



(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)



1-0029418

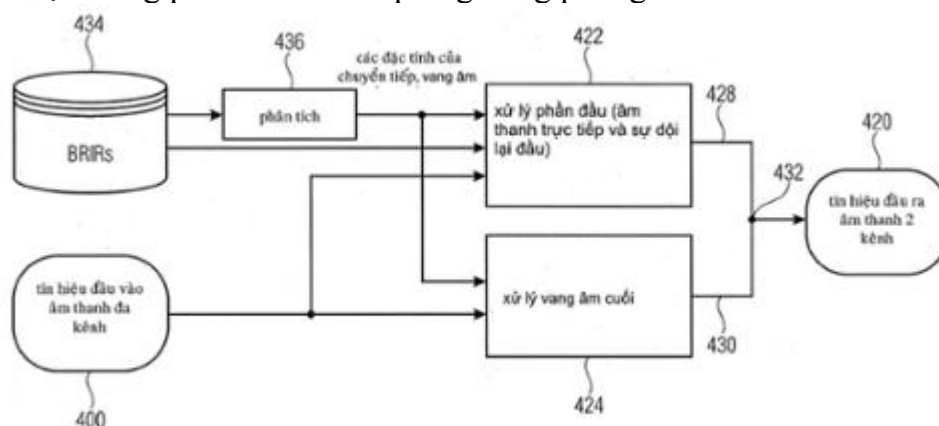
(51)<sup>7</sup> G10K 15/12; G01H 7/00

(13) B

- (21) 1-2016-00584 (22) 16/07/2014  
(86) PCT/EP2014/065227 16/07/2014 (87) WO 2015/010983 A1 29/01/2015  
(30) 13177362.4 22/07/2013 EP; 13189230.9 18/10/2013 EP  
(45) 25/09/2021 402 (43) 25/04/2016 337A  
(73) FRAUNHOFER-GESELLSCHAFT ZUR FOERDERUNG DER ANGEWANDTEN  
FORSCHUNG E.V. (DE)  
Hansastraße 27c, 80686 Muenchen, Germany  
(72) FUEG, Simone (DE); PLOGSTIES, Jan (DE).  
(74) CÔNG TY LUẬT TRÁCH NHIỆM HỮU HẠN AMBYS HÀ NỘI (AMBYS  
HANOI)

(54) PHƯƠNG PHÁP XỬ LÝ TÍN HIỆU ÂM THANH THEO ĐÁP ỨNG XUNG PHÒNG, BỘ PHẬN XỬ LÝ TÍN HIỆU, BỘ MÃ HÓA ÂM THANH, BỘ GIẢI MÃ ÂM THANH, BỘ KẾT XUẤT LẬP THỂ

(57) Sáng chế đề cập đến phương pháp xử lý tín hiệu âm thanh (400) theo đáp ứng xung phòng (434). Tín hiệu âm thanh (400) được xử lý riêng biệt (422, 424) với phần đầu và âm vang cuối của đáp ứng xung phòng (434), và phần đầu được xử lý (428) của tín hiệu âm thanh và tín hiệu vang lại (430) được tổ hợp (432). Sự chuyển tiếp từ phần đầu sang âm vang cuối trong đáp ứng xung phòng đạt được khi số đo tương quan đạt đến ngưỡng, ngưỡng này được thiết lập phụ thuộc vào số đo tương quan cho một trong số các phản xạ đầu được lựa chọn trong phần đầu của đáp ứng xung phòng.



### **Lĩnh vực kỹ thuật được đề cập**

Sáng chế đề cập đến lĩnh vực mã hóa/giải mã âm thanh, cụ thể là sáng chế đề cập đến lập mã âm thanh trong không gian và lập mã đối tượng âm thanh trong không gian, ví dụ, lĩnh vực hệ thống mã hóa-giải mã âm thanh ba chiều. Các phương án của sáng chế đề cập đến phương pháp xử lý tín hiệu âm thanh theo đáp ứng xung phòng và xác định sự chuyển tiếp từ các phản xạ đầu sang âm vang cuối trong đáp ứng xung phòng này.

### **Tình trạng kỹ thuật của sáng chế**

Công cụ lập mã âm thanh trong không gian đã được biết đến trong lĩnh vực kỹ thuật tương ứng và được chuẩn hóa, ví dụ, trong vùng bao quanh MPEG. Lập mã âm thanh trong không gian bắt đầu từ nhiều đầu vào ban đầu, ví dụ 5 hoặc 7 kênh đầu vào, mà được nhận dạng bởi vị trí của chúng trong thiết lập tái tạo, ví dụ như kênh bên trái, kênh trung tâm, kênh bên phải, kênh bao quanh trái, kênh bao quanh phải và kênh tăng cường tần số thấp. Bộ mã hóa âm thanh trong không gian có thể suy ra một hoặc nhiều kênh trộn giảm từ kênh ban đầu và, ngoài ra, có thể suy ra dữ liệu tham số liên quan đến các tín hiệu trong không gian như độ chênh lệch mức giữa các kênh trong giá trị tổ hợp kênh, độ chênh lệch pha giữa các kênh, độ chênh lệch thời gian giữa các kênh, v.v. Một hoặc nhiều kênh trộn giảm được truyền cùng với thông tin phụ theo tham số biểu thị các tín hiệu trong không gian đến bộ giải mã âm thanh trong không gian để giải mã các kênh trộn giảm và dữ liệu tham số tổ hợp để cuối cùng thu được các kênh đầu ra là phiên bản gần đúng các kênh đầu vào ban đầu. Vị trí các kênh trong thiết lập đầu ra có thể được cố định, ví dụ định dạng 5.1, định dạng 7.1, v.v.

Tương tự, công cụ lập mã đối tượng âm thanh trong không gian đã được biết đến trong lĩnh vực kỹ thuật tương ứng và được chuẩn hóa, ví dụ, theo tiêu chuẩn MPEG SAOC (SAOC = spatial audio object coding - lập mã đối tượng âm thanh trong không gian). Trái với việc lập mã âm thanh trong không gian bắt đầu từ các kênh ban

đầu, lập mã đối tượng âm thanh trong không gian bắt đầu từ các đối tượng âm thanh mà không tự động dành riêng cho thiết lập tái tạo kết xuất. Đúng hơn là, vị trí của các đối tượng âm thanh trong màn hình tái tạo là linh hoạt và có thể được thiết lập bởi người dùng, ví dụ bằng cách nhập thông tin kết xuất xác định vào bộ giải mã lập mã đối tượng âm thanh trong không gian. Ngoài ra hoặc theo cách khác, thông tin kết xuất có thể được truyền như thông tin phụ hoặc siêu dữ liệu bổ sung; thông tin kết xuất có thể bao gồm thông tin mà tại đó vị trí trong thiết lập tái tạo đối tượng âm thanh xác định được bố trí (ví dụ, qua thời điểm nhất định). Để thu được sự nén dữ liệu nhất định, số lượng đối tượng âm thanh được mã hóa nhờ sử dụng bộ mã hóa SAOC để tính toán, từ các đối tượng đầu vào, một hoặc nhiều kênh vận chuyển bằng cách trộn giảm các đối tượng theo thông tin trộn giảm xác định. Hơn nữa, bộ mã hóa SAOC tính toán thông tin phụ theo tham số thể hiện tín hiệu giữa các đối tượng như độ lệch mức đối tượng (object level differences - OLD), giá trị tổ hợp đối tượng, v.v. Như trong SAC (SAC = Mã hóa âm thanh trong không gian), dữ liệu tham số giữa các đối tượng được tính toán trong dải thời gian/tần số riêng rẽ. Đối với khung xác định (ví dụ, các mẫu 1024 hoặc 2048) của tín hiệu âm thanh, nhiều băng tần (ví dụ các băng 24, 32, hoặc 64) được xem xét sao cho dữ liệu tham số được cung cấp cho mỗi khung và mỗi băng tần. Ví dụ, khi phần âm thanh có 20 khung và khi mỗi khung được chia nhỏ thành 32 băng tần, thì số lượng dải thời gian/tần số bằng 640.

Trong hệ thống âm thanh ba chiều, có thể mong muốn để tạo ra ấn tượng về không gian của tín hiệu âm thanh như nếu tín hiệu âm thanh được nghe trong phòng cụ thể. Trong trường hợp như vậy, đáp ứng xung phòng của một phòng cụ thể được đề xuất, ví dụ dựa trên phép đo, và được sử dụng để xử lý tín hiệu âm thanh khi thể hiện nó cho người nghe. Các tác giả sáng chế mong muốn xử lý âm thanh trực tiếp và các phản xạ đầu theo cách biểu diễn tách khỏi âm vang cuối. Điều này yêu cầu phải xác định vị trí mà các phản xạ đầu kết thúc và vị trí âm vang cuối bắt đầu.

#### *Tài liệu tham khảo*

- [1] T. Hidaka et al: "A new definition of boundary point between early reflections and latereverberation in room impulse responses". Forum Acusticum, 2005.

[2] Jot et al: “Analysis and synthesis of room reverberation based on a statistical time frequency model”.

[3] J.S. Abel, P. Huang: “A Simple, Robust Measure of Reverberation Echo Density”. AESConvention, San Francisco, 2006.

[4] R. Stewart, M. Sandler: “Statistical Measures of Early Reflections of Room ImpulseResponses”. DAFx, 2007.

[5] Reilly et al: “Using Auralisation for Creating Animated 3-D Sound Fields Across Multiple Speakers”. AES Convention, New York, 1995.

[6] Usher, J.: “An improved method to determine the onset timings of reflections in anacoustic impulse response”. Journal of the Acoustical Society of America, (2010, volume127) band 4, p. 172-177.

[7] Masri, P.: “Computer Modelling of Sound for Transformation and Synthesis ofMusical Signals“. PhD thesis, University of Bristol, 1996.

### **Bản chất kỹ thuật của sáng chế**

Mục đích của sáng chế là đề xuất phương pháp cải tiến để xử lý tín hiệu âm thanh theo đáp ứng xung phòng.

Mục đích này đạt được bằng phương pháp theo điểm 1 yêu cầu bảo hộ, bộ phận xử lý tín hiệu theo điểm 16 yêu cầu bảo hộ, bộ mã hóa âm thanh theo điểm 18 yêu cầu bảo hộ, bộ giải mã âm thanh theo điểm 20 yêu cầu bảo hộ và bộ kết xuất lập thể theo điểm 24 yêu cầu bảo hộ.

Sáng chế được dựa trên các phát hiện của tác giả sáng chế đó là theo phương pháp thông thường thì vấn đề tồn tại ở chỗ có các trường hợp trong đó việc xác định sự chuyển tiếp từ các phản xạ đầu sang âm vang cuối là quá sớm bởi vì tương quan được sử dụng để đánh giá xảy ra sự chuyển tiếp đã đạt đến ngưỡng trước khi phản xạ thứ nhất xảy ra hoặc va chạm. Dựa trên các phát hiện này và kể từ khi được biết rằng thời gian chuyển tiếp phải lớn hơn thời gian tới của phản xạ thứ nhất, bởi vì phản xạ thứ nhất là rõ rệt và chắc chắn không thể là vang âm khuếch tán cuối cùng, các tác giả sáng chế phát hiện rằng cần phải tránh sử dụng ngưỡng cố định, tốt hơn là, theo phương pháp của sáng chế thì ngưỡng được định rõ sao cho nó phụ thuộc vào tương

quan tại thời điểm va chạm của một trong số các phản xạ đầu. Điều này đảm bảo rằng phản xạ đầu luôn luôn được định vị trước thời gian chuyển tiếp.

(1) Sáng chế đề xuất phương pháp xử lý tín hiệu âm thanh theo đáp ứng xung phòng, phương pháp bao gồm các bước:

xử lý riêng biệt tín hiệu âm thanh với phần đầu và âm vang cuối của đáp ứng xung phòng; và

tổ hợp tín hiệu âm thanh được xử lý với phần đầu của đáp ứng xung phòng và tín hiệu vang lại,

trong đó, sự chuyển tiếp từ phần đầu sang âm vang cuối trong đáp ứng xung phòng được xác định bởi số đo tương quan đạt đến ngưỡng, ngưỡng này được đặt phụ thuộc vào số đo tương quan cho một trong số các phản xạ đầu được chọn trong phần đầu của đáp ứng xung phòng.

Phương pháp theo sáng chế là có lợi bởi vì nó cho phép xử lý tín hiệu âm thanh được cải thiện dựa trên điểm chuyển tiếp mạnh. Phương pháp theo sáng chế được thực hiện ngoài phòng, dù phương pháp lập thể được sử dụng hay không và độc lập với góc tới. Khi so với phương pháp của kỹ thuật hiện nay, thì phương pháp theo sáng chế còn có lợi hơn bởi vì nó không phụ thuộc nhiều vào góc phương vị của đáp ứng xung lập thể và mối tương quan giữa biên độ của âm thanh trực tiếp và phản xạ va chạm thứ nhất.

(2) Theo các phương án, số đo tương quan mô tả đối với đáp ứng xung phòng tính tương tự của sự phân rã năng lượng âm thanh bao gồm trạng thái ban đầu và sự phân rã năng lượng âm thanh bắt đầu tại thời điểm bất kỳ theo sau trạng thái ban đầu trong khoảng tần số định trước.

(3) Theo các phương án, bước xác định sự chuyển tiếp bao gồm xác định sự phân bố năng lượng âm thanh dựa trên đáp ứng xung phòng, và xác định nhiều số đo tương quan biểu thị, đối với nhiều phần phân bố được xác định, sự tương quan giữa năng lượng âm thanh trong phần phân bố được xác định tương ứng và năng lượng âm thanh ở trạng thái ban đầu.

(4) Theo các phương án, bước xác định sự phân bố bao gồm việc xác định sự phân bố thời gian-tần số của năng lượng âm thanh, trong đó phần phân bố bao gồm khối thời gian có độ dài định trước, trạng thái ban đầu được xác định bởi khối thời gian thứ nhất trong số nhiều khối thời gian của sự phân bố thời gian-tần số.

Điều này là có lợi bởi nó cho phép phân tích sự phân bố tần số tại các thời điểm khác nhau, nhờ đó mô tả các đặc tính của sự phân bố tần số qua thời điểm nhất định.

(5) Theo các phương án, bước xác định sự phân bố bao gồm tính toán sự giảm phân rã năng lượng EDR (*EDR – energy decay relief*) từ đáp ứng xung phòng, trong đó EDR được tính toán như sau:

$$E(t, \omega) = \left| \int_t^{\infty} h(\tau) e^{-j\omega\tau} d\tau \right|^2$$

trong đó,

$E(t, \omega)$  = sự giảm phân rã năng lượng,

$h(\tau)$  = đáp ứng xung phòng,

$\omega$  =  $2\pi f$ .

(6) Theo các phương án đáp ứng xung phòng có độ dài thực định trước, và trong đó bước xác định sự phân bố thời gian-tần số bao gồm tính toán quang phổ FFT của đáp ứng xung phòng sử dụng cửa sổ có độ dài tương ứng với độ dài thực của đáp ứng xung phòng.

Điều này có lợi khi FFT/DFT được định rõ và có các thuật toán hiệu quả để tính toán các giá trị quang phổ. Nếu các giá trị này trong cửa sổ là đã biết thì FFT/DFT có thể được tính toán theo cách đơn giản.

(7) Theo các phương án, năng lượng âm thanh ở trạng thái ban đầu được xác định bằng cách lấy toàn bộ độ dài thực của đáp ứng xung phòng, tính toán quang phổ FFT và lấy bình phương các giá trị tuyệt đối, và năng lượng âm thanh của khối thời gian được xác định bằng cách di chuyển cửa sổ theo thời gian kết hợp với khối thời gian, đệm 0 các mẫu được tạo cửa sổ đến độ dài thực, tính toán FFT và lấy bình phương các giá trị tuyệt đối.

Điều này có lợi, khi không cần giàn lọc bổ sung hoặc loại tương tự để tính toán bằng hợp của  $EDR$ ; thì chỉ cần di chuyển cửa sổ.

(8) Theo các phương án, số đo tương quan là số đo tương quan mô tả tính tương tự của sự phân rã năng lượng âm thanh bao gồm trạng thái ban đầu và sự phân rã năng lượng âm thanh bắt đầu tại thời điểm bất kỳ sau trạng thái ban đầu. Số đo tương quan có thể được tính toán như sau:

$$\rho(t) = \frac{\sum_{\omega} (E(1, \omega) - \bar{E}(1, \omega)) \cdot \sum_{\omega} (E(t, \omega) - \bar{E}(t, \omega))}{\sqrt{\sum_{\omega} (E(1, \omega) - \bar{E}(1, \omega))^2} \cdot \sqrt{\sum_{\omega} (E(t, \omega) - \bar{E}(t, \omega))^2}}$$

trong đó,

$\rho(t)$  = số đo tương quan,

$E(1, \omega)$  = sự giảm phân rã năng lượng trong toàn bộ khoảng tần số tại tần số  $f$ ,

$\bar{E}(1, \omega)$  = giá trị trung bình trên tất cả các tần số của sự giảm phân rã năng lượng trong toàn bộ khoảng tần số,

$E(t, \omega)$  = sự giảm phân rã năng lượng tại tần số  $f$  bắt đầu tại thời gian  $t$ ,

$\bar{E}(t, \omega)$  = giá trị trung bình trên tất cả các tần số của sự giảm phân rã năng lượng trong toàn bộ khoảng bắt đầu tại thời gian  $t$ ,

$\omega$  =  $2\pi f$ .

Điều này có lợi khi công thức nêu trên đề cập đến hệ số tương quan Pearson đã biết (tương quan momen tích Pearson). Hệ số tương quan có thể được tính toán trực tiếp từ  $EDR$ .

(9) Theo các phương án, ngưỡng được xác định dựa trên giá trị không đổi và số đo tương quan đối với một trong số các phản xạ đầu được chọn. Ngưỡng này có thể được định rõ như sau:

$$\rho(t) = c \cdot \rho(t_F)$$

trong đó,

$\rho(t_F)$  = số đo tương quan đối với một trong số các phản xạ đầu được chọn,

$t_F$  = chỉ số thời gian trong đó một trong số các phản xạ đầu được chọn sau âm thanh trực tiếp đi tới,

$c$  = giá trị không đổi dựa trên  $\frac{1}{e}$ ,  $e$  là số Euler.

Điều này có lợi khi ngưỡng không phải là hằng số, mà phụ thuộc vào phản xạ đầu được chọn để đảm bảo rằng tương quan không nằm trong khoảng dưới ngưỡng.

(10) Theo các phương án, thời gian của một trong số các phản xạ đầu được chọn được xác định, ví dụ bằng cách thực hiện toán tử độ nhọn, bằng sự phát hiện ngưỡng hoặc phát hiện va chạm.

Điều này có lợi khi thời gian va chạm của phản xạ có thể được tính toán trực tiếp và tự động từ các mẫu miền thời gian của đáp ứng xung.

(11) Theo các phương án, một trong số các phản xạ đầu được lựa chọn là phản xạ thứ nhất theo sau âm thanh trực tiếp.

(12) Sáng chế đề xuất bộ phận xử lý tín hiệu, bao gồm đầu vào để nhận tín hiệu âm thanh, bộ xử lý được cấu hình hoặc được lập trình để xử lý tín hiệu âm thanh được nhận theo đáp ứng xung phòng theo phương pháp của sáng chế, và đầu ra để tổ hợp phần đầu được xử lý của tín hiệu âm thanh được nhận và tín hiệu vang lại vào tín hiệu âm thanh đầu ra. Bộ phận xử lý tín hiệu có thể bao gồm bộ xử lý phần đầu để xử lý tín hiệu âm thanh được nhận theo phần đầu của đáp ứng xung phòng, và bộ xử lý âm vang cuối để xử lý tín hiệu âm thanh được nhận theo âm vang cuối của đáp ứng xung phòng.

(13) Sáng chế đề xuất bộ mã hóa âm thanh để mã hóa tín hiệu âm thanh, trong đó bộ mã hóa âm thanh được cấu hình hoặc được lập trình để xử lý tín hiệu âm thanh được mã hóa theo đáp ứng xung phòng theo phương pháp của sáng chế. Bộ mã hóa âm thanh có thể bao gồm bộ phận xử lý tín hiệu theo sáng chế.

(14) Sáng chế đề xuất bộ giải mã âm thanh để giải mã tín hiệu âm thanh được mã hóa, trong đó bộ giải mã âm thanh được cấu hình hoặc được lập trình để xử lý tín

hiệu âm thanh được giải mã theo đáp ứng xung phòng theo phương pháp của sáng chế. Bộ giải mã âm thanh có thể bao gồm bộ phận xử lý tín hiệu theo sáng chế. Bộ giải mã âm thanh có thể bao gồm bộ kết xuất, như bộ kết xuất lập thể, được cấu hình hoặc được lập trình để nhận tín hiệu âm thanh được giải mã và để kết xuất tín hiệu đầu ra dựa trên đáp ứng xung phòng.

(15) Sáng chế đề xuất bộ kết xuất lập thể bao gồm bộ phận xử lý tín hiệu theo sáng chế.

