    6.716062e-01f,  1.011837e+00f,  1.799065e+00f,  2.824915e+00f,  4.800879e+00f,
    9.206031e+00f,  1.469832e+01f,  2.566224e+01f
};

dithering_ele_scale[8] = {
    6.716062e-01f,  1.011804e+00f,  1.796875e+00f,  2.804382e+00f,  4.623130e+00f,
    7.802667e+00f,  1.045446e+01f,  1.379538e+01f
};
```

Ngoài ra, độ mạnh hòa sắc có thể cũng được hướng phụ thuộc vào bản chất của tín hiệu trộn giảm. Thực tế, tín hiệu rất âm có xu hướng được nhận biết như nguồn được cục bộ hóa hơn như các tín hiệu không âm. Do đó, sự hòa sắc có thể được điều chỉnh khi đó trong hàm của âm điệu của sự trộn giảm được truyền dẫn, bằng cách giảm ảnh hưởng hòa sắc cho các mục âm. Âm điệu có thể được đo, ví dụ, trong miền thời gian

bằng cách tính toán độ khuếch đại dự báo dài hạn hoặc miền tần số bằng cách đo độ phẳng phổ.

Viện dẫn đến Fig.6a và Fig.6b, các phương án còn viện dẫn đến phương pháp giải mã ngữ cảnh âm thanh được mã hóa DirAC (xem, Fig.6a, phương pháp 200) và bộ giải mã 17 cho ngữ cảnh âm thanh được mã Hóa DirAC (xem Fig.6b) sẽ được thảo luận.

Fig.6a minh họa phương pháp 200 mới bao gồm các bước 110, 120 và 130 của phương pháp 100 và bước giải mã 210 bổ sung. Bước giải mã cho phép sự giải mã ngữ cảnh âm thanh được mã hóa DirAC bao gồm sự trộn giảm (không được thể hiện) bằng cách sử dụng tập hợp thứ nhất gồm các tham số âm thanh không gian và tập hợp thứ hai gồm các tham số âm thanh không gian, trong đó ở đây, tập hợp thứ hai được thay thế được sử dụng, được xuất ra bởi bước 130. Nguyên lý này được sử dụng bởi thiết bị 17, được thể hiện bởi Fig.6b. Fig.6b thể hiện bộ giải mã 70 bao gồm bộ xử lý che giấu tổn hao của các tham số âm thanh không gian 15 và bộ giải mã DirAC 72. Bộ giải mã DirAC 72 hoặc, cụ thể, bộ xử lý của bộ giải mã DirAC 72, nhận tín hiệu trộn giảm và tập hợp gồm các tham số âm thanh không gian, ví dụ, trực tiếp từ giao diện 52 và/hoặc được xử lý bởi bộ xử lý 52 theo phương pháp được thảo luận trên đây.

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

Các tín hiệu âm thanh được mã hóa theo sáng chế có thể được lưu trữ trên môi trường lưu trữ kỹ thuật số hoặc có thể được truyền trong môi trường truyền như môi trường truyền không dây hoặc môi trường truyền có dây như Internet.

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 ghi lưu trữ số, ví dụ, đĩa mềm, DVD, Blu-ray, CD, ROM, PROM, EPROM, EEPROM hoặc bộ nhớ FLASH, có các tín hiệu điều khiển có thể đọc bằng điện tử 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. Do đó, vật ghi lưu trữ số có thể đọc được bằng máy tính.

Một số phương án theo sáng chế bao gồm vật mang dữ liệu có các tín hiệu điều khiển có thể đọc được bằng điện tử, mà có khả năng kết hợp với hệ thống máy tính 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 các phương pháp được mô tả ở đây, được lưu trữ trên vật mang có thể đọc được bằng máy.

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. Vật mang dữ liệu, vật ghi lưu trữ số hoặc vật ghi đã được ghi thường là hữu hình và/hoặc không tạm thời.

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

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

Theo một số phương án, thiết bị logic lập trình được (ví dụ, mảng công 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 công lập trình được dạng trường có thể kết hợp với bộ vi xử lý để thực hiện một trong số các phương pháp được mô tả ở đây. Thông thường, các phương pháp tốt hơn là được thực hiện bởi thiết bị phần cứng bất kỳ.

Các phương án được mô tả ở trên chỉ mang tính minh họa cho các nguyên lý của sáng chế. Cần hiểu rằng các 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. Phương pháp (100) che giấu tổn hao các tham số âm thanh không gian, các tham số âm thanh không gian bao gồm ít nhất thông tin hướng tới, phương pháp bao gồm các bước sau:

nhận (110) tập hợp thứ nhất gồm các tham số không gian bao gồm ít nhất thông tin hướng tới thứ nhất ($azi1$, $ele1$);

nhận (120) tập hợp thứ hai gồm các tham số âm thanh không gian, bao gồm ít nhất thông tin hướng tới thứ hai ($azi2$, $ele2$); và

thay thế hướng thông tin tới thứ hai ($azi2$, $ele2$) của tập hợp thứ hai bằng thông tin hướng tới thay thế được suy ra từ thông tin hướng tới thứ nhất ($azi1$, $ele1$), nếu ít nhất thông tin hướng tới thứ hai ($azi2$, $ele2$) hoặc một phần thông tin hướng hướng tới thứ hai ($azi2$, $ele2$) bị mất hoặc hư hại.

trong đó bước thay thế bao gồm bước hòa sắc thông tin hướng tới thay thế; trong đó nhiều âm ngẫu nhiên sẽ được đưa vào phụ thuộc vào thông tin khuếch tán thứ nhất và/hoặc thứ hai ($\Psi1$, $\Psi2$); và/hoặc trong đó nhiều âm ngẫu nhiên sẽ được đưa vào được định tỉ lệ bởi thừa số phụ thuộc vào thông tin độ khuếch tán thứ nhất và/hoặc thứ hai ($\Psi1$, $\Psi2$); hoặc

trong đó bước thay thế bao gồm bước hòa sắc thông tin hướng tới thay thế; trong đó bước đưa vào được thực hiện, nếu thông tin độ khuếch tán thứ nhất hoặc thứ hai ($\Psi1$, $\Psi2$) biểu thị độ khuếch tán cao; và/hoặc nếu thông tin độ khuếch tán thứ nhất hoặc thứ hai ($\Psi1$, $\Psi2$) ở trên ngưỡng được định trước cho thông tin độ khuếch tán.