### **Mô tả vắn tắt các hình vẽ**

Các phương án của sáng chế sẽ được mô tả khi tham chiếu các hình vẽ kèm theo, trong đó:

Fig.1 minh họa tổng quan về bộ mã hóa âm thanh ba chiều của hệ thống âm thanh ba chiều;

Fig.2 minh họa tổng quan về bộ giải mã âm thanh ba chiều của hệ thống âm thanh ba chiều;

Fig.3 minh họa ví dụ thực hiện bộ chuyển đổi định dạng mà có thể được thực hiện trong bộ giải mã âm thanh ba chiều trên Fig.2;

Fig.4 minh họa phương án thể hiện bộ kết xuất lập thể mà có thể được thực hiện trong bộ giải mã âm thanh ba chiều trên Fig.2;

Fig.5 minh họa ví dụ về đáp ứng xung phòng  $h(t)$ ;

Fig.6(A) minh họa biểu đồ khối của bộ phận xử lý tín hiệu (ví dụ trong bộ kết xuất lập thể trên Fig.4) để xử lý riêng tín hiệu âm thanh với phần đầu và âm vang cuối của xung phòng theo phương án của sáng chế;

Fig.6(B) minh họa biểu đồ khối của một bộ phận xử lý tín hiệu khác (ví dụ trong bộ kết xuất lập thể trên Fig.4) để xử lý riêng biệt tín hiệu âm thanh với phần đầu và âm vang cuối của xung phòng theo phương án khác của sáng chế;

Fig.7 minh họa lưu đồ của phương pháp xác định thời gian chuyển tiếp giữa các phản xạ đầu và âm vang cuối trong đáp ứng xung phòng theo phương án của sáng chế;

Fig.8 minh họa sự giảm phân rã năng lượng (energy decay relief- EDR) thu được đối với đáp ứng xung được xác định theo phương pháp dựa trên FFT;

Fig.9 minh họa cách xác định thời gian chuyển tiếp theo một phương án của sáng chế;

Fig.10 minh họa các thời gian chuyển tiếp cho kênh bên trái và kênh bên phải đối với đáp ứng xung phòng lập thể đo được được xác định bằng cách sử dụng phương pháp thông thường;

Fig.11 minh họa các thời gian chuyển tiếp cho kênh bên trái và kênh bên phải đối với đáp ứng xung phòng xung phòng đo được được xác định bằng cách sử dụng phương pháp theo sáng chế;

Fig.12 minh họa theo biểu đồ quy trình xử lý xung phòng các tín hiệu âm thanh trong bộ kết xuất lập thể theo một phương án của sáng chế; và

Fig.13 minh họa theo biểu đồ quy trình xử lý trong phần vang âm miền tần số của bộ kết xuất lập thể trên Fig.12 theo một phương án của sáng chế.

### **Mô tả chi tiết sáng chế**

Các phương án thể hiện phương pháp theo sáng chế để xử lý tín hiệu âm thanh theo đáp ứng xung phòng và để xác định, trong đáp ứng xung phòng, sự chuyển tiếp từ các phản xạ đầu sang âm vang cuối sẽ được mô tả. Phần mô tả sau đây sẽ bắt đầu bằng sự mô tả tổng quan về hệ thống mã hóa-giải mã âm thanh ba chiều trong đó phương pháp theo sáng chế có thể được thực hiện.

Các hình vẽ Fig.1 và Fig.2 thể hiện các khối thuật toán của hệ thống âm thanh ba chiều theo các phương án. Cụ thể hơn, Fig.1 là hình tổng quan về bộ mã hóa âm thanh ba chiều 100. Bộ mã hóa âm thanh 100 nhận tại mạch bộ kết xuất trước/bộ trộn 102, mà có thể tùy ý được cung cấp, tín hiệu đầu vào, cụ thể hơn là nhiều kênh đầu vào được cung cấp cho bộ mã hóa âm thanh 100 nhiều tín hiệu kênh 104, nhiều tín hiệu đối tượng 106 và siêu dữ liệu đối tượng 108 tương ứng. Các tín hiệu đối tượng 106 được xử lý bằng bộ kết xuất trước/bộ trộn 102 (xem các tín hiệu 110) có thể được cung cấp cho bộ mã hóa SAOC 112 (SAOC - spatial audio object coding = Mã hóa đối tượng âm thanh trong không gian). Bộ mã hóa SAOC 112 tạo ra các kênh vận chuyển SAOC 114 được cung cấp cho bộ mã hóa USAC 116 (USAC = Unified Speech và

Audio Coding). Ngoài ra, tín hiệu SAOC-SI 118 (SAOC-SI = SAOC thông tin phụ) cũng được cung cấp cho bộ mã hóa USAC 116. Bộ mã hóa USAC 116 còn nhận tín hiệu đối tượng 120 trực tiếp từ bộ kết xuất trước/bộ trộn cũng như các tín hiệu kênh và tín hiệu đối tượng được kết xuất trước 122. Thông tin siêu dữ liệu đối tượng 108 được sử dụng cho bộ mã hóa OAM 124 (OAM = object metadata – siêu dữ liệu đối tượng) cung cấp thông tin siêu dữ liệu đối tượng được nén 126 cho bộ mã hóa USAC. Bộ mã hóa USAC 116, dựa trên các tín hiệu đầu vào nêu trên, tạo ra tín hiệu đầu ra được nén mp4, như được thể hiện bởi số chỉ dẫn 128.

Fig.2 là hình tổng quan về bộ giải mã âm thanh ba chiều 200 của hệ thống âm thanh ba chiều. Tín hiệu được mã hóa 128 (mp4) được tạo ra bởi bộ mã hóa âm thanh 100 trên Fig.1 được thu nhận tại bộ giải mã âm thanh 200, cụ thể hơn là tại bộ giải mã USAC 202. Bộ giải mã USAC 202 giải mã tín hiệu được nhận 128 thành tín hiệu kênh 204, tín hiệu đối tượng được kết xuất trước 206, tín hiệu đối tượng 208, và tín hiệu kênh vận chuyển SAOC 210. Ngoài ra, thông tin siêu dữ liệu đối tượng được nén 212 và tín hiệu SAOC-SI 214 được xuất ra bởi bộ giải mã USAC 202. Tín hiệu đối tượng 208 được cung cấp cho bộ kết xuất đối tượng 216 cung cấp tín hiệu đối tượng được kết xuất 218. Các tín hiệu kênh vận chuyển SAOC 210 được cung cấp đến bộ giải mã SAOC 220 cung cấp các tín hiệu đối tượng được kết xuất 222. Thông tin siêu đối tượng được nén 212 được cung cấp đến bộ giải mã OAM 224 đưa các tín hiệu điều khiển tương ứng đến bộ kết xuất đối tượng 216 và bộ giải mã SAOC 220 để tạo ra tín hiệu đối tượng được kết xuất 218 và tín hiệu đối tượng được kết xuất 222. Bộ giải mã còn bao gồm bộ trộn 226, như được thể hiện trên Fig.2, nhận các tín hiệu đầu vào 204, 206, 218 và 222 để xuất ra các tín hiệu kênh 228. Các tín hiệu kênh có thể được trực tiếp đưa đến loa phát thanh, ví dụ loa phát thanh kênh 32, như được biểu thị tại 230. Tín hiệu 228 có thể được cung cấp đến mạch chuyển đổi định dạng 232 nhận, để làm phân đầu vào điều khiển, tín hiệu sắp xếp tái tạo biểu diễn cách mà tín hiệu kênh 228 được chuyển đổi. Theo phương án được mô tả trên Fig.2, giả thiết rằng sự chuyển đổi được thực hiện theo cách sao cho các tín hiệu có thể được cung cấp đến hệ thống loa 5.1 như được biểu thị tại 234. Tương tự, tín hiệu kênh 228 có thể được cung cấp đến bộ kết xuất lập thể 236 tạo ra hai tín hiệu đầu ra, ví dụ dùng cho ống nghe, như được biểu thị tại 238.

Theo phương án của sáng chế, hệ thống mã hóa/giải mã được mô tả trên Fig.1 và Fig.2 dựa trên bộ mã hóa-giải mã MPEG-D USAC để mã hóa các tín hiệu đối tượng và tín hiệu kênh (xem các tín hiệu 104 và 106). Để gia tăng hiệu suất để lập mã lượng lớn đối tượng, kỹ thuật MPEG SAOC có thể được sử dụng. Ba loại bộ kết xuất có thể thực hiện các nhiệm vụ kết xuất đối tượng tới các kênh, kết xuất kênh tới ống nghe hoặc kết xuất kênh sang thiết lập loa phát thanh khác (xem Fig.2, các số tham chiếu 230, 234 và 238). Khi tín hiệu đối tượng được truyền rõ ràng hoặc được mã hóa tham số bằng cách sử dụng SAOC, thông tin siêu dữ liệu đối tượng tương ứng 108 được nén (xem tín hiệu 126) và được dồn kênh thành dòng bit âm thanh ba chiều 128.

Các khối thuật toán của toàn bộ hệ thống âm thanh ba chiều được thể hiện trên Fig.1 và Fig.2 sẽ được mô tả chi tiết hơn dưới đây.

Bộ kết xuất trước/bộ trộn 102 có thể tùy ý được cung cấp để chuyển đổi màn hình đầu vào đối tượng và kênh thành màn hình kênh trước khi mã hóa. Về mặt chức năng, nó giống với bộ kết xuất/bộ trộn đối tượng mà được mô tả dưới đây. Cần thực hiện kết xuất trước đối tượng để đảm bảo entropy tín hiệu quyết định tại bộ mã hóa đầu vào mà về cơ bản là độc lập với số lượng tín hiệu đối tượng hoạt động đồng thời. Với việc kết xuất trước các đối tượng, thì không cần phải truyền siêu dữ liệu đối tượng. Các tín hiệu đối tượng rời rạc được kết xuất thành bố trí kênh mà bộ mã hóa được cấu hình để sử dụng. Trọng số của các đối tượng cho mỗi kênh thu được từ siêu dữ liệu đối tượng đi kèm (associated object metadata - OAM).

Bộ mã hóa USAC 116 là bộ mã hóa-giải mã chính cho các kênh kênh loa phát thanh, các tín hiệu đối tượng rời rạc, các tín hiệu trộn giảm đối tượng và các tín hiệu được kết xuất trước. Điều đó dựa trên kỹ thuật MPEG-DUSAC. Kỹ thuật này xử lý mã hóa các tín hiệu nêu trên bằng cách tạo ra thông tin ánh xạ kênh-và-đối tượng dựa trên thông tin hình học và ngữ nghĩa của sự phân định kênh đầu vào và đối tượng. Thông tin ánh xạ này mô tả phương pháp mà các kênh đầu vào và đối tượng được ánh xạ đến các yếu tố kênh USAC, như yếu tố hai kênh (CPE - channel pair elements), yếu tố một kênh (SCE - single channel elements), ảnh hưởng tần số thấp (LFE - low frequency effects) và yếu tố bốn kênh (QCE - quad channel elements) và CPE, SCE và LFE, và thông tin tương ứng được truyền đến bộ giải mã. Tất cả tài hữu ích bổ sung như dữ

liệu SAOC 114, 118 hoặc siêu dữ liệu đối tượng 126 được xem xét khi điều chỉnh tốc độ của bộ mã hóa. Mã hóa đối tượng có thể thực hiện theo nhiều cách khác nhau, phụ thuộc vào các yêu cầu về tốc độ/biến dạng và các yêu cầu về tính tương tác cho bộ kết xuất. Theo các phương án, các biến thể mã hóa đối tượng có thể là:

- *Các đối tượng được kết xuất trước:* Các tín hiệu đối tượng được kết xuất trước và được trộn với tín hiệu kênh 22.2 trước khi mã hóa. Chuỗi mã hóa kế tiếp thấy được các tín hiệu kênh 22.2.
- *Các dạng sóng đối tượng rời rạc:* Các đối tượng được cung cấp như dạng sóng đơn âm cho bộ mã hóa. Bộ mã hóa sử dụng các bộ phận một kênh (SCE) để truyền đối tượng ngoài các tín hiệu kênh. Đối tượng được giải mã được kết xuất và trộn tại phía bộ nhận. Thông tin siêu dữ liệu đối tượng được nén được truyền đến bộ nhận/bộ kết xuất.
- *Các dạng sóng đối tượng tham số:* Các đặc tính của đối tượng và mối tương quan của chúng với nhau được mô tả bằng các tham số SAOC. Trộn giảm của tín hiệu đối tượng được mã hóa với USAC. Thông tin tham số được truyền kế bên. Số lượng kênh trộn giảm được chọn phụ thuộc vào số lượng đối tượng và tốc độ dữ liệu tổng. Thông tin siêu dữ liệu đối tượng được nén được truyền đến bộ kết xuất SAOC.

Bộ mã hóa SAOC 112 và bộ giải mã SAOC 220 cho các tín hiệu đối tượng có thể dựa trên kỹ thuật MPEG SAOC. Hệ thống này có khả năng tái tạo, biến đổi và kết xuất số lượng các đối tượng âm thanh dựa trên số lượng các kênh được truyền nhỏ hơn và dữ liệu theo tham số bổ sung, như OLD, IOC (Inter Object Coherence), DMG (Encoding Gains). Dữ liệu theo tham số bổ sung thể hiện tốc độ dữ liệu thấp hơn một cách đáng kể so với yêu cầu cần có để truyền riêng rẽ tất cả các đối tượng, khiến cho việc mã hóa hiệu quả. Bộ mã hóa SAOC 112 lấy các tín hiệu đối tượng/kênh làm dạng sóng đơn âm và cung cấp thông tin theo tham số (mà được tập hợp thành dòng bit âm thanh ba chiều 128) và kênh vận chuyển SAOC (mà được mã hóa bằng cách sử dụng các bộ phận một kênh và được truyền). Bộ giải mã SAOC 220 khôi phục tín hiệu đối tượng/kênh từ kênh vận chuyển SAOC được giải mã 210 và thông tin theo tham số 214, và tạo ra màn hình âm thanh đầu ra dựa trên sự bố trí tái tạo, thông tin siêu dữ liệu đối tượng được giải nén và tùy ý dựa trên thông tin tương tác của người dùng.

Bộ mã hóa-giải mã siêu dữ liệu đối tượng (xem bộ mã hóa OAM 124 và bộ giải mã OAM 224) được đề xuất sao cho, đối với mỗi đối tượng, siêu dữ liệu đi kèm sẽ cụ thể hóa vị trí hình học và âm lượng của đối tượng trong không gian ba chiều được mã hóa một cách hiệu quả bằng sự lượng tử hóa các đặc tính đối tượng theo thời gian và không gian. Siêu dữ liệu đối tượng được nén cOAM 126 được truyền đến bộ nhận 200 để làm thông tin phụ.

Bộ kết xuất đối tượng 216 sử dụng siêu dữ liệu đối tượng được nén để tạo ra các dạng sóng đối tượng theo định dạng tái tạo đã xác định. Mỗi đối tượng được kết xuất đến kênh đầu ra xác định theo thông tin siêu dữ liệu của nó. Đầu ra của khối này thu được từ tổng các kết quả thành phần. Nếu nội dung dựa trên cả hai kênh cũng như các đối tượng tham số/rời rạc được giải mã, thì dạng sóng dựa trên kênh và dạng sóng đối tượng được kết xuất được trộn bằng bộ trộn 226 trước khi tạo ra dạng sóng cuối cùng 228 hoặc trước khi cấp chúng đến mô-đun bộ xử lý sau như bộ kết xuất lập thể 236 hoặc mô-đun bộ kết xuất loa phát thanh 232.

Mô-đun bộ kết xuất lập thể 236 tạo ra trộn giảm lập thể vật liệu âm thanh đa kênh sao cho mỗi kênh đầu vào được thể hiện bởi một nguồn âm thanh ảo. Việc xử lý được thực hiện trên khung trong miền QMF (QMF - Quadrature Mirror Filterbank), và lập thể hóa âm thanh dựa trên đáp ứng xung phòng lập thể đo được.

Bộ kết xuất loa phát thanh 232 chuyển đổi giữa cấu hình kênh được truyền 228 và định dạng tái tạo mong muốn. Nó cũng có thể được gọi là “bộ chuyển đổi định dạng”. Bộ chuyển đổi định dạng thực hiện việc chuyển đổi sang số lượng các kênh đầu ra thấp hơn, tức là, nó tạo ra quá trình trộn giảm.

Fig.3 thể hiện ví dụ thực hiện bộ chuyển đổi định dạng 232. Bộ chuyển đổi định dạng 232, cũng được đề cập đến như bộ kết xuất loa phát thanh, chuyển đổi giữa cấu hình kênh máy phát và định dạng tái tạo mong muốn. Bộ chuyển đổi định dạng 232 thực hiện chuyển đổi sang số lượng kênh đầu ra thấp hơn, tức là, nó thực hiện quy trình trộn giảm (DMX) 240. Bộ trộn giảm 240, tốt hơn là thực hiện trong miền QMF, nhận các tín hiệu đầu ra của bộ trộn 228 và tạo ra tín hiệu loa phát thanh 234. Bộ cấu hình 242, cũng được đề cập đến là bộ điều khiển, có thể được tạo ra, làm đầu vào điều khiển, để nhận tín hiệu 246 biểu thị bố trí đầu ra bộ trộn, tức là, sự bố trí mà dữ liệu

được thể hiện bởi tín hiệu đầu ra của bộ trộn 228 được xác định, và tín hiệu 248 biểu thị bố trí tái tạo mong muốn. Dựa trên thông tin này, bộ điều khiển 242, tốt hơn là tự động, tạo ra ma trận trộn giảm tối ưu hóa để tổ hợp các định dạng đầu vào và đầu ra và sử dụng các ma trận này cho bộ trộn giảm 240. Bộ chuyển đổi định dạng 232 cho phép các cấu hình loa phát thanh tiêu chuẩn cũng như các cấu hình ngẫu nhiên với vị trí loa phát thanh không tiêu chuẩn.

Fig.4 minh họa phương án thể hiện bộ kết xuất lập thể 236 trên Fig.2. Mô-đun bộ kết xuất lập thể có thể tạo ra việc trộn giảm lập thể vật liệu âm thanh đa kênh. Lập thể hóa âm thanh có thể dựa trên đáp ứng xung phòng lập thể đo được. Đáp ứng xung phòng có thể được xem như “công nghệ phân tích âm thanh” của các đặc tính âm thanh của phòng thật. Đáp ứng xung phòng được đo và được lưu trữ, và các tín hiệu âm thanh tùy ý có thể được tạo ra với “công nghệ phân tích âm thanh”, nhờ đó cho phép mô phỏng các đặc tính âm thanh của phòng tổ hợp với đáp ứng xung phòng tại phía người nghe. Bộ kết xuất lập thể 236 có thể được lập trình hoặc được cấu hình để kết xuất các kênh đầu ra thành các kênh lập thể sử dụng chức năng chuyển phân đầu hoặc đáp ứng xung phòng lập thể (BRIR). Ví dụ, đối với thiết bị di động việc kết xuất lập thể cần phải có cho ống nghe hoặc loa phát thanh được đính kèm với thiết bị di động. Trong thiết bị di động như vậy, do các hạn chế nên có thể cần giới hạn bộ giải mã và độ phức tạp kết xuất. Ngoài việc loại bỏ giải tương quan theo cách xử lý như vậy, có thể được ưu tiên là thực hiện trước tiên trộn giảm sử dụng bộ trộn giảm 250 đối với tín hiệu giải mã trung gian 252, tức là đối với số lượng các kênh đầu ra thấp hơn mà tạo ra số lượng các kênh đầu vào thấp hơn cho bộ chuyển đổi lập thể thực tế 254. Ví dụ, vật liệu kênh 22.2 có thể được trộn giảm bằng bộ trộn giảm 250 đối với giải mã trung gian 5.1 hoặc, theo cách khác, trộn giảm trung gian có thể trực tiếp được tính toán bằng bộ giải mã SAOC 220 trên Fig.2 ở trạng thái “rút gọn”. Sau đó, việc kết xuất lập thể phải áp dụng 10 HRTF (HRTF - Head Related Transfer Functions) hoặc chức năng BRIR để kết xuất 5 kênh riêng rẽ tại các vị trí khác nhau trái với việc áp dụng 44 chức năng HRTF hoặc BRIR nếu các kênh đầu vào 22.2 được trực tiếp kết xuất. Thao tác nhân chấp cần thiết để kết xuất lập thể cần nhiều năng lượng xử lý và, do đó, làm giảm năng lượng xử lý này khi vẫn thu được chất lượng âm thanh chấp nhận được hữu ích dùng cho thiết bị di động. Bộ kết xuất lập thể 236 tạo ra trộn giảm

kép 238 của vật liệu âm thanh đa kênh 228, sao cho mỗi kênh đầu vào (không bao gồm các kênh LFE) được thể hiện bằng một nguồn âm thanh ảo. Quy trình xử lý có thể được thực hiện trên khung trong miền QMF. Lập thể hóa âm thanh dựa trên đáp ứng xung phòng lập thể đo được, và âm thanh trực tiếp và các phản xạ đầu có thể được ghi vào vật liệu âm thanh thông qua phương pháp nhân chập trong miền FFT giả sử dụng phép nhân chập nhanh của miền QMF, còn âm vang cuối có thể được xử lý riêng rẽ.

Fig.5 thể hiện ví dụ về đáp ứng xung phòng  $h(t)$  300. Đáp ứng xung phòng bao gồm ba thành phần, âm thanh trực tiếp 301, các phản xạ đầu 302 và âm vang cuối 304. Do đó, đáp ứng xung phòng mô tả phương pháp phản xạ không gian âm thanh vang âm khi tiến hành xung. Các phản xạ đầu 302 là các phản xạ rời rạc có mật độ gia tăng, và phần đáp ứng xung trong đó các phản xạ riêng rẽ không tách rời được gọi là âm vang cuối 304. Âm thanh trực tiếp 301 có thể dễ dàng được nhận dạng trong đáp ứng xung phòng và có thể được tách khỏi các phản xạ đầu, tuy nhiên, sự chuyển tiếp từ phản xạ đầu 302 sang âm vang cuối 304 ít rõ ràng hơn.

Các phương án của sáng chế dưới đây được mô tả chi tiết hơn. Theo các phương án của sáng chế, tín hiệu âm thanh được xử lý riêng rẽ với phần đầu và âm vang cuối của đáp ứng xung phòng. Tín hiệu âm thanh được xử lý với phần đầu của đáp ứng xung phòng và tín hiệu vang âm được tổ hợp và đi ra làm tín hiệu âm thanh đầu ra. Cần biết đến quy trình xử lý riêng rẽ sự chuyển tiếp trong đáp ứng xung phòng từ phần đầu sang âm vang cuối. Sự chuyển tiếp được xác định bằng số đo tương quan mà đạt đến ngưỡng, trong đó ngưỡng này được thiết lập phụ thuộc vào số đo tương quan đối với một trong số các phản xạ đầu được chọn trong phần đầu của đáp ứng xung phòng. Số đo tương quan có thể mô tả, đối với đáp ứng xung phòng, tính tương tự của sự suy yếu trong năng lượng âm thanh bao gồm trạng thái ban đầu và sự suy yếu trong năng lượng âm thanh bắt đầu tại thời điểm bất kỳ sau trạng thái ban đầu trong khoảng tần số định trước.

Theo các phương án, quy trình xử lý tín hiệu âm thanh riêng rẽ bao gồm xử lý tín hiệu âm thanh với phần phản xạ đầu 301, 302 của đáp ứng xung phòng trong quá trình xử lý thứ nhất, và xử lý tín hiệu âm thanh với vang âm khuếch tán 304 của đáp ứng xung phòng trong suốt quy trình thứ hai mà là quy trình khác với và tách khỏi quy

trình thứ nhất. Việc thay đổi từ quy trình thứ nhất sang quy trình thứ hai xảy ra tại thời gian chuyển tiếp. Theo các phương án khác, trong quy trình thứ hai vang âm khuếch tán (cuối cùng) 304 có thể được thay thế bằng vang âm tổng hợp. Trong trường hợp này, đáp ứng xung phòng được tạo ra có thể chứa chỉ phần phản xạ đầu 301, 302 (xem Fig.5) và không chứa vang âm khuếch tán cuối cùng 304.

Fig.6(A) thể hiện biểu đồ khối minh họa bộ phận xử lý tín hiệu thứ nhất để xử lý riêng tín hiệu âm thanh với phần đầu và âm vang cuối của xung phòng theo phương án của sáng chế. Xử lý tín hiệu âm thanh theo các phần đáp ứng xung phòng khác có thể được thực hiện trong bộ kết xuất lập thể 236 được mô tả ở trên. Tín hiệu đầu vào âm thanh 400 có thể là vật liệu âm thanh không vang âm, ví dụ tín hiệu đầu vào âm thanh đa kênh, được tổ hợp với đáp ứng xung phòng, ví dụ đáp ứng xung phòng được đo bằng cách sử dụng micrô nhân tạo trong tai hoặc qua đầu. Phép nhân chập cho phép thu được ấn tượng về không gian của vật liệu âm thanh không vang âm cứ như thể vật liệu âm thanh này được nghe trong phòng tổ hợp với đáp ứng xung phòng. Ví dụ, trong bộ kết xuất lập thể 236 nêu trên, cần xử lý tín hiệu âm thanh với âm thanh trực tiếp 301 và phản xạ đầu 302 trong đáp ứng xung phòng và để xử lý tín hiệu âm thanh với âm vang cuối 304 một cách riêng rẽ. Để xử lý tín hiệu đầu vào âm thanh 400, khối 402 để xử lý âm thanh trực tiếp, khối 404 để xử lý các phản xạ đầu và khối 406 để xử lý âm vang cuối được đề xuất. Tín hiệu đầu ra 408 và 410 của các khối 402 đến 406 tương ứng được tổ hợp bởi bộ công thứ nhất 412 để tạo ra tín hiệu được xử lý trước 414. Tín hiệu được xử lý trước 414 và tín hiệu vang âm 416 được tạo ra bởi bộ xử lý 406 được tổ hợp bởi bộ cộng thứ hai 418 để tạo ra tín hiệu đầu ra âm thanh 420 mà tạo ấn tượng cho người nghe cứ như thể tín hiệu âm thanh được nghe thấy ở trong phòng tổ hợp với đáp ứng xung phòng.

Xử lý âm vang cuối 302 tách khỏi âm thanh trực tiếp và các phản xạ đầu là có lợi do độ phức tạp tính toán được giảm đi. Cụ thể hơn, sử dụng phép nhân chập cho toàn bộ đáp ứng xung được tính toán là rất tốn kém. Do đó, thuật toán vang âm với độ phức tạp thấp hơn được sử dụng để xử lý tín hiệu âm thanh để mô phỏng âm vang cuối. Phần âm thanh trực tiếp và phản xạ đầu của đáp ứng xung được tính toán chính xác hơn, ví dụ bằng kết hợp/nhân chập. Một ưu điểm nữa đó là khả năng điều chỉnh vang âm. Khả năng này cho phép âm vang cuối được biến đổi, ví dụ, phụ thuộc phụ

thuộc đầu vào của người dùng, tham số phòng đo được hoặc phụ thuộc vào nội dung tín hiệu âm thanh. Để đạt được các ưu điểm nêu trên, cần biết đến sự chuyển tiếp (ví dụ điểm thời gian) trong đó các phản xạ đầu 302 kết thúc và trong đó âm vang cuối 304 bắt đầu. Khi quy trình xử lý âm vang cuối bắt đầu quá sớm, tín hiệu âm thanh có thể có chất lượng thấp hơn khi con người có thể nghe thấy các phản xạ đầu bị thiếu. Mặt khác, nếu thời gian chuyển tiếp được phát hiện quá muộn, hiệu suất tính toán sẽ không đạt được, do việc xử lý các phản xạ đầu tốn kém hơn việc xử lý âm vang cuối. Sự chuyển tiếp, ví dụ trong các mẫu miền thời gian, có thể được cấp đến bộ kết xuất lập thể để làm tham số đầu vào mà sau đó, phụ thuộc vào sự chuyển tiếp được nhận, điều khiển các bộ xử lý 402 đến 406 để xử lý riêng tín hiệu âm thanh.

Fig.6(B) minh họa biểu đồ khối thể hiện bộ phận xử lý tín hiệu làm ví dụ khác để xử lý riêng tín hiệu âm thanh với phần đầu và âm vang cuối của xung phòng theo phương án khác của sáng chế. Tín hiệu đầu vào 400, ví dụ tín hiệu đầu vào âm thanh đa kênh, được nhận và sử dụng cho bộ xử lý thứ nhất 422 để xử lý phần đầu, cụ thể là để xử lý tín hiệu âm thanh theo âm thanh trực tiếp 301 và các phản xạ đầu 302 trong đáp ứng xung phòng 300 được thể hiện trên Fig.5. Tín hiệu đầu vào âm thanh đa kênh 400 cũng được sử dụng cho bộ xử lý thứ hai 424 để xử lý tín hiệu âm thanh theo âm vang cuối 304 của đáp ứng xung phòng. Trong bộ kết xuất lập thể, như nêu trên, cần phải xử lý âm thanh trực tiếp và các phản xạ đầu tách khỏi âm vang cuối, chủ yếu là vì độ phức tạp tính toán được giảm đi. Xử lý âm thanh trực tiếp và các phản xạ đầu, ví dụ, có thể được ghi thành tín hiệu âm thanh bằng cách nhân chập được thực hiện bởi bộ xử lý thứ nhất 422, còn âm vang cuối có thể được thay thế bằng vang âm tổng hợp được tạo ra bởi bộ xử lý thứ hai 424. Tín hiệu đầu ra lập thể tổng 420 sau đó là sự tổ hợp của kết quả nhân chập 428 được tạo ra bởi bộ xử lý 422 và tín hiệu vang âm tổng hợp 430 được tạo ra bởi bộ xử lý 424. Theo các phương án, các tín hiệu 428 và 430 được tổ hợp bởi bộ cộng 432 cung cấp tín hiệu đầu ra lập thể tổng 420.

Như đã nêu, bộ xử lý thứ nhất 422 có thể tạo ra nhân chập của tín hiệu đầu vào âm thanh 400 với âm thanh trực tiếp và các phản xạ đầu của đáp ứng xung phòng mà có thể được tạo ra cho bộ xử lý thứ nhất 422 từ cơ sở dữ liệu bên ngoài 434 chứa nhiều đáp ứng xung phòng lập thể được ghi. Bộ xử lý thứ hai hoặc phần vang âm 424 có thể thực hiện dựa trên tham số phần vang âm, như vang âm RT60 và năng lượng vang âm,

mà có thể thu được từ đáp ứng xung phòng lập thể được lưu trữ bằng phân tích 436. Lưu ý rằng, phân tích 436 không nhất thiết là một phần của bộ kết xuất, thay vào đó điều này cho thấy rằng từ các xung tương ứng được lưu trữ trong cơ sở dữ liệu 434 thì các tham số vang âm tương ứng có thể được tạo ra; điều này có thể được thực hiện bên ngoài. Tham số phần vang âm có thể được xác định, ví dụ, bằng cách tính toán năng lượng và thời gian vang âm RT60 theo phân tích giàn lọc octa hoặc 1 phần 3 octa, hoặc có thể là các giá trị trung bình của các kết quả của nhiều phân tích đáp ứng xung.

Ngoài ra, cả hai bộ xử lý 422 và 424 nhận từ cơ sở dữ liệu 434 – trực tiếp hoặc thông qua phân tích 436 – là tham số đầu vào thông tin về sự chuyển tiếp trong đáp ứng xung phòng từ phần đầu sang âm vang cuối. Sự chuyển tiếp có thể được xác định theo cách như được mô tả chi tiết dưới đây.

Theo các phương án, phân tích chuyển tiếp có thể được sử dụng để tách các phản xạ đầu và âm vang cuối. Nó có thể được cung cấp đến bộ kết xuất lập thể như tham số đầu vào (ví dụ nó có thể được đọc từ tệp tin/giao diện riêng cùng với các giá trị RT60 và giá trị năng lượng mà được sử dụng để cấu hình phần vang âm). Phân tích này có thể dựa trên một tập hợp các đáp ứng xung phòng lập thể (tập hợp các cặp BRIR cho vô số các góc phương vị và góc nâng). Phân tích này có thể là bước xử lý trước mà được thực hiện riêng rẽ cho mỗi đáp ứng xung và sau đó lấy số trung bình của các giá trị chuyển tiếp làm giá trị chuyển tiếp tổng của một tập hợp BRIR. Giá trị chuyển tiếp tổng này có thể được sử dụng để tách các phản xạ đầu từ âm vang cuối theo tính toán tín hiệu đầu ra lập thể.

Tuy nhiên, các phương pháp để xác định sự chuyển tiếp đã biết có bất lợi như được mô tả ngay sau đây. Theo tài liệu [1] của kỹ thuật hiện nay, phương pháp được mô tả mà sử dụng sự giảm phân rã năng lượng (*EDR*) và số đo tương quan để xác định thời gian chuyển tiếp từ các phản xạ đầu sang âm vang cuối. Tuy nhiên, phương pháp được mô tả trong tài liệu [1] là bất lợi.

1. Phương pháp này phụ thuộc chủ yếu vào góc phương vị của đáp ứng xung lập thể và mối tương quan giữa biên độ của âm thanh trực tiếp và phản xạ va chạm thứ nhất.

2. Thời gian chuyển tiếp được tính toán trong bảng tần tùy ý. Không có hiểu biết thông thường về giải tần số nào đem lại thời gian chuyển tiếp đúng để được sử dụng cho đáp ứng xung tổng.

3. Không có thông tin nào về bước tương quan cần thiết của phương pháp này.

Phương pháp đã biết khác mô tả các phản xạ đầu bằng cách phân tán tiếng vọng trong không gian, ví dụ bằng số lượng trung bình của các phản xạ trên mỗi giây, và xác định thời điểm bắt đầu của âm vang cuối khi số lượng này vượt quá ngưỡng định trước ngưỡng (xem tài liệu của kỹ thuật hiện nay [2]). Phương pháp này phụ thuộc vào đặc tính của phòng, cụ thể là thể tích phòng, mà thường không được biết đến. Thể tích phòng không thể dễ dàng được chiết xuất khỏi đáp ứng xung đo được. Do đó, phương pháp không thể áp dụng để tính toán sự chuyển tiếp từ đáp ứng xung đo được. Tương tự, không có hiểu biết thông thường nào về độ dày đặc mà các phản xạ cần có để được gọi là âm vang cuối.

Một khả năng khác, được mô tả trong tài liệu của kỹ thuật hiện nay [3], là so sánh sự phân bố thực tế tại thời gian trong cửa sổ đáp ứng xung với sự phân bố Gaussian trong miền thời gian. Âm vang cuối được giả thiết là có sự phân bố thông thường. Trong sự phân bố thông thường, gần một phần ba (chính xác là  $1/e$ ) của các mẫu nằm ngoài độ lệch chuẩn của giá trị trung bình và hai phần ba các mẫu nằm trong khoảng độ lệch chuẩn của giá trị trung bình. Các phản xạ đầu rõ rệt có nhiều mẫu nằm trong khoảng độ lệch chuẩn và có ít hơn ở bên ngoài. Tỷ lệ các mẫu nằm ngoài độ lệch chuẩn so với các mẫu nằm trong khoảng độ lệch chuẩn có thể được sử dụng để xác định thời gian chuyển tiếp. Tuy nhiên, điểm bất lợi của phương pháp này đó là sự chuyển tiếp khó có thể được xác định bằng phương pháp này, bởi vì tỷ lệ đôi khi dao động quanh ngưỡng. Phương pháp này cũng phụ thuộc chủ yếu vào kích thước và loại cửa sổ trượt trong đó các tỷ lệ được tính toán.

Ngoài các phương pháp nêu trên, Độ nhọn (nửa bất biến cao hơn của tín hiệu ngẫu nhiên) cũng có thể được sử dụng để xác định thời gian chuyển tiếp. Nó làm giảm nhanh khi tiếp cận về phía phần cuối cùng của đáp ứng xung, như được tóm tắt trong tài liệu của kỹ thuật hiện nay [4]. Tuy nhiên, định nghĩa về ngưỡng đối với sự chuyển

tiếp (sử dụng sự giảm nhanh hoặc thời gian khi lần thứ nhất nó đạt đến 0) là không rõ ràng.

Tuy nhiên, có phương pháp khác mà không phụ thuộc vào phân tích đáp ứng xung đo được, nhưng phụ thuộc vào thể tích phòng, như được mô tả trong tài liệu [2]. Phương pháp này giả thiết rằng thời gian chuyển tiếp chỉ phụ thuộc vào thể tích, mà không xem xét đến các đặc tính khuếch tán của các vùng biên. Do đó, kết quả có thể chỉ là giá trị gần của thời gian chuyển tiếp và không chính xác như cần thiết để tránh các bất lợi nêu trên khi không xác định chính xác thời gian chuyển tiếp. Hơn nữa, thể tích của phòng thường không được biết đến và không thể dễ dàng được chiết xuất từ đáp ứng xung đo được.

Các phương pháp khác đã biết hoàn toàn bỏ qua yếu tố môi trường và xác định thời gian chuyển tiếp đơn giản là 80ms, ví dụ xem tài liệu của kỹ thuật hiện nay [5]. Tuy nhiên, con số này hoàn toàn tách rời với các đặc tính của phòng hoặc đáp ứng xung đo được và, do đó, không chính xác để tách âm vang cuối khỏi phần nhắc của đáp ứng xung.

Theo các phương án, ngoài quy trình xử lý tín hiệu âm thanh được cải thiện, sáng chế cũng đề xuất phương pháp được cải thiện để xác định thời gian chuyển tiếp giữa các phản xạ đầu và âm vang cuối trong đáp ứng xung phòng tạo ra cách xác định thời gian chuyển tiếp chính xác hơn. Các phương án, sẽ được mô tả dưới đây, tạo ra khả năng hiệu quả và đơn giản để tính toán thời gian chuyển tiếp từ đáp ứng xung đo được bằng cách sử dụng phân tích FFT.

Fig.7 là lưu đồ thể hiện phương pháp để xác định thời gian chuyển tiếp giữa các phản xạ đầu và âm vang cuối trong đáp ứng xung phòng theo phương án của sáng chế. Để xác định thời gian chuyển tiếp từ các phản xạ đầu sang âm vang cuối, ở bước thứ nhất 500 sự phân bố thời gian-tần số của năng lượng âm thanh được xác định. Ví dụ, theo các phương án sự giảm phân rã năng lượng ( $E(t,f)$ ,  $EDR$ ) có thể được tính toán ở bước 500.  $EDR$  có thể được tính toán trực tiếp từ đáp ứng xung phòng đo được (ví dụ đáp ứng lập thể) và có thể được diễn giải là sự mở rộng miền-tần số của đường cong suy yếu năng lượng thường được sử dụng (Schroeder integration, EDC (d)) mà thể hiện năng lượng còn lại trong đáp ứng xung sau thời gian  $t$ . Thay vì sử dụng đáp ứng

xung bằng thông rộng, *EDR* được suy ra từ sự biểu diễn thời gian-tần số và nhiều phép biểu diễn thời gian-tần số khác có thể được sử dụng cho mục đích này. Khi sự phân bố thời gian-tần số của năng lượng âm thanh được xác định ở bước 500, thì ở bước 502 số đo tương quan giữa năng lượng âm thanh tại khối thời gian phân bố thời gian-tần số và tổng năng lượng âm thanh ở trạng thái ban đầu được xác định. Ở bước 504, xác định xem liệu số đo tương quan đạt đến ngưỡng xác định (ví dụ rơi vào khoảng dưới ngưỡng xác định) hay không. Nếu số đo tương quan không đạt đến ngưỡng, phương pháp tiến hành tiếp tục bước 506 trong đó khối thời gian tiếp theo và sự phân bố sau khối thời gian hiện tại được chọn và các bước 502 và 504 được lặp lại cho khối thời gian tiếp theo. Do đó, theo các bước 502 đến 506 thì số đo tương quan được sử dụng để tính toán giá trị tương quan giữa mỗi khối thời gian của *EDR* được xác định ở bước 500 với tổng năng lượng ở trạng thái ban đầu. Thời gian chuyển tiếp đạt được khi số đo tương quan đạt đến ngưỡng xác định (ví dụ rơi vào khoảng dưới ngưỡng xác định). Nói cách khác, khi xác định theo bước 504 rằng đối với khối thời gian hiện tại thì số đo tương quan thấp hơn ngưỡng, phương pháp tiến hành tiếp đến bước 508 trong đó thời gian của khối thời gian hiện tại được coi như thời gian chuyển tiếp.

Sau đây, phương pháp theo phương án của sáng chế sẽ được mô tả chi tiết hơn. Lúc đầu, đáp ứng xung lập thể đo được có thể được sử dụng làm đầu vào để tính toán thời gian chuyển tiếp. Sau đó, phân bố Page hoặc Levin được sử dụng để tính toán sự giảm phân rã năng lượng (*EDR*). Sự phân bố Page đề cập đến đạo hàm của phổ trước và sự phân bố Page của tín hiệu ngược thời điểm được gọi là sự phân bố Levin (xem tài liệu của kỹ thuật hiện nay [2]). Sự phân bố này mô tả phổ công suất tức thì, và *EDR* của đáp ứng xung  $h(t)$  (xem, ví dụ, Fig.5) được tính toán như sau:

$$E(t, \omega) = \left| \int_t^\infty h(\tau) e^{-j\omega\tau} d\tau \right|^2$$

trong đó,

$E(t, \omega)$  = sự giảm phân rã năng lượng,

$h(\tau)$  = đáp ứng xung phòng,

$\omega$  =  $2\pi f$ .

Tính toán theo phương trình nêu trên bắt đầu tại âm thanh trực tiếp 301 (xem Fig.5), và với thời gian gia tăng thì sự giảm phân rã năng lượng chứa ít các phản xạ khác biệt hơn và nhiều vang âm ngẫu nhiên hơn. Theo phương án được mô tả, sự giảm phân rã năng lượng được tính toán cho khối thời gian có độ dài là 1ms để dễ dàng tính toán. Bằng cách sử dụng chức năng nêu trên, sự phân bố thời gian-tần số của năng lượng âm thanh được xác định như được mô tả theo bước 500 trên Fig.7.