2. Phương pháp (100) theo điểm 1, trong đó các tập hợp thứ nhất và các tập hợp thứ hai gồm các tham số âm thanh không gian lần lượt bao gồm thông tin độ khuếch tán thứ nhất và thứ hai ($\Psi1$, $\Psi2$).

3. Phương pháp (100) theo điểm 2, trong đó thông tin độ khuếch tán thứ nhất hoặc thứ hai ($\Psi1$, $\Psi2$) được suy ra từ ít nhất một tỉ lệ năng lượng tương ứng với ít nhất một thông tin hướng tới.

4. Phương pháp (100) theo điểm 2 hoặc 3, trong đó phương pháp còn bao gồm thay thế thông tin độ khuếch tán thứ hai (Ψ_2) của tập hợp thứ hai bằng thông tin độ khuếch tán thay thế được suy ra từ thông tin độ khuếch tán thứ nhất (Ψ_1).

5. Phương pháp (100) theo một điểm trong số các điểm nêu trên, trong đó thông tin hướng tới thay thế tuân theo thông tin hướng tới thứ nhất (azi_1 , ele_1).

6. Phương pháp (100) theo một điểm trong số các điểm nêu trên, trong đó bước thay thế bao gồm đưa nhiều âm ngẫu nhiên vào thông tin hướng tới (azi_1 , ele_1) để thu được thông tin hướng tới thay thế.

7. Phương pháp (100) theo điểm 6, trong đó thông tin độ khuếch tán bao gồm hoặc dựa trên tỉ lệ giữa các thành phần định hướng và không định hướng của ngữ cảnh âm thanh được mô tả bởi tập hợp thứ nhất và/hoặc tập hợp thứ hai gồm các tham số âm thanh không gian.

8. Phương pháp (100) theo một điểm trong số các điểm 1 đến 7, phương pháp còn bao gồm bước phân tích âm điệu của ngữ cảnh âm thanh được mô tả bởi tập hợp thứ nhất và/hoặc thứ hai gồm các tham số âm thanh không gian hoặc phân tích âm điệu của sự trộn giảm được truyền dẫn thuộc về tập hợp thứ nhất và/hoặc thứ hai gồm các tham số âm thanh không gian để thu được giá trị âm điệu mô tả âm điệu; và

trong đó nhiều âm ngẫu nhiên sẽ được đưa vào phụ thuộc vào giá trị âm điệu.

9. Phương pháp (100) theo điểm 8, trong đó nhiều âm ngẫu nhiên được định tỉ lệ bởi thừa số giảm cùng với sự nghịch đảo của giá trị âm điệu hoặc âm điệu tăng.

10. Phương pháp (100) theo một điểm trong số các điểm nêu trên, trong đó phương pháp (100) bao gồm bước ngoại suy thông tin hướng tới thứ nhất (azi_1 , ele_1) để thu được thông tin hướng tới thay thế.

11. Phương pháp (100) theo điểm 10, trong đó sự ngoại suy dựa trên một hoặc nhiều thông tin hướng tới bổ sung thuộc về một hoặc nhiều tập hợp gồm các tham số âm thanh không gian.

12. Phương pháp (100) theo điểm 10 hoặc 11, trong đó phép ngoại suy được thực hiện, nếu thông tin độ khuếch tán thứ nhất hoặc thứ hai (Ψ_1 , Ψ_2) biểu thị độ khuếch tán

thấp; hoặc nếu thông tin độ khuếch tán thứ nhất và/hoặc thứ hai (Ψ_1 , Ψ_2) ở dưới ngưỡng được định trước cho thông tin độ khuếch tán.

13. Phương pháp (100) theo một điểm trong số các điểm nêu trên, trong đó tập hợp thứ nhất gồm các tham số âm thanh không gian thuộc về điểm thứ nhất theo thời gian và/hoặc khung thứ nhất và trong đó tập hợp thứ hai gồm các tham số âm thanh không gian thuộc về điểm thứ hai theo thời gian và/hoặc khung thứ hai; hoặc

trong đó tập hợp thứ nhất gồm các tham số âm thanh không gian thuộc về điểm thứ nhất theo thời gian và trong đó điểm thứ hai theo thời gian là tiếp theo điểm thứ nhất theo thời gian hoặc trong đó khung thứ hai tiếp theo khung thứ nhất.

14. Phương pháp (100) theo một điểm trong số các điểm nêu trên, trong đó tập hợp thứ nhất gồm các tham số âm thanh không gian trước đó, trong đó tập con thứ nhất gồm các tham số âm thanh không gian cho băng tần thứ nhất và tập con thứ hai gồm các tham số âm thanh không gian cho băng tần thứ hai; và/hoặc

trong đó tập hợp thứ hai gồm các tham số âm thanh không gian trước đó, trong đó tập con thứ nhất khác gồm các tham số âm thanh không gian cho băng tần thứ nhất và tập con thứ hai khác gồm các tham số âm thanh không gian cho băng tần thứ hai.

15. Phương pháp (200) giải mã ngữ cảnh âm thanh được mã hóa DirAC, phương pháp bao gồm các bước sau:

giải mã ngữ cảnh âm thanh được mã hóa DirAC bao gồm sự trộn giảm, tập hợp thứ nhất gồm các tham số âm thanh không gian và tập hợp thứ hai gồm các tham số âm thanh không gian;

thực hiện phương pháp theo một điểm trong số các bước trước đó.