Sau đó, như được mô tả theo các bước 502 đến 506 trên Fig.7, số đo tương quan  $\rho(t)$  mà dựa trên tương quan momen tích (còn được biết đến là hệ số tương quan) được xác định. Cụ thể hơn, tương quan của năng lượng âm thanh đối với mỗi khối thời gian với tổng năng lượng ở trạng thái ban đầu được xác định, theo các phương án, là như sau:

$$\rho(t) = \frac{\sum_{\omega}(E(1, \omega) - \bar{E}(1, \omega)) \cdot \sum_{\omega}(E(t, \omega) - \bar{E}(t, \omega))}{\sqrt{\sum_{\omega}(E(1, \omega) - \bar{E}(1, \omega))^2} \cdot \sqrt{\sum_{\omega}(E(t, \omega) - \bar{E}(t, \omega))^2}}$$

trong đó,

$E(1, \omega) =$  sự giảm phân rã năng lượng trong toàn bộ khoảng tần số tại tần số  $f$ ,

$\bar{E}(1, \omega) =$  giá trị trung bình của tổng tần số của sự giảm phân rã năng lượng trong toàn bộ khoảng ban đầu,

$E(t, \omega) =$  sự giảm phân rã năng lượng tại tần số  $f$  bắt đầu tại thời gian  $t$ ,

$\bar{E}(t, \omega) =$  giá trị trung bình của tổng tần số của sự giảm phân rã năng lượng trong toàn bộ khoảng bắt đầu tại thời gian  $t$ ,

$$\omega = 2\pi f.$$

Tương quan nêu trên mô tả sự giống nhau của sự suy yếu bao gồm trạng thái ban đầu và sự suy yếu bắt đầu tại thời gian  $t$ . Nó được tính toán từ băng thông rộng *EDR*, sử dụng toàn bộ khoảng tần số của *EDR* để tính toán, nhờ đó so sánh trường hợp năng lượng ban đầu đầy đủ với trường hợp tại thời gian  $t$ .

Sáng chế không giới hạn ở việc tính toán sự tương quan đối với tất cả các tần số. Tốt hơn là, sự tương quan cũng có thể được tính toán trong khoảng tần số định trước. Khoảng tần số này có thể được xác định từ tín hiệu âm thanh cần được xử lý. Ví dụ, đối với các tín hiệu âm thanh cụ thể, khoảng tần số có thể giới hạn ở khoảng định trước, ví dụ khoảng tần số có thể nghe thấy được. Theo các phương án, khoảng tần số này có thể nằm trong khoảng từ 20 Hz đến 20 kHz. Lưu ý rằng các khoảng khác cũng có thể được chọn, ví dụ bằng các nghiên cứu thực nghiệm.

Theo phương án, việc thực hiện dựa trên FFT hiệu quả của *EDR* có thể được sử dụng. Cửa sổ có độ dài thực của đáp ứng xung đo được áp dụng, và giả thiết rằng đáp ứng xung đo được có độ dài thực là  $2^{13}$  bằng với 8192 ngăn tần số. Trong khi tính toán, cửa sổ này được chuyển bởi độ dài rời rạc của một khối thời gian, và phần dưới của cửa sổ được đệm 0. Theo các phương án, sử dụng độ dài khối thời gian là 1ms, và để tính toán *EDR* đơn giản và hiệu quả thì phương pháp sau đây được sử dụng:

(1) Lấy tổng độ dài thực của đáp ứng xung đo được để tính quang phổ FFT, và các giá trị tuyệt đối được lấy bình phương  $E(l, \omega)$ .

(2) Cho đến khi kết thúc của đáp ứng xung được đạt, cửa sổ được di chuyển bởi độ dài khối thời gian rời rạc là 1ms về phía cuối của đáp ứng xung, các mẫu được tạo cửa sổ được tạo đệm 0 đối với độ dài thực (tức là, các mẫu ở xa độ dài thực được làm thành 0), và sau đó phổ FFT được tính toán mà thu được  $E(t, \omega)$ .

Phương pháp nêu trên là có lợi, khi không cần phải có giàn lọc bỏ sung hoặc loại tương tự để tính toán bằng hợp của *EDR*; thì chỉ cần chuyển cửa sổ. Fig.8 thể hiện ví dụ về sự giảm phân rã năng lượng đạt được cho đáp ứng xung theo phương pháp dựa trên FFT nêu trên.

Như được mô tả trên Fig.7 theo các bước 504 và 508, tương quan được xác định theo cách nêu trên sau đó sẽ được so sánh với ngưỡng định trước. Ngưỡng càng nhỏ, thì thời gian chuyển tiếp càng di chuyển hướng đến phần cuối của đáp ứng xung hơn. Ví dụ, đối với đáp ứng xung lập thể, nếu ngưỡng được chọn là  $1/e \approx 0,3679$  (xem tài liệu của kỹ thuật hiện nay [2]), thì sự chuyển tiếp là quá sớm tại các góc phương vị, bởi vì sự tương quan nằm trong khoảng dưới ngưỡng trước khi phản xạ thứ nhất xảy ra

hoặc va chạm. Tuy nhiên, do đã biết rằng thời gian chuyển tiếp phải muộn hơn thời gian tới của phản xạ thứ nhất, bởi vì phản xạ thứ nhất là rõ rệt và chắc chắn có thể không phải là vang âm khuếch tán cuối cùng, theo các phương án, nên ngưỡng không được xác định là ngưỡng cố định. Tốt hơn là, theo phương pháp của sáng chế, ngưỡng được định rõ sao cho nó phụ thuộc vào tương quan tại thời gian va chạm của phản xạ thứ nhất. Với sự định rõ này, giả thiết rằng phản xạ thứ nhất luôn luôn được định vị trước thời gian chuyển tiếp. Theo các phương án, thời gian chuyển tiếp, như được thể hiện ở bước 508, được xem xét là đạt đến khi áp dụng phương trình sau:

$$\rho(t) = c \cdot \rho(t_F)$$

trong đó,

$\rho(t_F)$  = số đo tương quan cho một trong số các phản xạ đầu được chọn,

$t_F$  = chỉ số thời gian trong đó một trong số các phản xạ đầu được chọn sau khi âm thanh trực tiếp va chạm,

$c$  = giá trị không đổi mà dựa trên  $\frac{1}{e}$ ,  $e$  là số Euler.

Tuy nhiên, theo các phương án, giá trị không đổi có thể là  $\frac{1}{e}$ , sáng chế không chỉ giới hạn ở giá trị này. Theo các phương án, giá trị không đổi có thể gần  $\frac{1}{e}$ , ví dụ bằng cách làm tròn hoặc cắt đi  $\frac{1}{e}$  theo vị trí thập phân định trước (xem dưới đây).

Theo phương án được mô tả,  $t_F$  là chỉ số thời gian trong đó phản xạ thứ nhất cuối cùng khi âm thanh trực tiếp va chạm.

Fig.9 thể hiện việc xác định thời gian chuyển tiếp theo phương pháp của sáng chế trong đó ngưỡng được tính toán phụ thuộc vào đáp ứng xung bằng cách nhân số tương quan tại điểm va chạm của phản xạ thứ nhất và giá trị cố định hoặc giá trị không đổi  $1/e$ . Biên độ của đáp ứng xung phòng 600 được thể hiện qua số lượng các mẫu, và phản xạ thứ nhất 602 cũng được xác định. Dạng sóng 604 biểu thị các giá trị tương quan thu được bằng phương trình (2). Tại 606 giá trị tương quan tại phản xạ thứ nhất được thể hiện, trong ví dụ đã mô tả, có giá trị là 0,58. Tương tự, ngưỡng cố định thường được sử dụng  $1/e$  được thể hiện tại 608. Giá trị tương quan 606 cho phản xạ

thứ nhất và giá trị cố định ban đầu  $1/e$  được sử dụng cho bộ nhân 610 mà tạo ra ngưỡng mới phụ thuộc vào giá trị tương quan tại phản xạ thứ nhất và, theo phương án được mô tả, có giá trị là 0,21 như được thể hiện tại 612. Do đó, khi so sánh với các phương pháp thông thường, thì điểm chuyển tiếp 614 được dịch chuyển hơn nữa về phía phải sao cho tất cả các mẫu sau điểm chuyển tiếp 614 đều được xem là âm vang cuối 304 và tất cả các mẫu trước đó được xem là phản xạ đầu 302. Có thể thấy rằng thời gian quyết định cuối cùng 614 mạnh hơn. Ví dụ, trong đáp ứng xung phòng lập thể điều này có nghĩa là thời gian chuyển tiếp được tính toán ổn định hơn nhiều tại các góc phương vị. Điều này có thể thấy được từ sự so sánh các Fig.10 và Fig.11. Fig.10 thể hiện thời gian chuyển tiếp khi sử dụng phương pháp được mô tả trong tài liệu của kỹ thuật hiện nay [1] đối với kênh bên trái 700 và kênh bên phải 702 cho đáp ứng xung phòng lập thể đo được bằng cách thực hiện EDC như nêu trên nhưng với ngưỡng cố định  $1/e$ . Sự phụ thuộc vào tai và góc phương vị là nhìn thấy rõ ràng cũng như quãng trễ sâu trong thời gian chuyển tiếp giảm xuống ít hơn 10ms là do giá trị tương quan  $\rho(t)$  nằm trong khoảng dưới ngưỡng trước khi phản xạ thứ nhất va chạm. Fig.11 thể hiện thời gian chuyển tiếp đối với kênh bên trái 700 và kênh bên phải 702 khi được tính toán theo phương pháp của sáng chế. Có thể thấy rằng thời gian chuyển tiếp thu được ít phụ thuộc vào tai và góc phương vị hơn khi so với phương pháp thông thường được mô tả theo Fig.10.

Theo các phương án, thời gian chuyển tiếp được xem là đã đạt được khi giá trị tương quan nằm trong khoảng dưới hoặc bằng giá trị ngưỡng cho lần thứ nhất và không gia tăng lần nữa đối với ngưỡng sau đó. Giá trị thời gian mà kết hợp với mẫu này theo chức năng tương quan được tính toán là thời gian trong đó âm vang cuối của đáp ứng xung được coi là bắt đầu. Theo phương pháp của sáng chế, thời gian va chạm của phản xạ thứ nhất có thể được xác định bằng cách thực hiện toán tử độ nhọn, như được mô tả trong tài liệu của kỹ thuật hiện nay [6]. Theo cách khác, phản xạ thứ nhất có thể đó thấy bằng các phương pháp khác, ví dụ, bằng cách dò ngưỡng hoặc dò va chạm, ví dụ, được mô tả trong tài liệu của kỹ thuật hiện nay [7].

Theo các phương án,  $e^{-1} = 0,3679$  được sử dụng làm giá trị để biểu thị tương quan thấp theo các quy trình ngẫu nhiên, ví dụ, như được đề cập đến trong tài liệu của kỹ thuật hiện nay [1]. Theo các phương án, giá trị này được sử dụng với 4 chữ số thập

phân sao cho  $e^{-1}$  gần bằng 0,3679. Theo các phương án khác, nhiều hoặc ít số thập phân hơn cũng có thể được sử dụng và người ta quan sát thấy rằng thời gian chuyển tiếp dò thấy thay đổi theo độ lệch so với số chính xác  $e^{-1}$ . Ví dụ, khi sử dụng giá trị 0,368 thì sẽ tạo ra thay đổi rất nhỏ trong thời gian chuyển tiếp dưới 1ms.

Theo các phương án khác, đáp ứng xung có thể được băng giới hạn, và trong trường hợp này, *EDR* có thể được tính toán trong khoảng tần số giới hạn và số đo tương quan cũng có thể được tính toán trong khoảng tần số giới hạn của *EDR*. Các biến đổi tần số giàn lọc thay thế cũng có thể được sử dụng, ví dụ, để đạt đến sự vận hành hoàn toàn trong miền FFT, nhờ đó tránh các biến đổi bổ sung, ví dụ khi sử dụng FFT dựa trên phép nhân chập/lọc.

Lưu ý rằng trong phần mô tả các phương án nêu trên, tham chiếu tới giá trị của giá trị tương quan cho phản xạ thứ nhất. Tuy nhiên, các phương án khác có thể sử dụng giá trị tương quan được tính toán cho một trong các phản xạ đầu khác.

Như được đề cập đến ở trên, phương pháp theo sáng chế, theo các phương án có thể được sử dụng trong bộ xử lý lập thể để xử lý lập thể các tín hiệu âm thanh. Sau đây, phương án xử lý lập thể các tín hiệu âm thanh sẽ được mô tả. Sự xử lý lập thể có thể tiến hành như quy trình giải mã chuyển đổi tín hiệu được giải mã thành tín hiệu trộn giảm lập thể mà cung cấp trải nghiệm âm thanh bao quanh khi được nghe qua ống nghe.

Fig.12 thể hiện sự biểu diễn dưới dạng biểu đồ bộ kết xuất lập thể 800 để xử lý lập thể tín hiệu âm thanh theo phương án của sáng chế. Fig.12 cũng thể hiện cái nhìn tổng quan về xử lý miền QMF trong bộ kết xuất lập thể. Tại đầu vào 802, bộ kết xuất lập thể 800 nhận tín hiệu âm thanh để xử lý, ví dụ tín hiệu đầu vào bao gồm các kênh N và 64 băng QMF. Ngoài ra, bộ kết xuất lập thể 800 nhận nhiều tham số đầu vào để điều chỉnh việc xử lý tín hiệu âm thanh. Các tham số đầu vào bao gồm đáp ứng xung phòng lập thể (BRIR) 804 cho các kênh  $2 \times N$  và 64 băng QMF, số chỉ dẫn  $K_{\max}$  806 của băng cực đại được sử dụng để nhân chập của tín hiệu đầu vào âm thanh có phần phản xạ đầu của BRIR 804, và tham số phần vang âm 808 và 810 nêu trên (RT60 và năng lượng vang âm). Bộ kết xuất lập thể 800 bao gồm bộ xử lý nhân chập nhanh 812 để xử lý tín hiệu âm thanh đầu vào 802 với phần đầu của BRIR được nhận 804. Tại

đầu ra, bộ xử lý 812 tạo ra tín hiệu được xử lý trước 814 bao gồm hai kênh và các băng  $K_{\max}$  QMF. Bộ kết xuất lập thể 800 bao gồm, ngoài nhánh xử lý sớm có bộ xử lý nhân chập nhanh 812, nhánh vang âm bao gồm hai phần vang âm 816a và 816b trong đó mỗi phần nhận thông tin RT60 808 và thông tin năng lượng vang âm 810 làm tham số đầu vào. Nhánh vang âm còn bao gồm bộ xử lý trộn giảm âm lập thể 818 và bộ xử lý phân tích tương quan 820 đều nhận tín hiệu âm thanh đầu vào 802. Ngoài ra, hai giai đoạn tăng 821a và 821b được tạo ra giữa bộ xử lý trộn giảm âm lập thể 818 và phần vang âm 816a và 816b tương ứng để điều khiển sự gia tăng của tín hiệu trộn giảm 822 được tạo ra bởi bộ xử lý trộn giảm âm lập thể 818. Bộ xử lý trộn giảm âm lập thể 818 dựa trên tín hiệu đầu vào 802 tạo ra tín hiệu trộn giảm 822 có hai băng và 64 băng QMF. Sự gia tăng của hai thời kỳ 821a và 821b được điều khiển bởi các tín hiệu điều khiển tương ứng 824a và 824b được tạo ra bởi bộ xử lý phân tích tương quan 820. Tín hiệu được trộn giảm được điều khiển được nhập vào phần vang âm 816a và 816b tương ứng tạo ra các tín hiệu vang âm 826a, 826b tương ứng. Tín hiệu được xử lý sớm 814 và các tín hiệu vang âm 826a, 826b được nhận bởi bộ trộn 828 mà tổ hợp các tín hiệu được nhận vào tín hiệu âm thanh đầu ra 830 có hai kênh và 64 băng QMF. Ngoài ra, theo sáng chế bộ xử lý nhân chập 812 và phần vang âm 816a và 816b nhận tham số đầu vào bổ sung 832 chỉ sự chuyển tiếp trong đáp ứng xung phòng 804 từ phần đầu sang âm vang cuối được xác định như trên.

Mô-đun bộ kết xuất lập thể 800 (ví dụ bộ kết xuất lập thể 236 trên Fig.2 hoặc Fig.4) có dòng dữ liệu được giải mã làm đầu vào 802. Tín hiệu được xử lý bởi giàn lọc phân tích QMF như được thể hiện theo ISO/IEC 14496-3:2009, điều khoản 4.B.18.2 với phần sửa đổi được nêu tại ISO/IEC 14496-3:2009, điều khoản 8.6.4.2. Mô-đun bộ kết xuất 800 cũng có thể xử lý dữ liệu đầu vào miền QMF; trong trường hợp này giàn lọc phân tích có thể được loại bỏ. Đáp ứng xung phòng lập thể (BRIR) 804 được thể hiện là các bộ lọc miền QMF phức. Sự chuyển đổi từ đáp ứng xung phòng lập thể miền thời gian sang sự biểu diễn bộ lọc QMF phức được thể hiện theo ISO/IEC FDIS 23003-1:2006, phụ lục B. BRIR 804 giới hạn ở số lượng khe thời gian nhất định trong miền QMF phức, sao cho chúng chỉ chứa phần phản xạ đầu 301, 302 (xem Fig.5) và không bao gồm vang âm khuếch tán cuối cùng 304. Điểm chuyển tiếp 832 từ các phản xạ đầu sang âm vang cuối được xác định như trên, ví dụ bằng phân tích BRIR 804 ở

bước xử lý trước của quy trình xử lý lập thể. Tín hiệu âm thanh miền QMF 802 và BRIR miền QMF 804 sau đó được xử lý bằng cách nhân chập nhanh theo băng 812 để thực hiện xử lý lập thể. Phần vang âm miền QMF 816a, 816b được sử dụng để tạo ra âm vang cuối miền QMF 2 kênh 826a, 826b. Mô-đun vang âm 816a, 816b sử dụng tập hợp các lần vang âm phụ thuộc vào tần số 808 và giá trị năng lượng 810 để đáp ứng các đặc tính của vang âm. Dạng sóng của vang âm dựa trên trộn giảm âm thanh nổi 818 của tín hiệu đầu vào âm thanh 802 và được chia tỉ lệ thích hợp 821a, 821b theo biên độ phụ thuộc vào phân tích tương quan 820 của tín hiệu âm thanh đa kênh 802. Kết quả nhân chập miền QMF 2-kênh 814 và vang âm miền QMF 2 kênh 816a, 816b sau đó được tổ hợp 828 và cuối cùng, hai giàn lọc tổng hợp QMF tính toán các tín hiệu đầu ra miền thời gian lập thể 830 như được xác định theo ISO/IEC 14496-3:2009, điều khoản 4.6.18.4.2. Bộ kết xuất cũng có thể tạo ra dữ liệu đầu ra miền QMF; giàn lọc tổng hợp sau đó được loại bỏ.

#### Các định nghĩa

Tín hiệu âm thanh 802 mà được cấp vào mô-đun bộ kết xuất lập thể 800 được đề cập đến là các *tín hiệu đầu vào* như dưới đây. Các tín hiệu âm thanh 830 là kết quả của việc xử lý lập thể được đề cập đến là *các tín hiệu đầu ra*. Các tín hiệu đầu vào 802 của mô-đun bộ kết xuất lập thể 800 là tín hiệu đầu ra âm thanh của bộ giải mã chính (ví dụ, xem tín hiệu 228 trên Fig.2). Các định nghĩa sau đây được sử dụng:

|           |                                                                                                                                             |
|-----------|---------------------------------------------------------------------------------------------------------------------------------------------|
| $N_{in}$  | Số lượng kênh đầu vào                                                                                                                       |
| $N_{out}$ | Số lượng kênh đầu ra, $N_{out} = 2$                                                                                                         |
| $M_{DMX}$ | Ma trận trộn giảm chứa các hệ số trộn giảm không âm giá trị thực (tăng trộn độ khuếch đại). $M_{DMX}$ là kích thước $N_{out} \times N_{in}$ |
| $L$       | Độ dài khung được đo trong các mẫu âm thanh miền thời gian.                                                                                 |
| $\nu$     | Chỉ số mẫu miền thời gian                                                                                                                   |
| $n$       | Chỉ số khe thời gian QMF (giá trị mẫu phân giải)                                                                                            |
| $L_n$     | Độ dài khung đo được trong các khe thời gian QMF                                                                                            |
| $F$       | Chỉ số khung (số khung)                                                                                                                     |
| $K$       | Số lượng băng tần QMF, $K = 64$                                                                                                             |

|                                       |                                                                                                         |
|---------------------------------------|---------------------------------------------------------------------------------------------------------|
| $k$                                   | Chỉ số băng QMF (1..64)                                                                                 |
| $A, B, ch$                            | Các chỉ số kênh (số lượng kênh của cấu hình kênh)                                                       |
| $L_{trans}$                           | Độ dài của phần phản xạ đầu của BRIR trong các mẫu miền thời gian                                       |
| $L_{trans,n}$                         | Độ dài của phần phản xạ đầu của BRIR trong các khe thời gian QMF                                        |
| $N_{BRIR}$                            | Số lượng cặp BRIR trong tập hợp dữ liệu BRIR                                                            |
| $L_{FFT}$                             | Độ dài của biến đổi FFT                                                                                 |
| $\Re(\cdot)$                          | Phần thực của tín hiệu có giá trị phức tạp                                                              |
| $\Im(\cdot)$                          | Phần ảo của tín hiệu có giá trị phức tạp                                                                |
| $\mathbf{m}_{conv}$                   | Vectơ phát tín hiệu mà chứa kênh tín hiệu đầu vào mà cặp BRIR trong tập hợp dữ liệu BRIR                |
| $f_{max}$                             | Tần số cực đại được sử dụng để xử lý lập thể                                                            |
| $f_{max,decoder}$                     | Tần số tín hiệu cực đại mà có trong tín hiệu đầu ra âm thanh của bộ giải mã                             |
| $K_{max}$                             | Bảng thông cực đại mà được sử dụng để nhân chập tín hiệu đầu vào âm thanh với phần phản xạ đầu của BRIR |
| $a$                                   | Hệ số ma trận trộn giảm                                                                                 |
| $c_{eq,k}$                            | Hệ số cân bằng năng lượng theo băng                                                                     |
| $\varepsilon$                         | Hằng số, $\varepsilon = 10^{-20}$                                                                       |
| $d$                                   | Độ trễ trong các khe thời gian miền QMF                                                                 |
| $\tilde{\mathbf{y}}_{ch}^{n',k}$      | Biểu diễn tín hiệu miền FFT giả trong băng tần $k$                                                      |
| $n'$                                  | Chỉ số tần số FFT giả                                                                                   |
| $\tilde{\mathbf{h}}^{n',k}$           | Biểu diễn miền FFT giả của BRIR trong băng tần $k$                                                      |
| $\tilde{\mathbf{z}}_{ch,conv}^{n',k}$ | Kết quả nhân chập miền FFT giả trong băng tần $k$                                                       |
| $\hat{\mathbf{z}}_{ch,conv}^{n,k}$    | Tín hiệu trung gian: kết quả nhân chập 2-kênh trong miền QMF                                            |
| $\hat{\mathbf{z}}_{ch,rev}^{n,k}$     | Tín hiệu trung gian: vang âm 2-kênh trong miền QMF                                                      |

|                              |                                                                                                                |
|------------------------------|----------------------------------------------------------------------------------------------------------------|
| $K_{ana}$                    | Số lượng các băng tần phân tích (được sử dụng cho phần vang âm)                                                |
| $f_{c,ana}$                  | Tần số giữa của các băng tần phân tích                                                                         |
| $N_{DMX,act}$                | Số lượng các kênh mà được trộn giảm thành một kênh trộn giảm âm lập thể và hoạt động trong khung tín hiệu thực |
| $c_{corr}$                   | Tổng hệ số tương quan cho một khung tín hiệu                                                                   |
| $c_{corr}^{A,B}$             | Hệ số tương quan để tổ hợp các kênh $A, B$                                                                     |
| $\sigma_{\hat{y}_{ch,A}^n}$  | Độ lệch chuẩn cho khe thời gian $n$ của tín hiệu $\hat{y}_{ch,A}^n$                                            |
| $\mathbf{c}_{scale}$         | Vectơ của hai bội suất                                                                                         |
| $\tilde{\mathbf{c}}_{scale}$ | Vectơ của hai bội suất, đều đặn qua thời điểm nhất định                                                        |

## Xử lý

Quy trình xử lý tín hiệu đầu vào được mô tả dưới đây. Mô-đun bộ kết xuất lập thể hoạt động trên các khung kế nhau, không chồng lấp các mẫu miền thời gian của các tín hiệu âm thanh đầu vào và xuất ra một khung của  $L$  mẫu cho mỗi khung đầu vào được xử lý có độ dài  $L$ .

### (1) Khởi tạo và xử lý trước

Tiến hành khởi tạo khối xử lý lập thể trước khi xử lý các mẫu âm thanh được phân phối bởi bộ giải mã chính (ví dụ, xem bộ giải mã 200 trên Fig.2). Quá trình khởi tạo bao gồm các bước xử lý sau.

#### (a) Đọc các giá trị phân tích

Mô-đun phân vang âm 816a, 816b lấy tập hợp các lần vang âm 808 phụ thuộc vào tần số và các giá trị năng lượng 810 làm các tham số đầu vào. Các giá trị này được đọc từ giao diện tại thời điểm khởi tạo mô-đun xử lý lập thể 800. Ngoài ra, thời gian chuyển tiếp 832 từ các phản xạ đầu sang âm vang cuối trong các mẫu miền thời gian cũng được đọc. Các giá trị có thể được lưu trữ trong tệp tin nhị phân được viết với 32 bit trên mỗi mẫu, giá trị thực, trình tự dạng little-endian. Các giá trị được đọc cần thiết để xử lý được thể hiện ở bảng dưới đây:

| Mô tả giá trị                                                                                               | Số lượng         | Loại dữ liệu |
|-------------------------------------------------------------------------------------------------------------|------------------|--------------|
| Độ dài chuyển tiếp $L_{\text{trans}}$                                                                       | 1                | Số nguyên    |
| Số lượng băng tần $K_{\text{ana}}$                                                                          | 1                | Số nguyên    |
| Tần số giữa $f_{c,\text{ana}}$ của băng tần                                                                 | $K_{\text{ana}}$ | Số thực      |
| Số lần vang âm RT60 theo giây                                                                               | $K_{\text{ana}}$ | Số thực      |
| Giá trị năng lượng mà biểu thị năng lượng (biên độ với công suất là hai) của phần âm vang cuối của một BRIR | $K_{\text{ana}}$ | Số thực      |

(b) *Đọc và xử lý trước BRIR*

Đáp ứng xung phòng lập thể 804 được đọc từ hai tệp tin chuyên dụng mà lưu trữ BRIR riêng cho tai trái và tai phải. Các mẫu miền thời gian của BRIR được lưu trữ trong các tệp tin dạng sóng số nguyên có độ phân giải là 24 bit đối với mỗi mẫu và 32 kênh. Trật tự BRIR trong tệp tin được thể hiện trong bảng dưới đây:

| Số lượng kênh | Loa       |
|---------------|-----------|
| 1             | CH_M_L045 |
| 2             | CH_M_R045 |
| 3             | CH_M_000  |
| 4             | CH_LFE1   |
| 5             | CH_M_L135 |
| 6             | CH_M_R135 |
| 7             | CH_M_L030 |
| 8             | CH_M_R030 |
| 9             | CH_M_180  |
| 10            | CH_LFE2   |
| 11            | CH_M_L090 |
| 12            | CH_M_R090 |
| 13            | CH_U_L045 |

|    |           |
|----|-----------|
| 14 | CH_U_R045 |
| 15 | CH_U_000  |
| 16 | CH_T_000  |
| 17 | CH_U_L135 |
| 18 | CH_U_R135 |
| 19 | CH_U_L090 |
| 20 | CH_U_R090 |
| 21 | CH_U_180  |
| 22 | CH_L_000  |
| 23 | CH_L_L045 |
| 24 | CH_L_R045 |
| 25 | CH_M_L060 |
| 26 | CH_M_R060 |
| 27 | CH_M_L110 |
| 28 | CH_M_R110 |
| 29 | CH_U_L030 |
| 30 | CH_U_R030 |
| 31 | CH_U_L110 |
| 32 | CH_U_R110 |

Nếu không có BRIR đo được tại một trong số các vị trí loa phát thanh, thì kênh tương ứng trong tệp tin sóng chứa các giá trị 0. Các kênh LFE không được sử dụng để xử lý lập thể.

Để làm bước xử lý trước, tập hợp đáp ứng xung phòng lập thể (BRIR) được biến đổi từ bộ lọc miền thời gian sang bộ lọc miền QMF có giá trị phức. Việc thiết lập bộ lọc miền thời gian trong miền QMF có giá trị phức được thực hiện theo ISO/IEC FDIS 23003-1:2006, phụ lục B. Hệ số bộ lọc mẫu để chuyển đổi bộ lọc được sử dụng theo ISO/IEC FDIS 23003-1:2006, phụ lục B, bảng B.1. Biểu diễn miền thời gian  $\tilde{\mathbf{h}}_{\text{ch}}^v = [\tilde{h}_1^v \dots \tilde{h}_{N_{\text{BRIR}}}^v]$  trong đó  $1 \leq v \leq L_{\text{trans}}$  được xử lý để thu được bộ lọc miền QMF có giá trị phức  $\hat{\mathbf{h}}_{\text{ch}}^{n,k} = [\hat{h}_1^{n,k} \dots \hat{h}_{N_{\text{BRIR}}}^{n,k}]$  trong đó  $1 \leq n \leq L_{\text{trans},n}$ .

## (2) Xử lý tín hiệu âm thanh

Khối xử lý âm thanh của mô-đun bộ kết xuất lập thể 800 thu được các mẫu âm thanh miền thời gian 802 cho các kênh đầu vào  $N_{in}$  từ bộ giải mã chính và tạo ra tín hiệu đầu ra lập thể 830 có  $N_{out} = 2$  kênh.

Quy trình xử lý lấy làm đầu vào

- dữ liệu âm thanh được giải mã 802 từ bộ giải mã lõi,
- biểu diễn miền QMF phức của phản phản xạ đầu của tập hợp BRIR 804, và
- tập hợp tham số phụ thuộc vào tần số 808, 810, 832 mà được sử dụng bởi phân vang âm miền QMF 816a, 816b để tạo ra âm vang cuối 826a, 826b.

## (a) phân tích QMF tín hiệu âm thanh

Như bước xử lý thứ nhất, mô-đun bộ kết xuất lập thể biến đổi các mẫu miền thời gian  $L = 2048$  của tín hiệu đầu vào miền thời gian  $N_{in}$ -kênh (từ bộ giải mã chính)  $[\tilde{y}_{ch,1}^v \cdots \tilde{y}_{ch,N_{in}}^v] = \tilde{y}_{ch}^v$  thành biểu diễn tín hiệu miền QMF  $N_{in}$ -kênh 802 có kích thước khe thời gian QMF  $L_n = 32$  (chỉ số khe  $n$ ) và  $K = 64$  băng tần (chỉ số băng  $k$ ).

Phân tích QMF như được mô tả trong ISO/IEC 14496-3:2009, điều khoản 4.B.18.2 được sửa đổi theo ISO/IEC 14496-3:2009, điều khoản 8.6.4.2. được thực hiện trên khung tín hiệu miền thời gian  $\tilde{y}_{ch}^v$  để thu được khung tín hiệu miền QMF  $[\hat{y}_{ch,1}^{n,k} \cdots \hat{y}_{ch,N_{in}}^{n,k}] = \hat{y}_{ch}^{n,k}$  trong đó  $1 \leq v \leq L$  và  $1 \leq n \leq L_n$ .

## (b) Nhân chấp nhanh tín hiệu âm thanh miền QMF và BRIR miền QMF

Tiếp theo, nhân chấp nhanh theo băng 812 được thực hiện để xử lý tín hiệu âm thanh miền QMF 802 và BRIR miền QMF 804. Phân tích FFT có thể được thực hiện cho mỗi băng tần QMF  $k$  đối với mỗi kênh tín hiệu đầu vào 802 và mỗi BRIR 804.

Do các giá trị phức trong miền QMF nên phân tích FFT được tiến hành trên phần thực tế của biểu diễn tín hiệu miền QMF và phân tích FFT trên các phần ảo của

phép biểu diễn tín hiệu miền QMF. Sau đó, các kết quả được tổ hợp để hình thành tín hiệu miền giả FFT giá trị phức theo bảng cuối

$$\tilde{\mathbf{y}}_{\text{ch}}^{n',k} = \text{FFT}(\hat{\mathbf{y}}_{\text{ch}}^{n',k}) = \text{FFT}(\Re(\hat{\mathbf{y}}_{\text{ch}}^{n',k})) + j \cdot \text{FFT}(\Im(\hat{\mathbf{y}}_{\text{ch}}^{n',k}))$$

và BRIR giá trị phức theo bảng

$$\tilde{\mathbf{h}}_1^{n',k} = \text{FFT}(\hat{\mathbf{h}}_1^{n',k}) = \text{FFT}(\Re(\hat{\mathbf{h}}_1^{n',k})) + j \cdot \text{FFT}(\Im(\hat{\mathbf{h}}_1^{n',k})) \text{ cho tai trái}$$

$$\tilde{\mathbf{h}}_2^{n',k} = \text{FFT}(\hat{\mathbf{h}}_2^{n',k}) = \text{FFT}(\Re(\hat{\mathbf{h}}_2^{n',k})) + j \cdot \text{FFT}(\Im(\hat{\mathbf{h}}_2^{n',k})) \text{ cho tai phải.}$$

Độ dài của biến đổi FFT được xác định theo độ dài của bộ lọc BRIR miền QMF có giá trị phức  $L_{\text{trans},n}$  và độ dài khung trong các khe thời gian miền QMF  $L_n$  sao cho

$$L_{\text{FFT}} = L_{\text{trans},n} + L_n - 1.$$

Tín hiệu miền FFT giả giá trị phức sau đó được nhân với các bộ lọc BRIR miền FFT giá trị phức để thu được các kết quả nhân chập nhanh. Vector  $\mathbf{m}_{\text{conv}}$  được sử dụng để báo hiệu kênh tín hiệu đầu vào tương ứng với cặp BRIR trong tập hợp dữ liệu BRIR.

Phép nhân này được thực hiện theo bảng cho tất cả các băng tần QMF  $k$  với  $1 \leq k \leq K_{\text{max}}$ . Bảng cực đại  $K_{\text{max}}$  được xác định bởi băng QMF thể hiện cả tần số 18 kHz hoặc tần số tín hiệu cực đại mà được thể hiện trong có trong tín hiệu âm thanh từ bộ giải mã chính

$$f_{\text{max}} = \min(f_{\text{max,decoder}}, 18\text{kHz}).$$

Phép nhân này thu được từ mỗi kênh đầu vào âm thanh với mỗi cặp BRIR được tính tổng trong mỗi băng tần QMF  $k$  trong đó  $1 \leq k \leq K_{\text{max}}$  tạo ra tín hiệu miền FFT giả  $K_{\text{max}}$ -băng 2-kênh trung gian.

$$\tilde{\mathbf{z}}_{\text{ch},1,\text{conv}}^{n',k} = \sum_{ch=1}^{ch=N_{\text{in}}} \tilde{\mathbf{y}}_{\text{ch},ch}^{n',k} \cdot \tilde{\mathbf{h}}_{1,\mathbf{m}_{\text{conv}}[ch]}^{n',k} \text{ và } \tilde{\mathbf{z}}_{\text{ch},2,\text{conv}}^{n',k} = \sum_{ch=1}^{ch=N_{\text{in}}} \tilde{\mathbf{y}}_{\text{ch},ch}^{n',k} \cdot \tilde{\mathbf{h}}_{2,\mathbf{m}_{\text{conv}}[ch]}^{n',k} \text{ là kết quả nhân chập}$$

FFT giả  $\tilde{\mathbf{z}}_{\text{ch},\text{conv}}^{n',k} = [\tilde{\mathbf{z}}_{\text{ch},1,\text{conv}}^{n',k}, \tilde{\mathbf{z}}_{\text{ch},2,\text{conv}}^{n',k}]$  trong băng tần miền QMF  $k$ .

Tiếp theo, tổng hợp FFT theo băng được thực hiện để biến đổi kết quả nhân chập ngược trở lại miền QMF tạo ra tín hiệu miền QMF  $K_{\max}$ -băng 2-kênh trung gian trong đó khe thời gian  $L_{\text{FFT}}$   $\hat{\mathbf{z}}_{\text{ch,conv}}^{n,k} = [\hat{z}_{\text{ch,1,conv}}^{n,k}, \hat{z}_{\text{ch,2,conv}}^{n,k}]$  trong đó  $1 \leq n \leq L_{\text{FFT}}$  và  $1 \leq k \leq K_{\max}$ .

Đối với mỗi khung tín hiệu đầu vào miền QMF với khe thời gian  $L=32$ , khung tín hiệu có kết quả nhân chập với các khe thời gian  $L=32$  được quay lại. Khe thời gian  $L_{\text{FFT}}-32$  còn lại được lưu trữ và xử lý bổ sung-chồng lắp được thực hiện trong (các) khung dưới đây.

(c) *Tạo ra âm vang cuối*

Để làm tín hiệu trung gian thứ hai 826a, 826b, tín hiệu vang âm được gọi là  $\hat{\mathbf{z}}_{\text{ch,rev}}^{n,k} = [\hat{z}_{\text{ch,1,rev}}^{n,k}, \hat{z}_{\text{ch,2,rev}}^{n,k}]$  được tạo ra bởi mô-đun bộ vang miền tần số 816a, 816b. Bộ vang miền tần số 816a, 816b lấy làm đầu vào.

- trộn giảm âm lập thể miền QMF 822 của một khung tín hiệu đầu vào,
- tập hợp tham số chứa các lần vang âm phụ thuộc vào tần số 808 và giá trị năng lượng 810.

Bộ vang miền tần số 816a, 816b quay lại phần đuôi vang âm miền QMF 2-kênh.

Số lượng băng được sử dụng tối đa của tập hợp tham số phụ thuộc vào tần số được tính toán phụ thuộc vào tần số cực đại.

Trước tiên, trộn giảm miền QMF 818 của một khung tín hiệu đầu vào  $\hat{\mathbf{y}}_{\text{ch}}^{n,k}$  được thực hiện để tạo ra đầu vào của phần vang âm bằng cách tính tổng trọng số của các kênh tín hiệu đầu vào. Trọng số thu được được chứa trong ma trận trộn giảm  $\mathbf{M}_{\text{DMX}}$ . Chúng là các giá trị thật và không âm và ma trận trộn giảm có kích thước  $N_{\text{out}} \times N_{\text{in}}$ . Nó chứa giá trị khác 0 trong đó kênh tín hiệu đầu vào được ánh đến một trong hai kênh đầu ra.

Các kênh mà thể hiện các vị trí loa phát thanh trên bán cầu trái được ánh xạ đến kênh đầu ra bên trái và các kênh mà thể hiện loa phát thanh được định vị trên bán cầu

phải được ánh xạ đến kênh đầu ra bên phải. Các tín hiệu của các kênh này được đo bằng hệ số 1. Các kênh mà thể hiện loa phát thanh ở mặt phẳng giữa được ánh xạ đến cả hai kênh đầu ra của tín hiệu lập thể. Tín hiệu đầu vào của các kênh này được đo bằng hệ số

$$a = 0.7071 \approx \frac{1}{\sqrt{2}}.$$

Ngoài ra, bước cân bằng năng lượng được thực hiện trong việc trộn giảm. Nó đáp ứng năng lượng theo băng của một kênh trộn giảm sao cho bằng với tổng năng lượng theo băng của các kênh tín hiệu đầu vào mà có trong kênh trộn giảm này. Bước cân bằng năng lượng này được thực hiện bằng phép nhân theo băng với hệ số có giá trị thật

$$c_{eq,k} = \sqrt{\frac{P_{in}^k}{P_{out}^k + \varepsilon}}.$$

Hằng số  $c_{eq,k}$  được giới hạn ở khoảng  $[0,5, 2]$ . Hằng số  $\varepsilon$  được sử dụng để tránh phép chia bằng 0. Trộn giảm cũng được giới hạn bằng đối với tần số  $f_{max}$ ; các giá trị trong tất cả các băng tần cao hơn đều được thiết lập thành 0.

Fig.13 là biểu đồ thể hiện việc xử lý trong bộ vang miền tần số 816a, 816b của bộ kết xuất lập thể 800 theo phương án của sáng chế.

Trong bộ vang miền tần số trộn giảm đơn của đầu vào âm lập thể được tính toán bằng cách sử dụng bộ trộn đầu vào 900. Việc này được thực hiện rồi rạc sử dụng chuyển pha  $90^\circ$  trên kênh đầu vào thứ hai.

Tín hiệu đơn này sau đó được nạp đến vòng lặp độ trễ phản hồi 902 trong mỗi băng tần  $k$ , mà tạo ra chuỗi xung suy yếu. tiếp theo là bộ giải tương quan FIR song song mà phân bố năng lượng tín hiệu theo cách suy yếu dần vào các khoảng giữa các xung và tạo ra sự rời rạc giữa các kênh đầu ra. Mật độ va đập bộ lọc suy yếu được sử dụng để tạo ra sự phân rã năng lượng. Các sự vận hành pha đập bộ lọc được hạn chế với 4 lựa chọn để thực hiện quá trình giải tương quan rời rạc và không có giá trị nhân.

Sau khi tính toán vang âm, hiệu chỉnh sự rời rạc giữa các kênh (ICC - inter-channel coherence) 904 bao gồm trong mô-đun bộ vang đối với mỗi băng tần QMF. Ở

bước hiệu chỉnh ICC phần thẳng dư trực tiếp phụ thuộc vào tần số  $g_{\text{direct}}$  và phần thẳng dư sau khi trộn  $g_{\text{cross}}$  được sử dụng để đáp ứng ICC.

Lượng năng lượng và số lần vang âm đối với các băng tần khác được chứa trong tập hợp tham số đầu vào. Các giá trị được xác định ra tại nhiều điểm tần số mà được ánh xạ bên trong thành băng tần QMF  $K = 64$ .

Hai trường hợp bộ vang miền tần số được sử dụng để tính toán tín hiệu trung gian cuối cùng  $\hat{\mathbf{z}}_{\text{ch,rev}}^{n,k} = [\hat{z}_{\text{ch,1,rev}}^{n,k}, \hat{z}_{\text{ch,2,rev}}^{n,k}]$ . Tín hiệu  $\hat{z}_{\text{ch,1,rev}}^{n,k}$  là kênh đầu ra thứ nhất của trường hợp thứ nhất của phần vang âm, và  $\hat{z}_{\text{ch,2,rev}}^{n,k}$  là kênh đầu ra thứ hai của trường hợp thứ hai của phần vang âm. Chúng được tổ hợp thành khung tín hiệu vang âm cuối cùng có kích thước 2 kênh, 64 băng và 32 khe thời gian.