16. Vật ghi lưu trữ số có thể đọc được bằng máy được lưu trữ trên đó chương trình máy tính có mã chương trình để thực hiện, khi chạy trên máy tính phương pháp (100) hoặc phương pháp (200) theo một trong số các điểm nêu trên.

17. Thiết bị che giấu tổn hao (50) để che giấu tổn hao các tham số âm thanh không gian, các tham số âm thanh không gian bao gồm ít nhất thông tin hướng tới, thiết bị bao gồm:

bộ nhận (52) để nhận (110) tập hợp thứ nhất gồm các tham số âm thanh không gian bao gồm thông tin hướng tới thứ nhất ($azi1, ele1$) và để nhận (120) tập hợp thứ hai gồm các tham số âm thanh không gian bao gồm thông tin hướng tới thứ hai ($azi2, ele2$);

bộ xử lý (54) để thay thế hướng thông tin tới thứ hai ($azi2, ele2$) của tập hợp thứ hai bằng thông tin hướng tới thay thế được suy ra từ thông tin hướng tới thứ nhất ($azi1, ele1$), nếu ít nhất thông tin hướng tới thứ hai ($azi2, ele2$) hoặc một phần thông tin hướng tới thứ hai ($azi2, ele2$) bị mất hoặc hư hại;

trong đó thay thế bao gồm bước hòa sắc thông tin hướng tới thay thế; trong đó nhiều âm ngẫu nhiên sẽ được đưa vào phụ thuộc vào thông tin khuếch tán thứ nhất và/hoặc thứ hai ($\Psi1, \Psi2$); và/hoặc trong