Trộn giảm âm lập thể 822 là được chia tỉ lệ hai lần 821a,b theo số đo tương quan 820 của khung tín hiệu đầu vào để đảm bảo chia tỉ lệ đúng của phần đầu ra của phần vang âm. Hệ số tỉ lệ được xác định là giá trị trong khoảng  $[\sqrt{N_{\text{DMX,act}}}, N_{\text{DMX,act}}]$  phụ thuộc vào hệ số tương quan  $c_{\text{corr}}$  giữa 0 và 1 trong đó

$$c_{\text{corr}} = \frac{1}{N_{\text{in}}^2} \cdot \sum_{A=1}^{A=N_{\text{DMX,act}}} \sum_{B=1}^{B=N_{\text{DMX,act}}} c_{\text{corr}}^{A,B} \text{ và}$$

$$c_{\text{corr}}^{A,B} = \left| \frac{1}{K-1} \cdot \frac{\sum_k \sum_n \hat{y}_{\text{ch,A}}^{n,k} \cdot \hat{y}_{\text{ch,B}}^{n,k*}}{\sum_n \sigma_{\hat{y}_{\text{ch,A}}^n} \cdot \sigma_{\hat{y}_{\text{ch,B}}^n}} \right|$$

trong đó,  $\sigma_{\hat{y}_{\text{ch,A}}^n}$  nghĩa là độ lệch chuẩn trong một khe thời gian  $n$  của kênh  $A$ , toán tử  $\{*\}$  biểu thị liên hợp phức tạp và  $\hat{y}$  là phiên bản trung bình 0 của tín hiệu miền QMF  $\hat{y}$  trong khung tín hiệu thực tế.

$c_{\text{corr}}$  được tính toán hai lần: một lần cho tất cả các kênh  $A, B$  mà hoạt động tại khung tín hiệu thực  $F$  và được chứa trong kênh bên trái của trộn giảm âm lập thể và một lần cho tất cả các kênh  $A, B$  mà hoạt động tại khung tín hiệu thực  $F$  và được chứa trong kênh bên phải của trộn giảm âm lập thể.

$N_{\text{DMX,act}}$  là số lượng kênh đầu vào mà được trộn giảm thành một kênh trộn giảm  $A$  (số lượng yếu tố ma trận trong hàng  $A$ th của ma trận trộn giảm  $\mathbf{M}_{\text{DMX}}$  mà không bằng 0) và hoạt động trong khung hiện tại.

Hệ số tỉ lệ là

$$\begin{aligned} \mathbf{c}_{\text{scale}} &= [c_{\text{scale},1}, c_{\text{scale},2}] \\ &= \left[ \sqrt{N_{\text{DMX,act},1}} + c_{\text{corr}} \cdot (N_{\text{DMX,act},1} - \sqrt{N_{\text{DMX,act},1}}), \sqrt{N_{\text{DMX,act},2}} + c_{\text{corr}} \cdot (N_{\text{DMX,act},2} - \sqrt{N_{\text{DMX,act},2}}) \right] \end{aligned}$$

Thừa số định tỉ lệ được làm nhẵn qua khung tín hiệu âm thanh bằng bộ lọc lấy hạ tần cũ hơn thứ nhất thu được các hệ số tỉ lệ được làm nhẵn  $\tilde{\mathbf{c}}_{\text{scale}} = [\tilde{c}_{\text{scale},1}, \tilde{c}_{\text{scale},2}]$ .

Thừa số định tỉ lệ được khởi tạo trong khung dữ liệu đầu vào âm thanh thứ nhất bằng cách phân tích tương quan miền –thời gian với các giá trị trung bình giống nhau.

Đầu vào trong trường hợp của phản vang âm được chia tỉ lệ với hệ số tỉ lệ  $\tilde{c}_{\text{scale},1}$  và đầu vào trong trường hợp của phản vang âm thứ hai được chia tỉ lệ với hệ số tỉ lệ  $\tilde{c}_{\text{scale},2}$ .

(d) *Tổ hợp kết quả nhân chập và âm vang cuối*

Tiếp theo, kết quả nhân chập 814,  $\hat{\mathbf{z}}_{\text{ch,conv}}^{n,k} = [\hat{\mathbf{z}}_{\text{ch},1,\text{conv}}^{n,k}, \hat{\mathbf{z}}_{\text{ch},2,\text{conv}}^{n,k}]$ , và đầu ra của phản vang âm 826a, 826b,  $\hat{\mathbf{z}}_{\text{ch,rev}}^{n,k} = [\hat{\mathbf{z}}_{\text{ch},1,\text{rev}}^{n,k}, \hat{\mathbf{z}}_{\text{ch},2,\text{rev}}^{n,k}]$ , đối với khung đầu vào âm thanh miền QMF được tổ hợp bằng quy trình trộn 828 trong đó theo bằng bổ sung thêm hai tín hiệu. Lưu ý rằng các băng ở trên cao hơn  $K_{\text{max}}$  là 0 theo  $\hat{\mathbf{z}}_{\text{ch,conv}}^{n,k}$  bởi vì phép nhân chập chỉ được thực hiện trong các băng đến  $K_{\text{max}}$ .

Đầu ra âm vang cuối bị làm trễ bởi  $d = ((L_{\text{trans}} - 20 \cdot 64 + 1) / 64 + 0.5) + 1$  lượng khe thời gian trong quy trình trộn.

Độ trễ  $d$  dùng để chỉ thời gian chuyển tiếp từ các phản xạ đầu sang các phản xạ cuối cùng trong BRIR và độ trễ ban đầu của phản vang âm của 20 khe thời gian QMF, cũng như giá trị trễ phân tích của các khe thời gian 0,5 QMF để phân tích QMF của

BRIR để đảm bảo đưa âm vang cuối vào tại khe thời gian hợp lý. Tín hiệu tổ hợp  $\hat{\mathbf{z}}_{\text{ch}}^{n,k}$  tại một khe thời gian  $n$  được tính toán bằng  $\hat{\mathbf{z}}_{\text{ch,conv}}^{n,k} + \hat{\mathbf{z}}_{\text{ch,rev}}^{n-d,k}$ .

(e) *Tổng hợp QMF của tín hiệu miền QMF lập thể*

Một khung 2-kênh của 32 khe thời gian của tín hiệu đầu ra miền QMF  $\hat{\mathbf{z}}_{\text{ch}}^{n,k}$  được biến đổi thành khung tín hiệu miền thời gian 2-kênh có độ dài  $L$  bằng cách tổng hợp QMF theo ISO/IEC 14496-3:2009, điều khoản 4.6.18.4.2. thu được tín hiệu đầu ra miền thời gian cuối cùng 830,  $\tilde{\mathbf{z}}_{\text{ch}}^v = [\tilde{z}_{\text{ch},1}^v \cdots \tilde{z}_{\text{ch},2}^v]$ .

Mặc dù các khía cạnh của sáng chế được mô tả theo ngữ cảnh thiết bị, nhưng rõ ràng rằng các khía cạnh này cũng thể hiện sự mô tả về các phương pháp tương ứng, trong đó khối hoặc thiết bị tương ứng với bước của phương pháp hoặc đặc tính của bước của phương pháp. Tương tự, các khía cạnh được mô tả theo ngữ cảnh của các bước của phương pháp cũng thể hiện sự mô tả khối, bộ phận hoặc đặc điểm của thiết bị tương ứng. Một số hoặc tất cả các bước có thể được thực hiện bởi (hoặc bằng cách sử dụng) thiết bị phần cứng, ví dụ như, bộ vi xử lý, máy tính có thể lập trình hoặc mạch điện tử. Theo một số phương án, một hoặc nhiều bước phương pháp quan trọng nhất có thể được thực hiện bởi thiết bị này.

Phụ thuộc vào yêu cầu thiết lập nhất định, các phương án của sáng chế có thể được thực hiện trên phần cứng hoặc phần mềm. Việc thiết lập có thể được thực hiện nhờ sử dụng vật ghi lưu trữ không chuyển tiếp như vật ghi lưu trữ kỹ thuật số, ví dụ đĩa mềm, DVD, Blu-Ray, CD, ROM, PROM, và EPROM, bộ nhớ EEPROM hoặc FLASH, có các tín hiệu điều khiển có thể đọc được bằng điện tử được lưu trữ trên đó, tổ hợp (hoặc có khả năng kết hợp p) với hệ thống máy tính có thể lập trình được sao cho phương pháp tương ứng được thực hiện. Do đó, vật ghi lưu trữ kỹ thuật số có thể đọc được trên 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 có thể lập trình được, sao cho một trong số các phương pháp được mô tả ở đây được thực hiện.

Nói chung, các phương án của có thể được thực hiện như sản phẩm chương trình máy tính có mã chương trình, mã chương trình này có tác dụng để thực hiện một trong số các phương pháp khi sản phẩm chương trình máy tính được chạy trên máy tính. Mã chương trình này, ví dụ, có thể 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.

Nói cách khác, phương án theo sáng chế, do đó, 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 được chạy trên máy tính.

Phương án khác theo sáng chế, do đó, là vật mang dữ liệu (hoặc vật ghi lưu trữ kỹ thuật số, hoặc vật ghi có thể đọc được trên máy tính) bao gồm, được ghi trên đó, chương trình máy tính để thực hiện một trong số các phương pháp được mô tả ở đây. Vật mang dữ liệu, vật ghi lưu trữ kỹ thuật số hoặc vật ghi được ghi thường là hữu hình và/hoặc không chuyển tiếp.

Phương án khác theo sáng chế, do đó, là dòng dữ liệu hoặc chuỗi tín hiệu thể hiện chương trình máy tính để thực hiện một trong số các phương pháp được mô tả ở đây. Dòng dữ liệu hoặc chuỗi tín hiệu, ví dụ, được cấu hình để được truyền qua kết nối truyền thông dữ liệu, ví dụ, thông qua liên mạng.

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

Phương án khác của sáng chế bao gồm máy tính có chương trình máy tính được thiết lập để thực hiện một trong số các phương pháp được mô tả ở đây.

Phương án khác theo sáng chế bao gồm thiết bị hoặc hệ thống được cấu hình để truyền (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, ví dụ, có thể là máy tính, thiết bị di động, thiết bị ghi nhớ hoặc thiết bị tương tự. Thiết bị hoặc hệ

thống này, ví dụ, có thể bao gồm máy chủ tệp tin để truyền chương trình máy tính đến bộ nhận.

Theo một số phương án, thiết bị logic có thể lập trình được (ví dụ, mảng cổng có thể lập trình) 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. Theo một số phương án, mảng cổng có thể lập trình có thể kết hợp với bộ vi xử lý để thực hiện một trong số các phương pháp được mô tả ở đây. Nói chung, các phương pháp được ưu tiên thực hiện bởi thiết bị phần cứng bất kỳ.

Các phương án nêu trên chỉ nhằm mục đích minh họa sáng chế. Cần hiểu rằng các biến đổi và thay đổi về cách bố trí và các chi tiết được mô tả ở đây sẽ là rõ ràng với người có hiểu biết trung bình trong lĩnh vực kỹ thuật. Do đó, sáng chế chỉ bị giới hạn bởi phạm vi yêu cầu bảo hộ kèm theo và không bị giới hạn ở phần mô tả chi tiết cụ thể được mô tả và giải thích theo các phương án ở đây.

## YÊU CẦU BẢO HỘ

1. Phương pháp xử lý tín hiệu âm thanh (400, 802) theo đáp ứng xung phòng (300, 804), phương pháp bao gồm:

xử lý một cách riêng biệt (402, 404, 406, 422, 424, 812, 816a, 816b) tín hiệu âm thanh (400, 802) với phần đầu (301, 302) của đáp ứng xung phòng (300, 804) và với âm vang cuối (304) của đáp ứng xung phòng (300, 804) hoặc âm vang tổng hợp; và

tổ hợp (418, 432, 828) tín hiệu âm thanh được xử lý với phần đầu (414, 428, 814) của đáp ứng xung phòng (300, 804) và tín hiệu âm thanh (416, 430, 826a, 826b) được xử lý với âm vang cuối (304) của đáp ứng xung phòng (300, 804) hoặc với âm vang tổng hợp,

xác định sự chuyển tiếp từ phần đầu (301, 302) tới âm vang cuối (304) trong đáp ứng xung phòng (300, 804) tại thời điểm khi số đo tương quan đạt ngưỡng,

trong đó số đo tương quan mô tả đối với đáp ứng xung phòng (300, 804) sự tương tự của sự phân rã trong năng lượng âm thanh gồm có trạng thái ban đầu và của sự phân rã trong năng lượng âm thanh bắt đầu tại thời điểm, thời điểm đã nêu theo sau trạng thái ban đầu trên khoảng tần số được định trước,

trong đó ngưỡng được thiết lập phụ thuộc vào số đo tương quan cho thời điểm đã nêu, thời điểm đã nêu là thời điểm của phản xạ được lựa chọn trong số các phản xạ trước (302, 602) trong phần đầu (301, 302) của đáp ứng xung phòng (300, 804), và

trong đó một phản xạ được lựa chọn trong số các phản xạ trước là phản xạ thứ nhất (602).

2. Phương pháp theo điểm 1, trong đó việc xác định sự chuyển tiếp bao gồm:

xác định (500) sự phân bố của năng lượng âm thanh dựa trên đáp ứng xung phòng (300, 804); và;

xác định (502) nhiều số đo tương quan biểu thị cho nhiều phần của việc phân bố được xác định tương quan giữa năng lượng âm thanh trong phần tương ứng của sự phân bố được xác định và năng lượng âm thanh tại trạng thái ban đầu.

3. Phương pháp theo điểm 2, trong đó:

xác định sự phân bố bao gồm việc xác định sự phân bố thời gian-tần số của năng lượng âm thanh, và

một phần của việc phân bố bao gồm khối thời gian có độ dài định trước, trạng thái ban đầu được định rõ bởi một khối thời gian thứ nhất trong số nhiều khối thời gian của việc phân bố thời gian-tần số.

4. Phương pháp theo điểm 2 hoặc 3, trong đó việc xác định việc phân bố bao gồm tính toán sự giảm phân rã năng lượng, *EDR*, từ đáp ứng xung phòng (300, 804).

5. Phương pháp theo điểm 4, trong đó sự giảm phân rã năng lượng, *EDR*, được tính toán như sau:

$$E(t, \omega) = \left| \int_t^{\infty} h(\tau) e^{-j\omega\tau} d\tau \right|^2$$

trong đó:

$$\begin{aligned} E(t, \omega) &= \text{sự giảm phân rã năng lượng,} \\ h(\tau) &= \text{đáp ứng xung phòng (300, 804),} \\ \omega &= 2\pi f. \end{aligned}$$

6. Phương pháp theo một trong số các điểm từ 3 đến 5, trong đó đáp ứng xung phòng (300, 804) có độ dài thực được định trước, và trong đó việc xác định sự phân bố thời gian-tần số bao gồm việc tính toán phổ FFT của đáp ứng xung phòng (300, 804) sử dụng cửa sổ có độ dài tương ứng với độ dài thực của đáp ứng xung phòng.

7. Phương pháp theo điểm 6, trong đó:

năng lượng âm thanh ở trạng thái ban đầu được xác định bằng việc lấy tổng độ dài thực của đáp ứng xung phòng (300, 804) (300, 804), tính toán phổ FFT và lấy bình phương các giá trị tuyệt đối, và

năng lượng âm thanh của khối thời gian được xác định bằng việc chuyển cửa sổ bởi thời gian được kết hợp với khối thời gian, đệm 0 các mẫu được tạo cửa sổ thành độ dài thực, tính toán FFT và lấy bình phương các giá trị tuyệt đối.

8. Phương pháp theo một trong số các điểm từ 1 đến 7, trong đó số đo tương quan được tính toán như sau:

$$\rho(t) = \frac{\sum_{\omega} (E(1, \omega) - \bar{E}(1, \omega)) \cdot \sum_{\omega} (E(t, \omega) - \bar{E}(t, \omega))}{\sqrt{\sum_{\omega} (E(1, \omega) - \bar{E}(1, \omega))^2} \cdot \sqrt{\sum_{\omega} (E(t, \omega) - \bar{E}(t, \omega))^2}}$$

trong đó:

- $\rho(t)$  = số đo tương quan,
- $E(1, \omega)$  = sự giảm phân rã năng lượng tại tần số  $f$ ,
- $\bar{E}(1, \omega)$  = giá trị trung bình trên tất cả các tần số của sự giảm phân rã năng lượng ban đầu,
- $E(t, \omega)$  = sự giảm phân rã năng lượng tại tần số  $f$  bắt đầu tại thời gian  $t$ ,
- $\bar{E}(t, \omega)$  = giá trị trung bình trên tất cả các tần số của sự giảm phân rã năng lượng bắt đầu tại thời gian  $t$ ,
- $\omega$  =  $2\pi f$ .

9. Phương pháp theo một trong số các điểm từ 1 đến 8, trong đó ngưỡng được xác định dựa trên giá trị không đổi và số đo tương quan đối với một phản xạ được lựa chọn trong số các phản xạ trước (302, 602).

10. Phương pháp theo điểm 9, trong đó hằng số là  $1/e$ , và trong đó ngưỡng được định rõ như sau

$$\rho(t) = c \cdot \rho(t_F)$$

trong đó:

- $\rho(t_F)$  = số đo tương quan đối với một phản xạ được lựa chọn trong số các phản xạ trước (302, 602),
- $t_F$  = chỉ số thời gian trong đó phản xạ được chọn trong số các phản xạ trước (302, 602) sau khi âm thanh trực tiếp (300) va chạm,
- $c$  = giá trị không đổi là  $\frac{1}{e}$ ,  $e$  là số Euler.

11. Sản phẩm chứa chương trình máy tính không hữu hình bao gồm vật ghi có thể đọc được bằng máy tính lưu trữ các lệnh mà, khi chạy trên máy tính, thực hiện phương pháp theo một trong số các điểm từ 1 đến 10.

12. Bộ phận xử lý tín hiệu, bao gồm:

đầu vào để nhận tín hiệu âm thanh (400, 802);

bộ xử lý được tạo cấu hình để xử lý tín hiệu âm thanh (400, 802) với phần đầu (301, 302) của đáp ứng xung phòng (300, 804) và âm vang cuối (304) của đáp ứng xung phòng (300, 804) hoặc âm vang tổng hợp; và

đầu ra để tổ hợp (418, 432, 828) tín hiệu âm thanh được xử lý với phần đầu (414, 428, 814) của đáp ứng xung phòng (300, 804) và tín hiệu âm thanh (416, 430, 826a, 826b) được xử lý với âm vang cuối (304) của đáp ứng xung phòng (300, 804) hoặc với âm vang tổng hợp, thành tín hiệu âm thanh đầu ra;

trong đó bộ xử lý được tạo cấu hình để xác định sự chuyển tiếp từ phần đầu (301, 302) tới âm vang cuối (304) trong đáp ứng xung phòng (300, 804) tại thời điểm khi số đo tương quan đạt ngưỡng,

trong đó số đo tương quan mô tả đối với đáp ứng xung phòng (300, 804) sự tương tự của sự phân rã trong năng lượng âm thanh gồm có trạng thái ban đầu và của sự phân rã trong năng lượng âm thanh bắt đầu tại thời điểm, thời điểm đã nêu theo sau trạng thái ban đầu trên khoảng tần số đã định trước,

trong đó ngưỡng được thiết lập phụ thuộc vào số đo tương quan cho thời điểm đã nêu, thời điểm đã nêu là thời điểm của phản xạ được lựa chọn trong số các phản xạ trước (302, 602) trong phần đầu (301, 302) của đáp ứng xung phòng (300, 804), và

trong đó một phản xạ được lựa chọn trong số các phản xạ trước là phản xạ thứ nhất (602).

13. Bộ phận xử lý tín hiệu theo điểm 12, còn bao gồm:

bộ xử lý phần đầu (402, 404, 422, 812) để xử lý tín hiệu âm thanh nhận được (400, 802) theo phần đầu (301, 302, 602) của đáp ứng xung phòng (300, 804); và

bộ xử lý âm vang cuối (403, 424, 816a, 816b) để xử lý tín hiệu âm thanh nhận được (400, 802) theo âm vang cuối (304) của đáp ứng xung phòng (300, 804).

14. Bộ mã hóa âm thanh để mã hóa tín hiệu âm thanh, trong đó:  
bộ mã hóa âm thanh được tạo cấu hình để xử lý tín hiệu âm thanh sẽ được mã hóa theo  
đáp ứng xung phòng (300, 804) theo một phương pháp trong số các phương pháp của  
các điểm từ 1 đến 10.
15. Bộ mã hóa âm thanh theo điểm 14, trong đó bộ mã hóa âm thanh bao gồm bộ  
phận xử lý tín hiệu theo điểm 12 hoặc 13.
16. Bộ giải mã âm thanh để giải mã tín hiệu âm thanh được mã hóa, trong đó:  
bộ giải mã âm thanh được tạo cấu hình để xử lý tín hiệu âm thanh được giải mã (400,  
802) theo đáp ứng xung phòng (300, 804) theo một phương pháp trong số các phương  
pháp của các điểm từ 1 đến 10.
17. Bộ giải mã âm thanh theo điểm 16, trong đó bộ giải mã âm thanh bao gồm bộ  
phận xử lý tín hiệu theo điểm 12 hoặc 13.
18. Bộ kết xuất âm lập thể bao gồm bộ phận xử lý tín hiệu của điểm 12 hoặc 13.