đó nhiều âm ngẫu nhiên sẽ được đưa vào được định tỉ lệ bởi thừa số phụ thuộc vào thông tin độ khuếch tán thứ nhất và/hoặc thứ hai ($\Psi1, \Psi2$); hoặc

trong đó thay thế bao gồm bước hòa sắc thông tin hướng tới thay thế; trong đó bước đưa vào được thực hiện, nếu thông tin độ khuếch tán thứ nhất hoặc thứ hai ($\Psi1, \Psi2$) biểu thị độ khuếch tán cao; và/hoặc nếu thông tin độ khuếch tán thứ nhất hoặc thứ hai ($\Psi1, \Psi2$) ở trên ngưỡng được định trước cho thông tin độ khuếch tán.

18. Bộ giải mã (70) cho ngữ cảnh âm thanh được mã hóa DirAC bao gồm thiết bị che giấu tổn hao theo điểm 17.

10

1/10

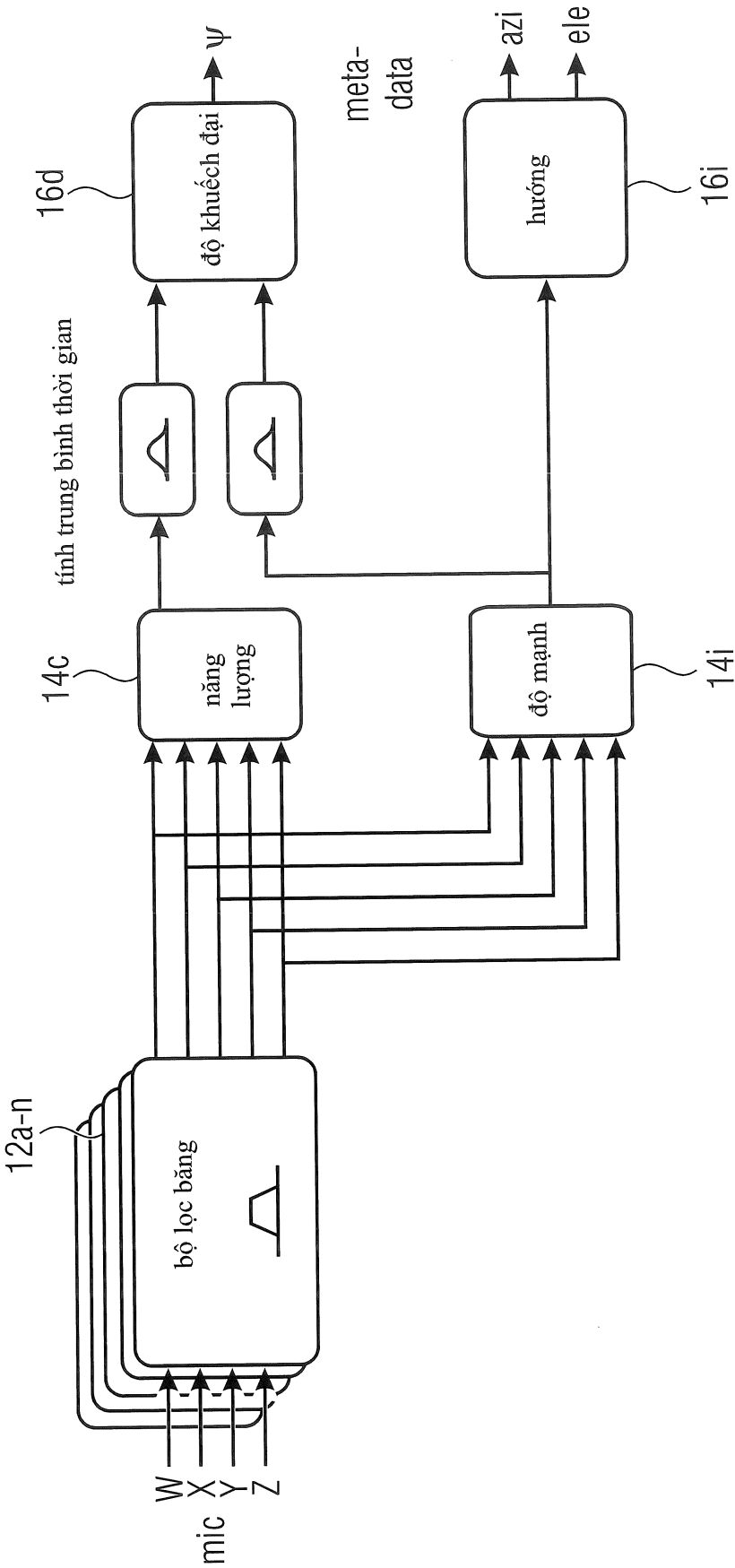


Fig. 1a

2/10

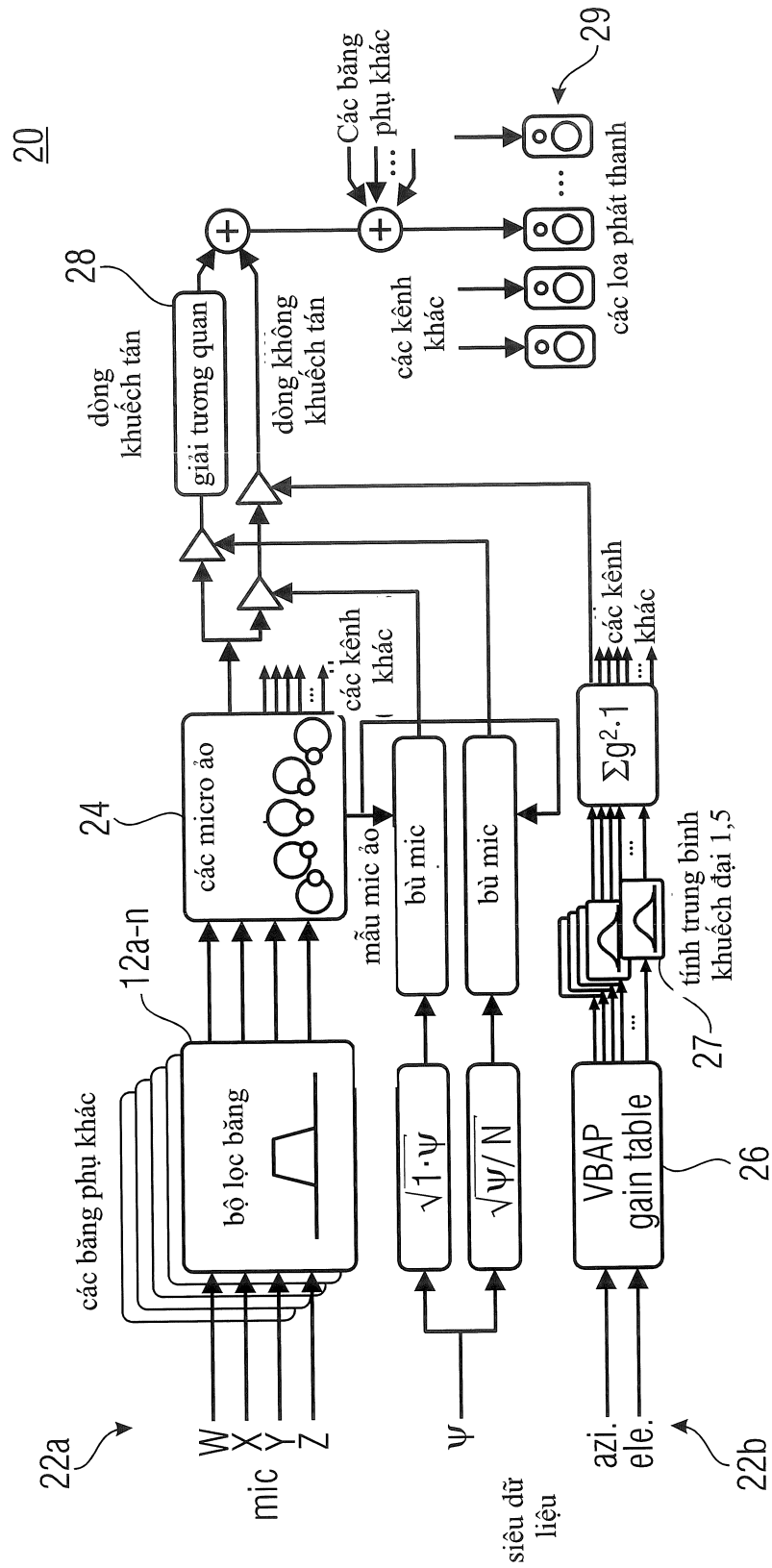


Fig. 1b

3/10

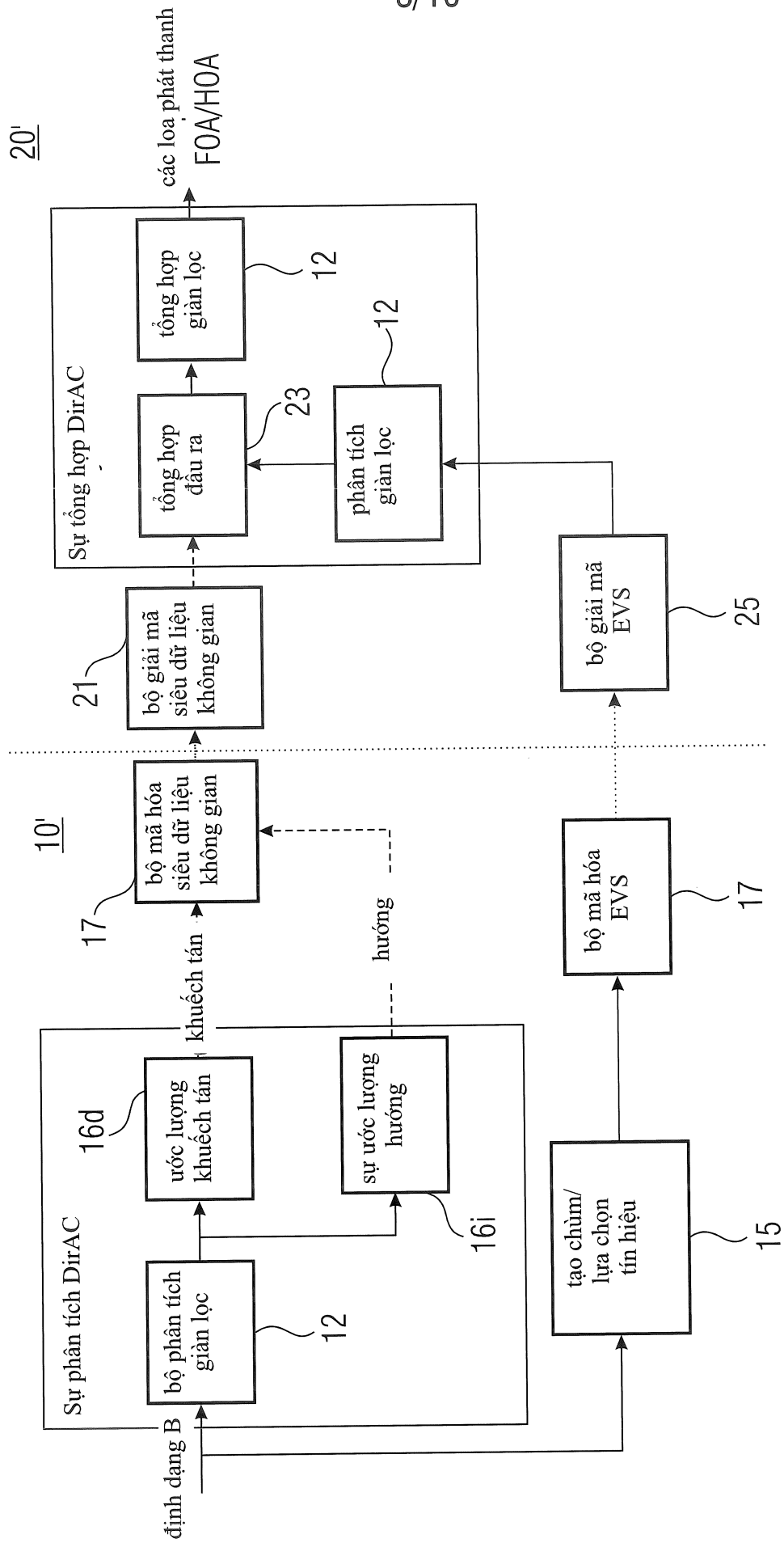


Fig. 2

4/10

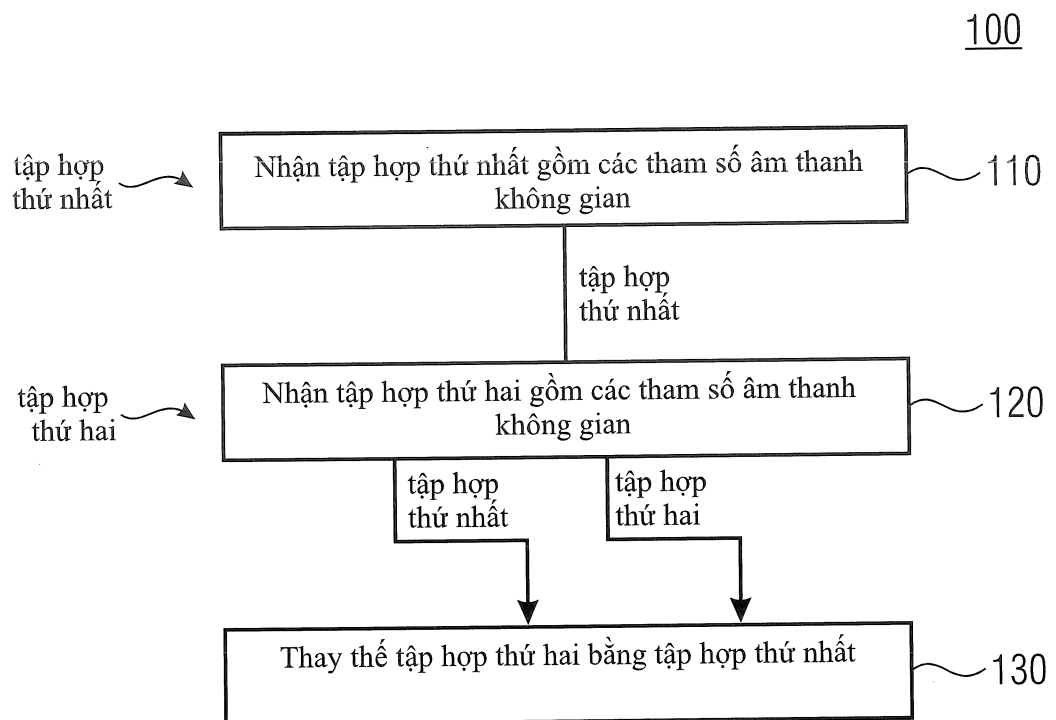


Fig. 3a

5/10

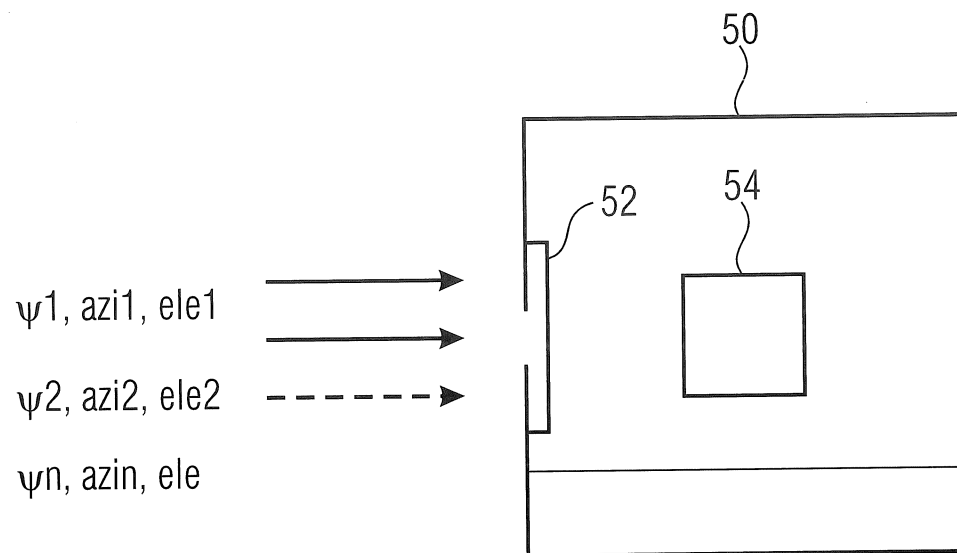


Fig. 3b

6/10

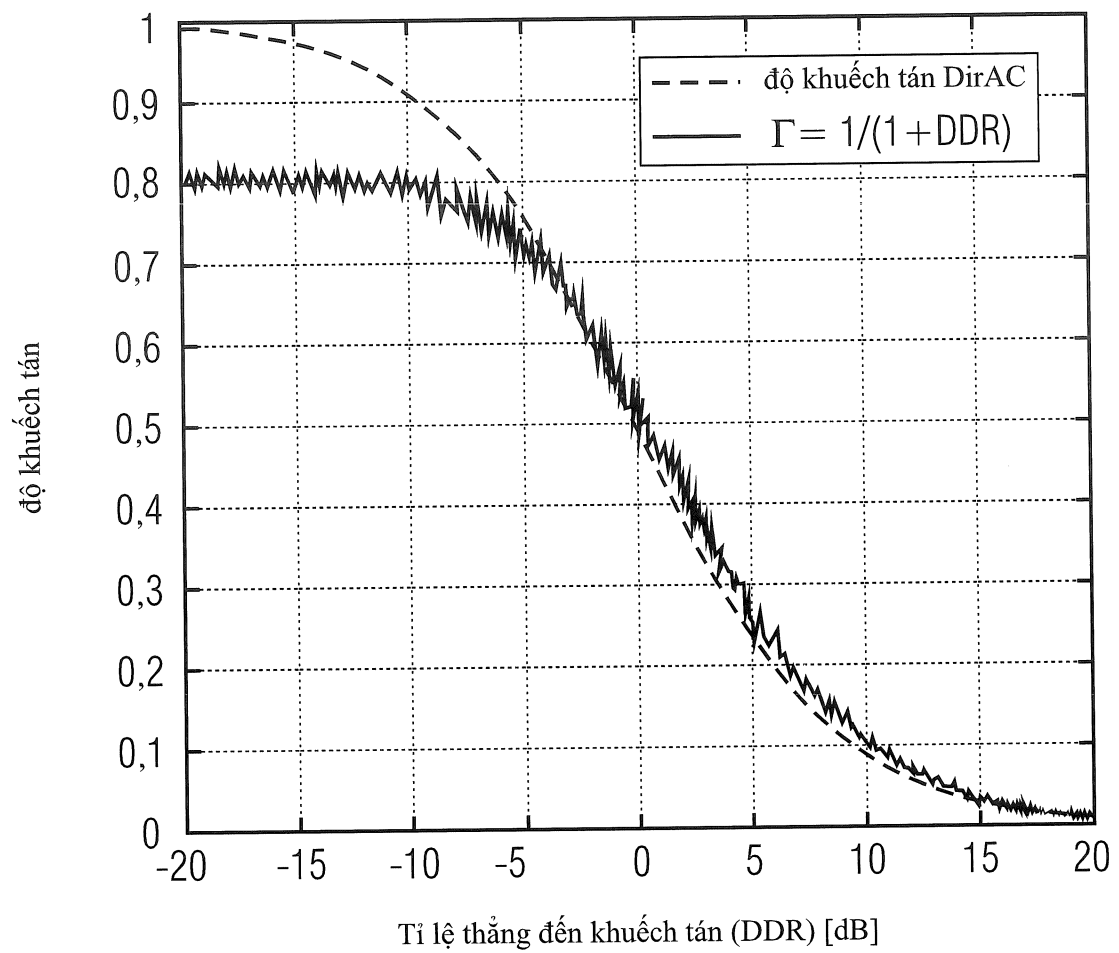


Fig. 4a

7/10

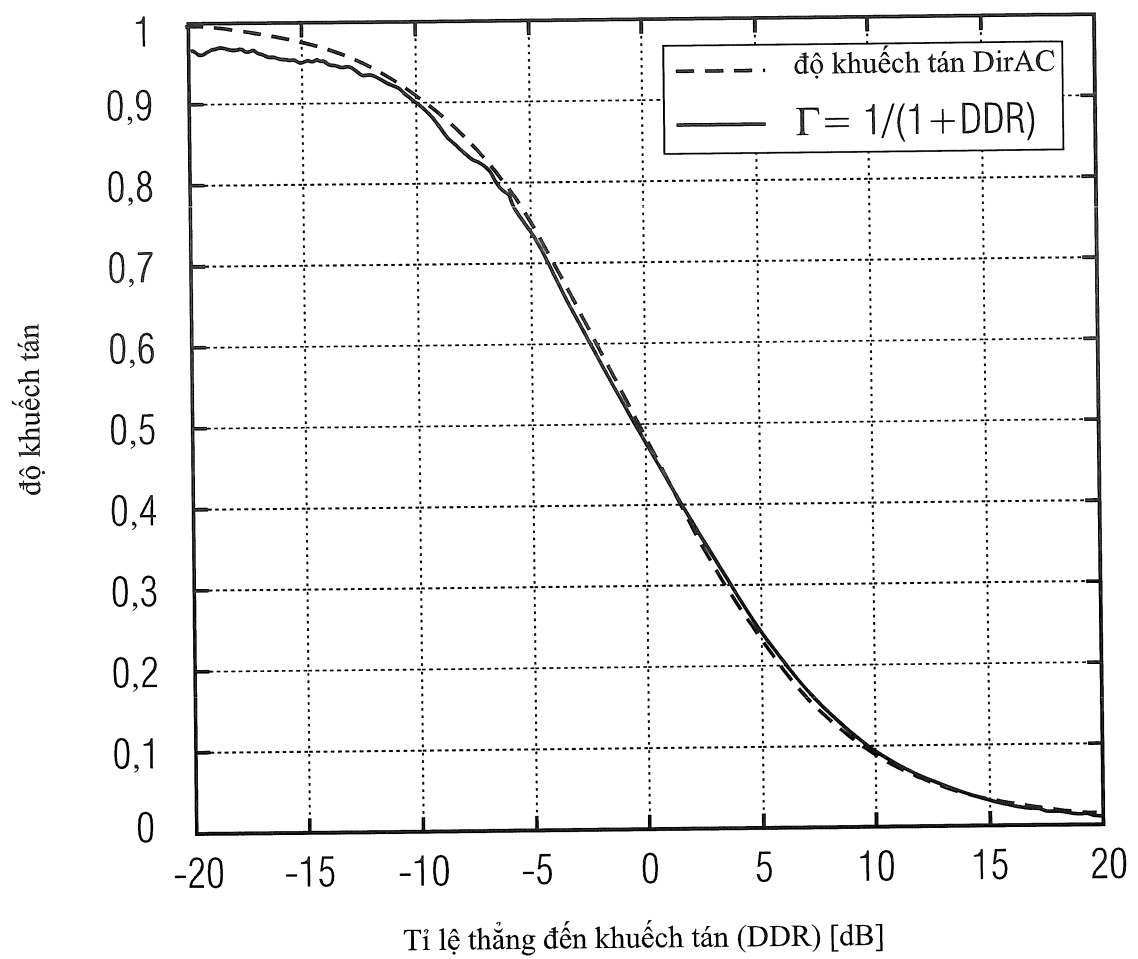


Fig. 4b

8/10

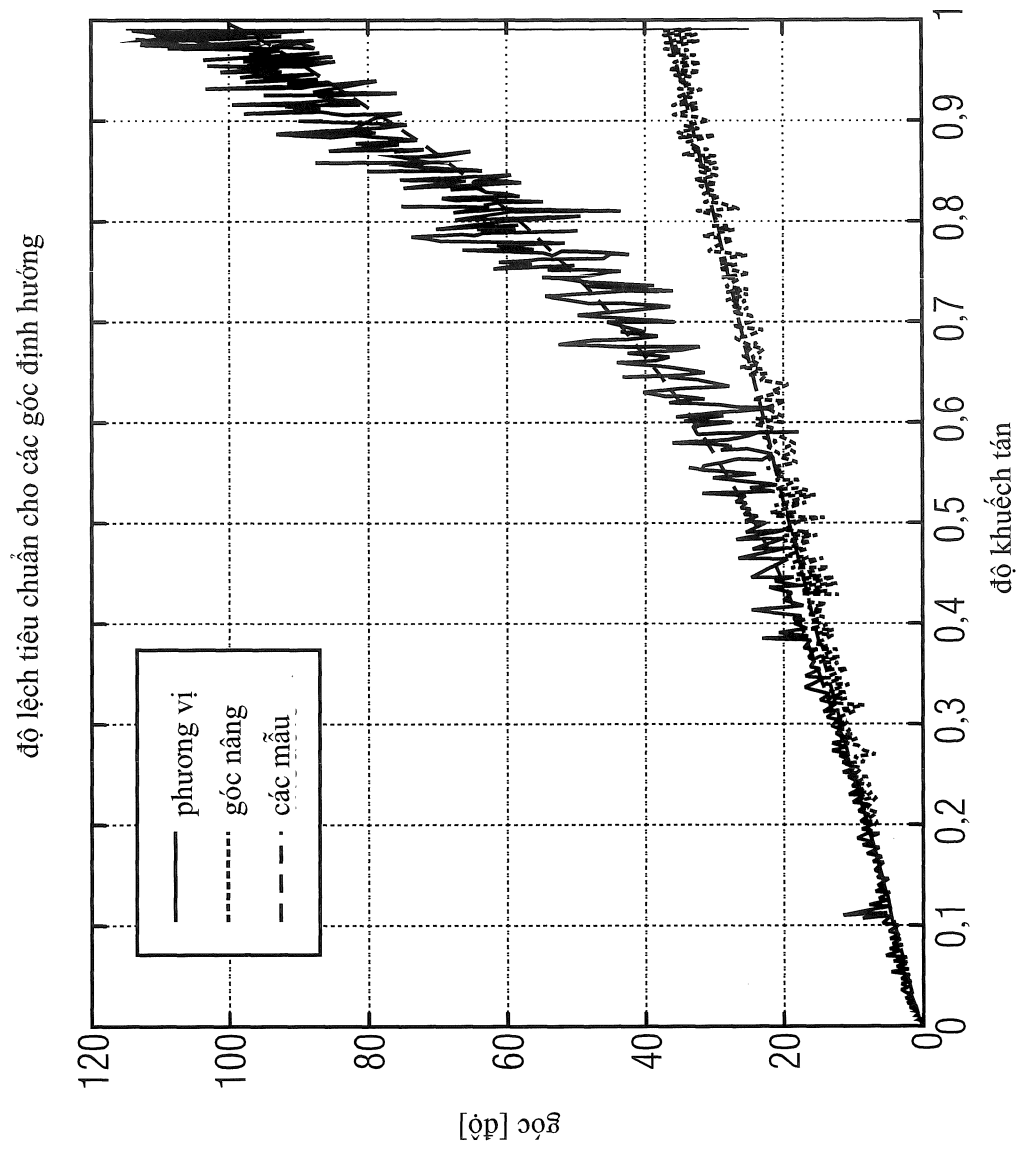


Fig. 5

9/10

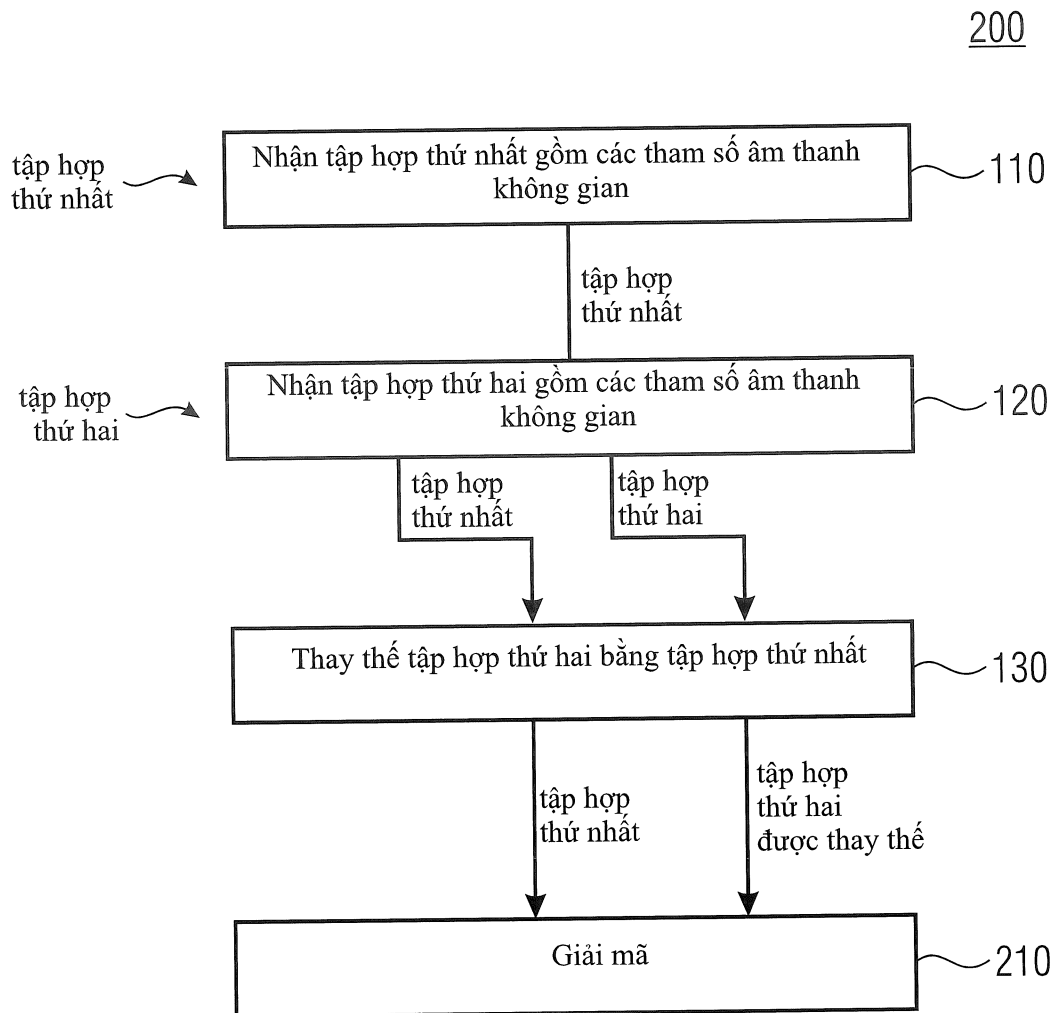


Fig. 6a

10/10

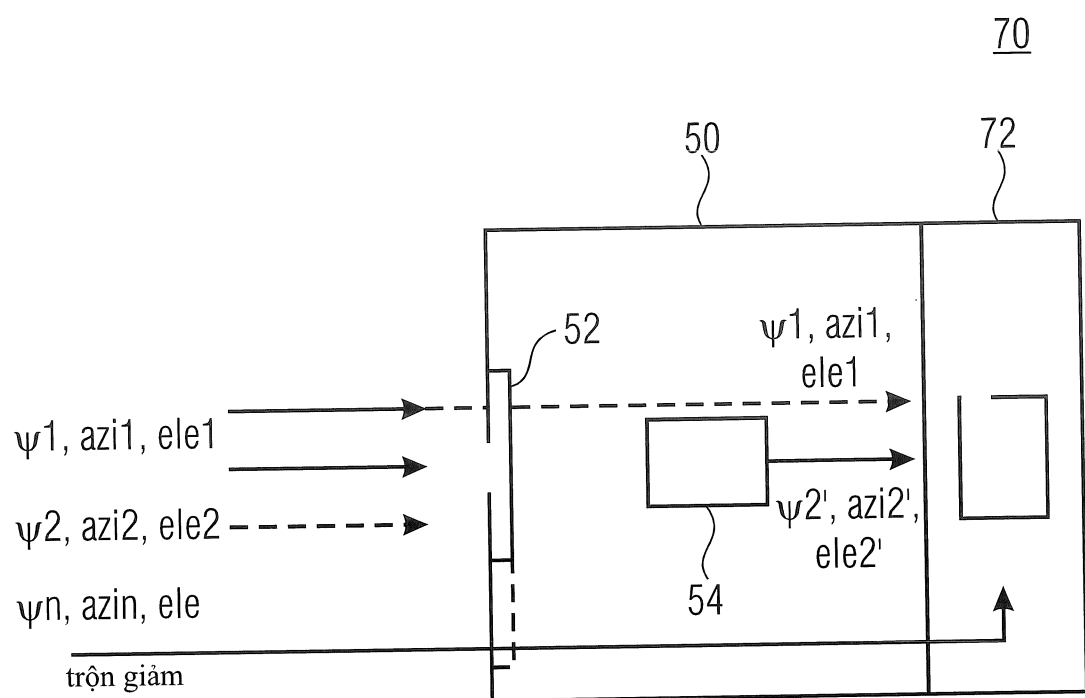


Fig. 6b