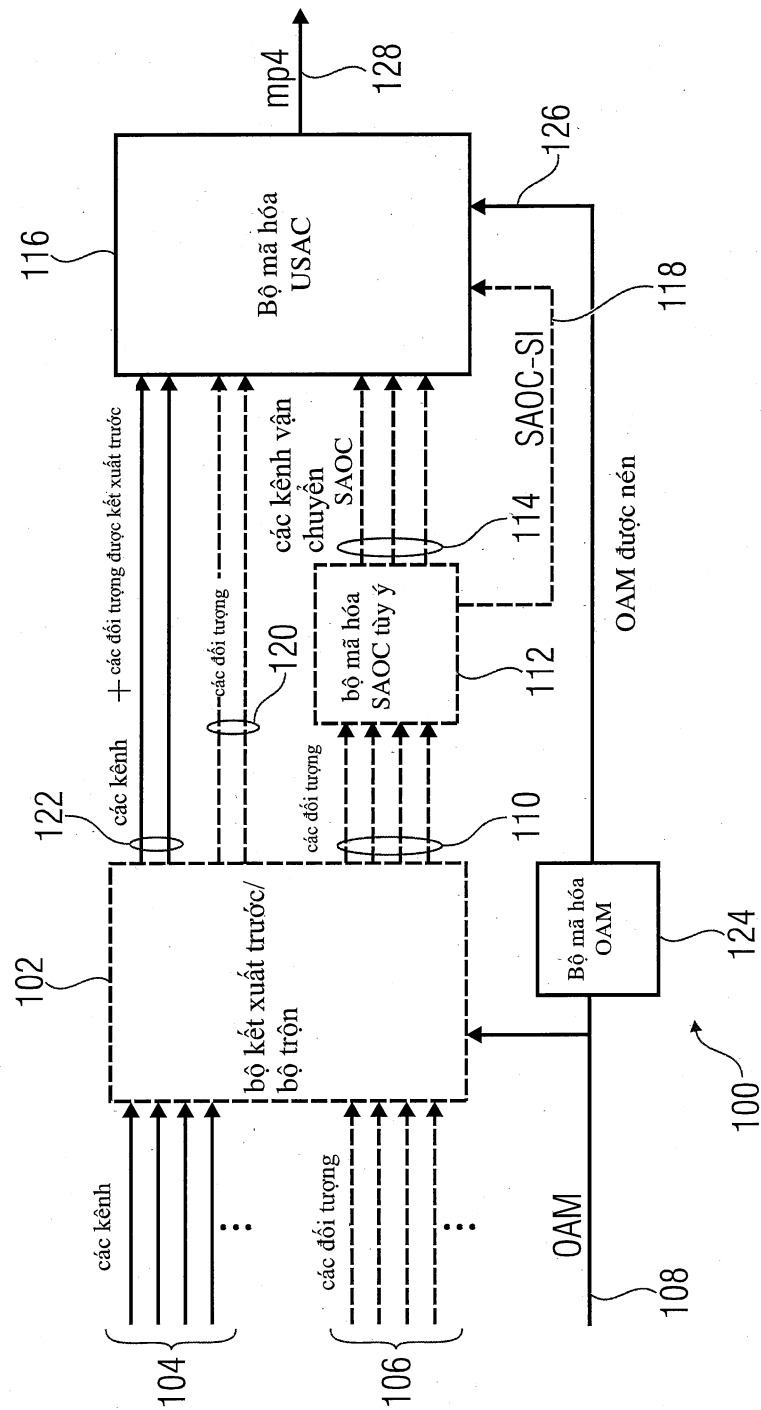


FIG 1

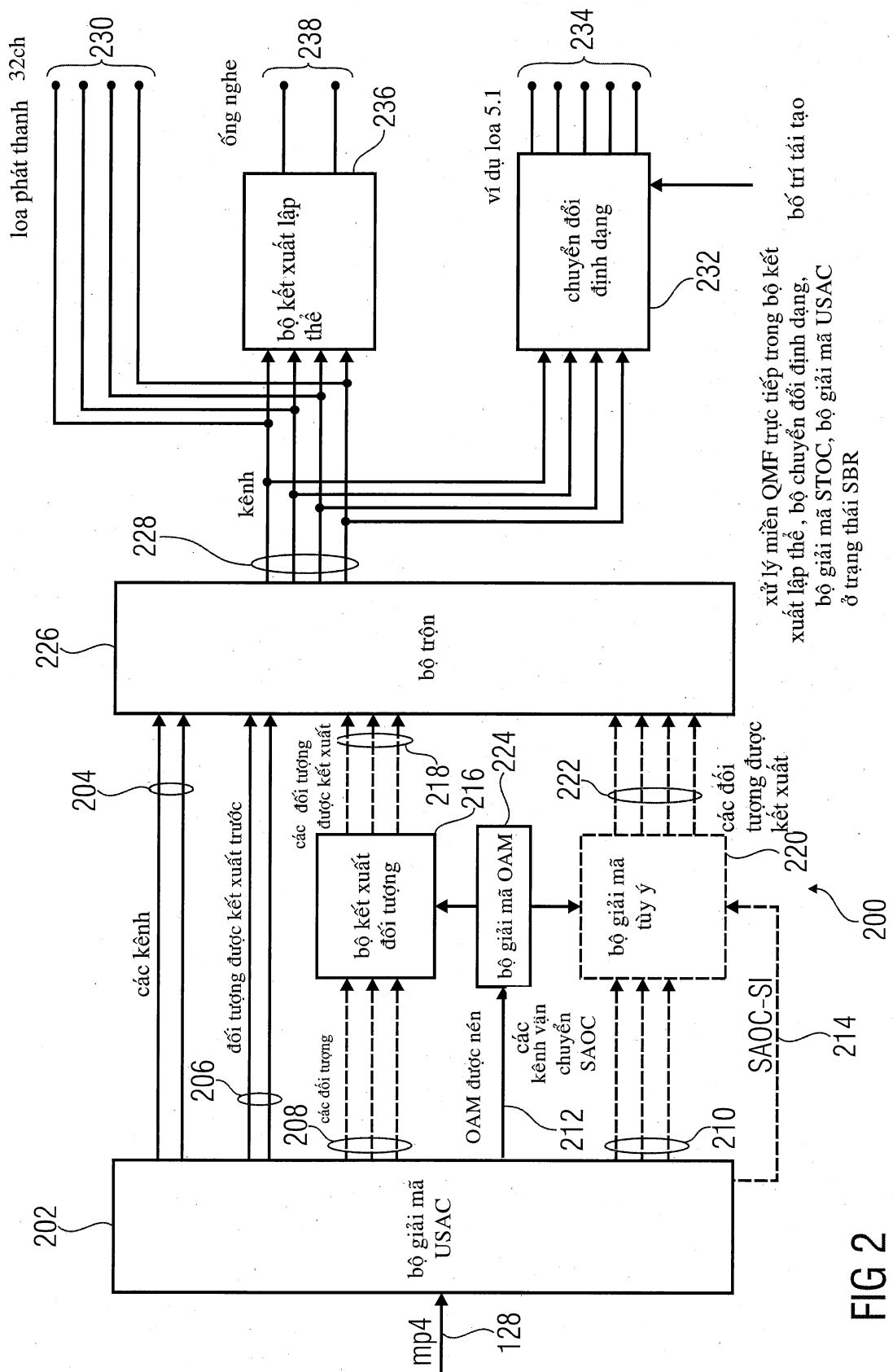


FIG 2

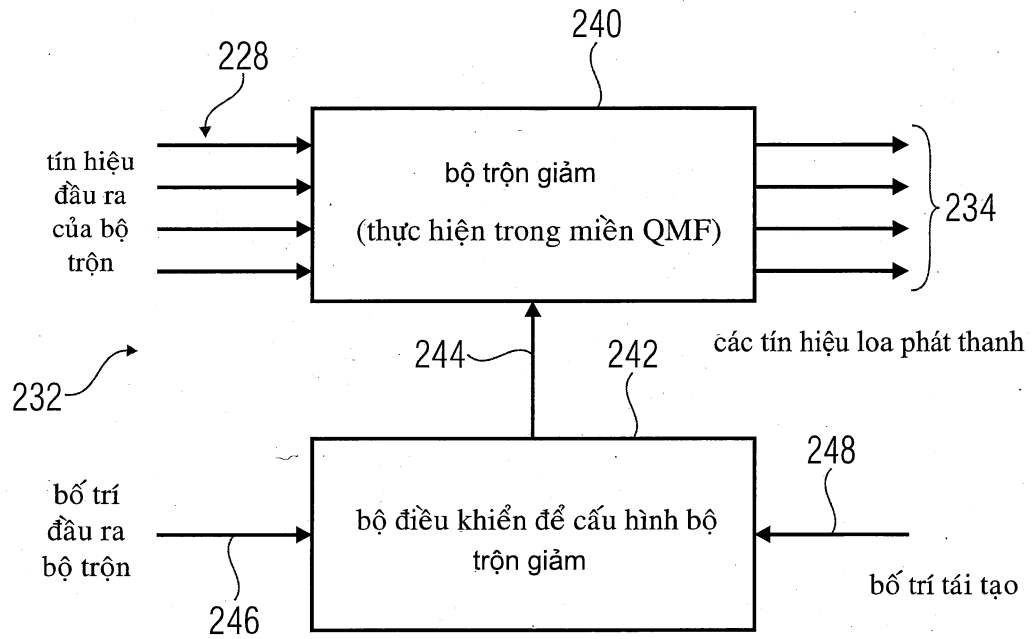


FIG 3

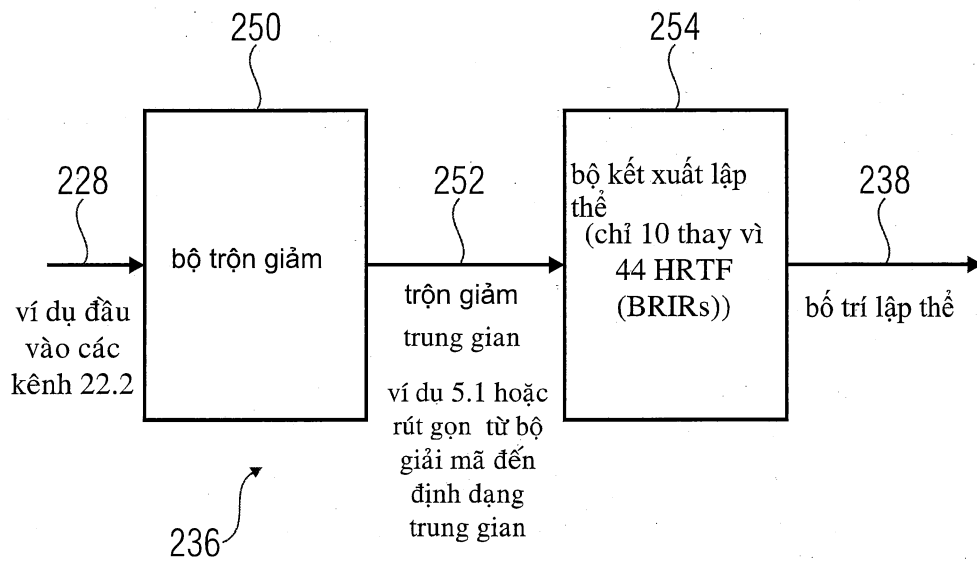


FIG 4

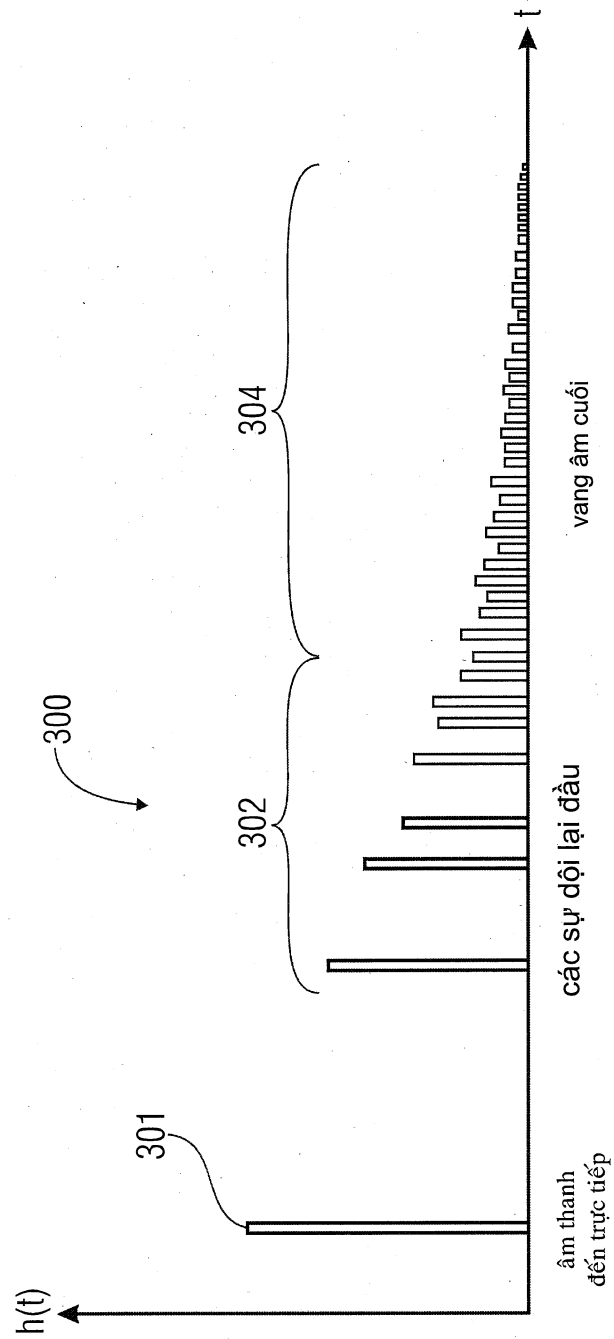


FIG 5

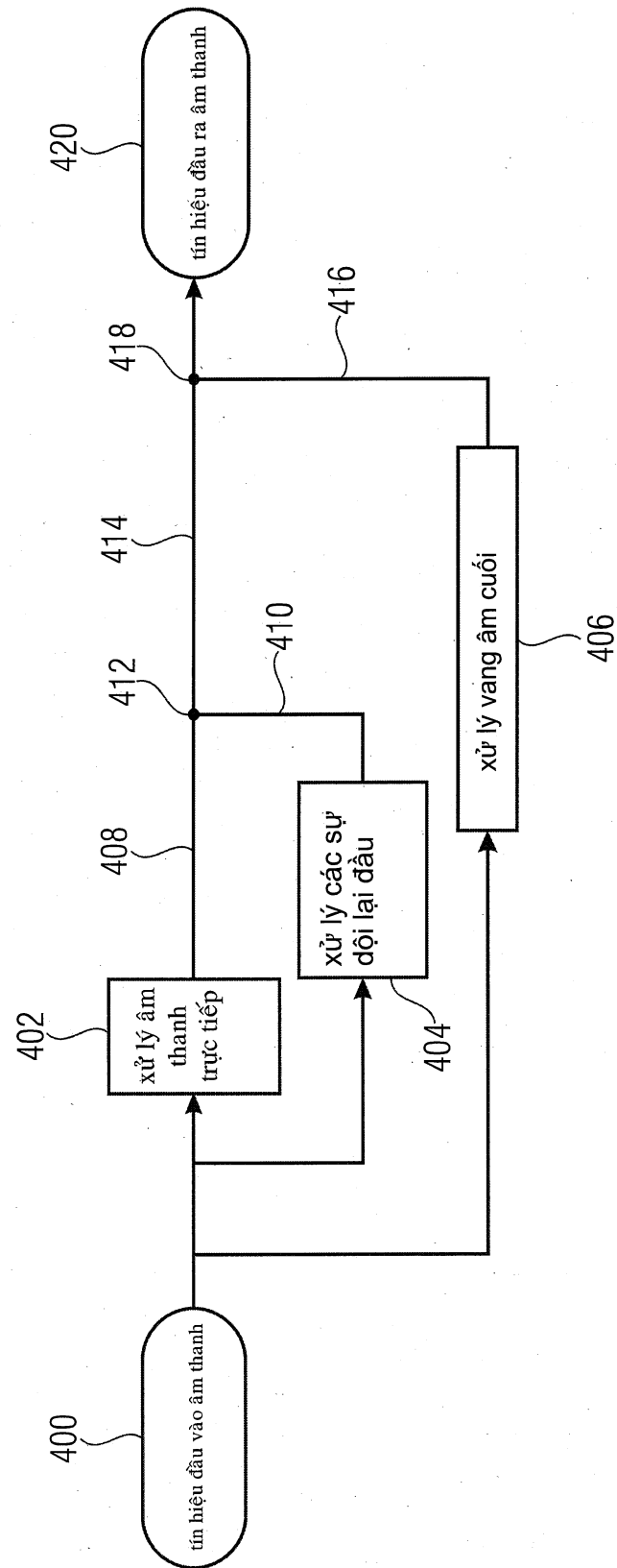


FIG 6A

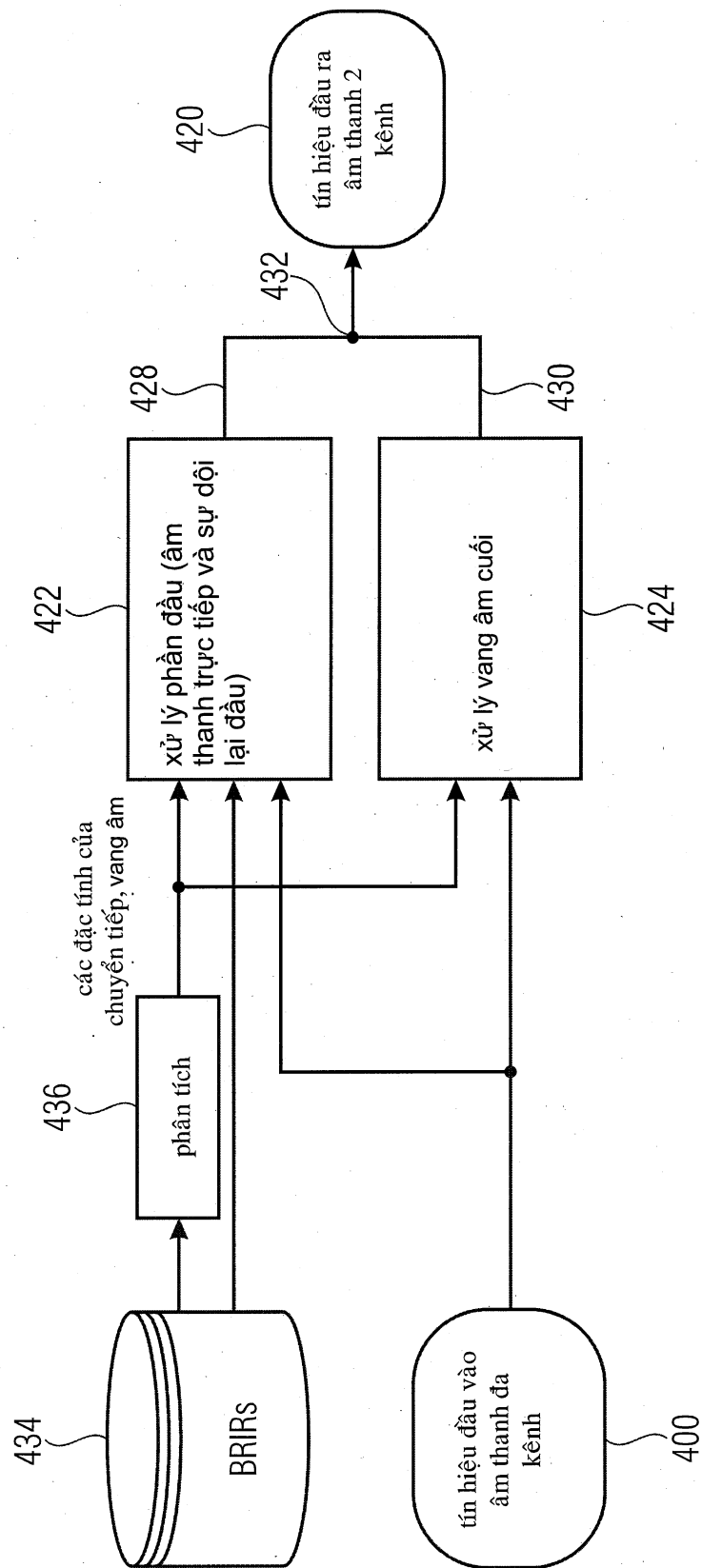


FIG 6B

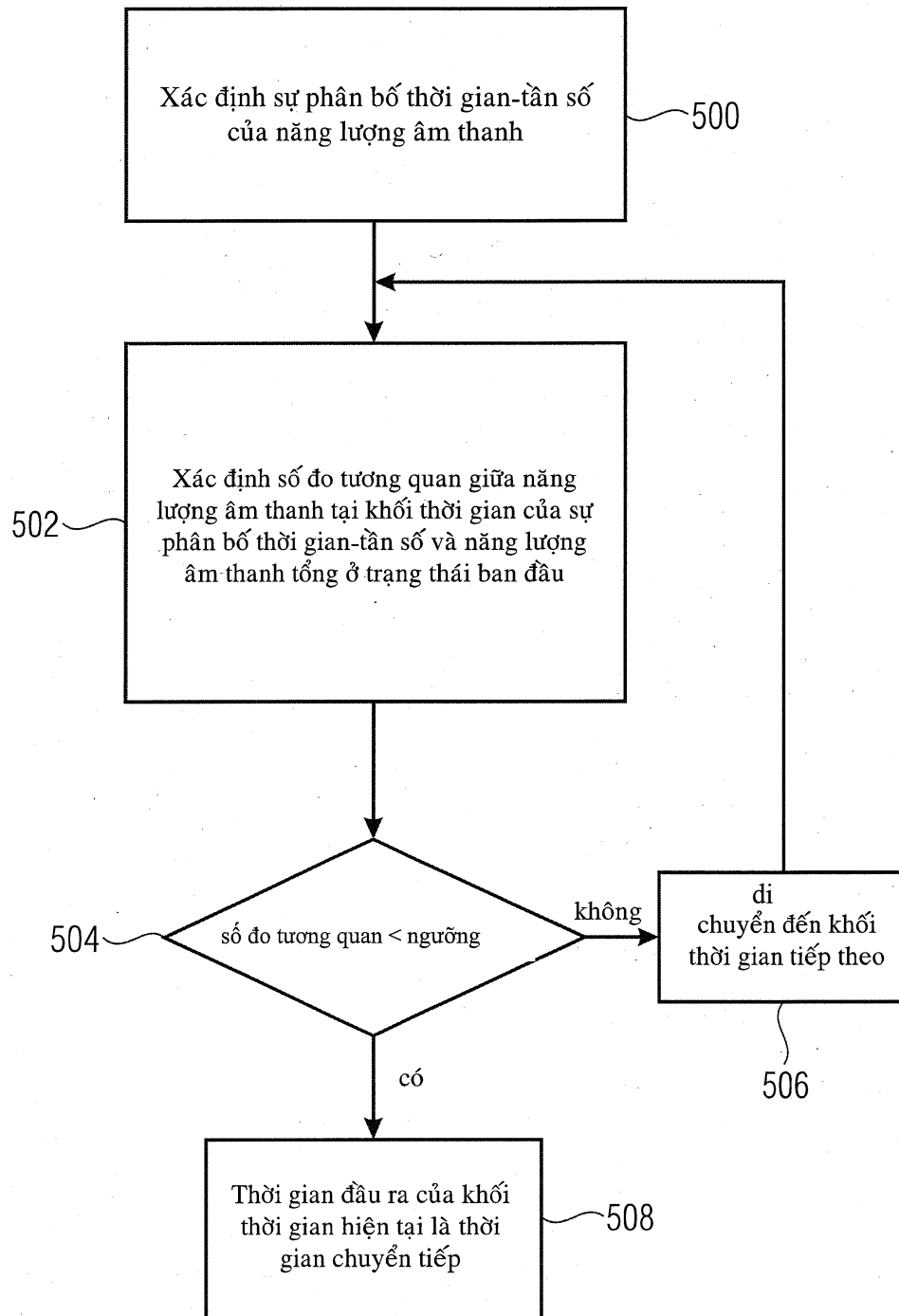


FIG 7

sự giảm phân rã năng lượng

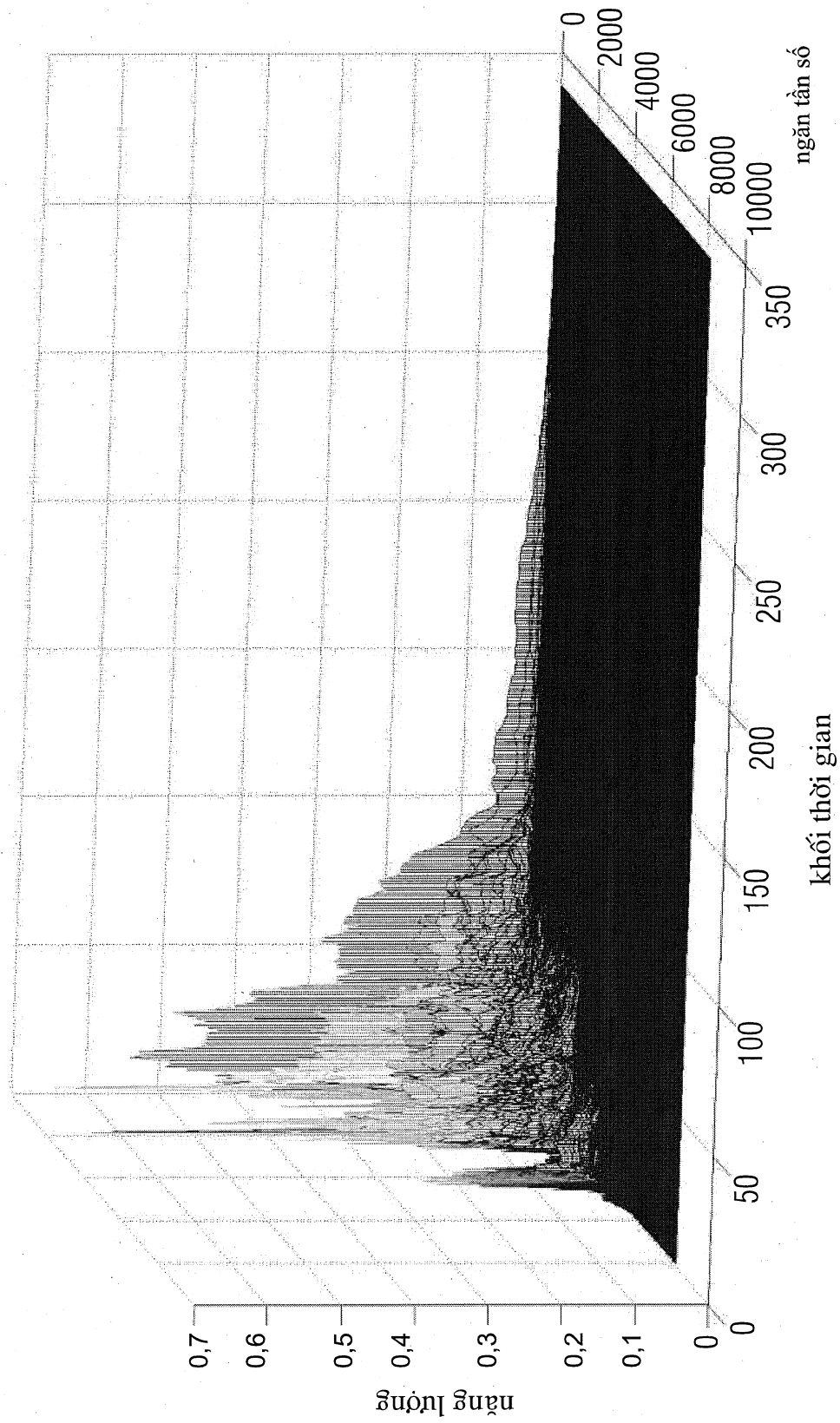


FIG 8

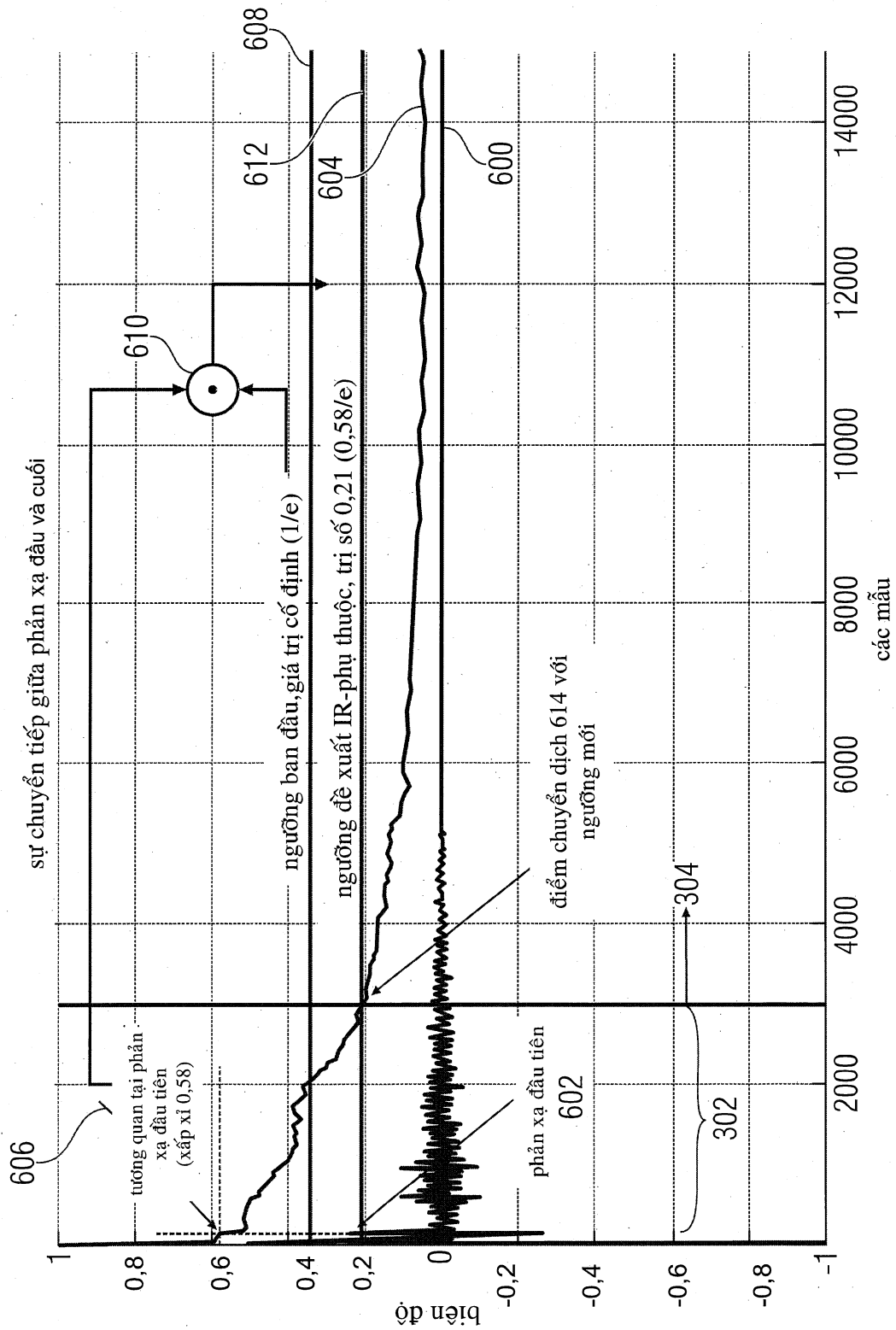


FIG 9

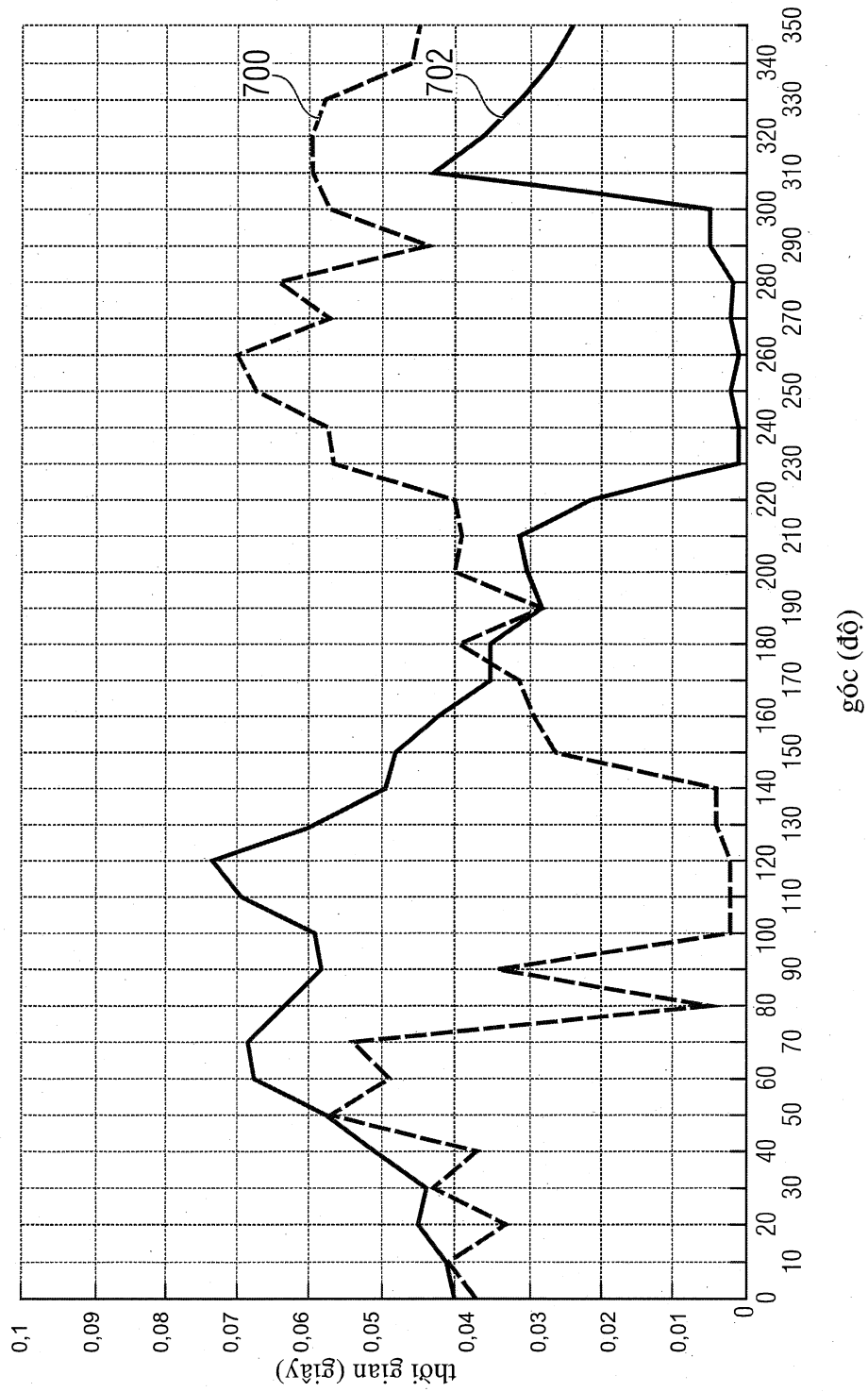


FIG 10

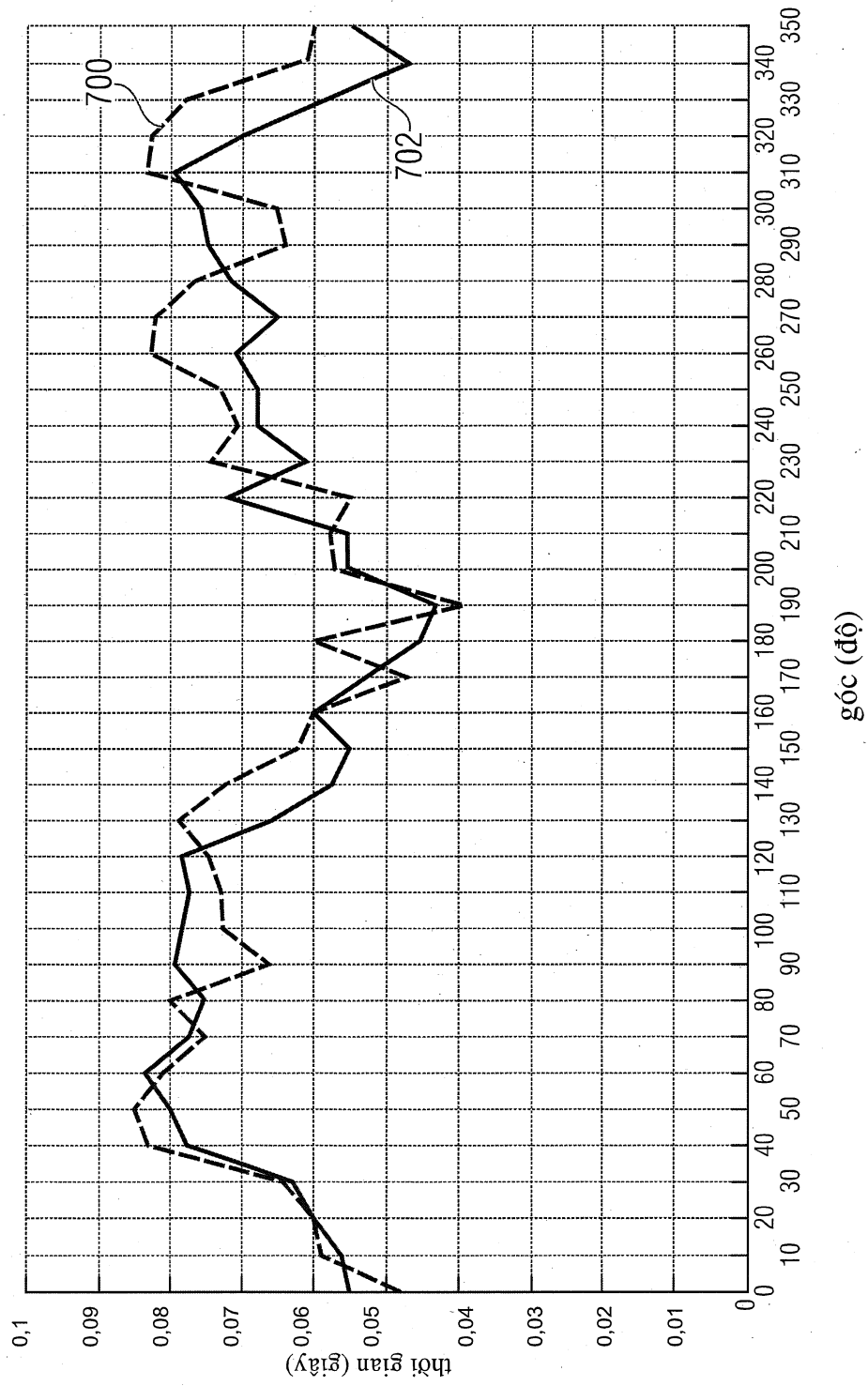


FIG 11

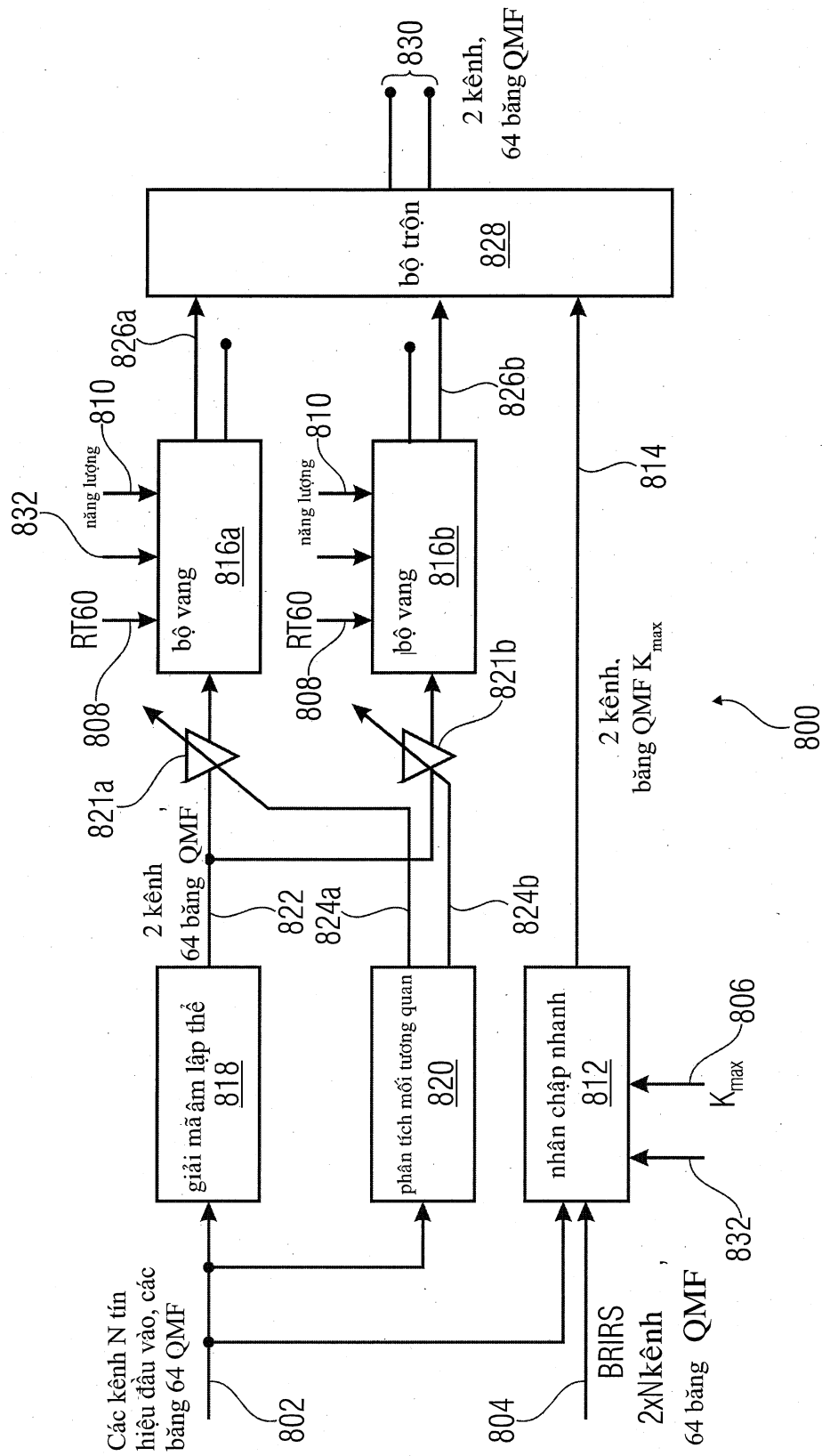


FIG 12

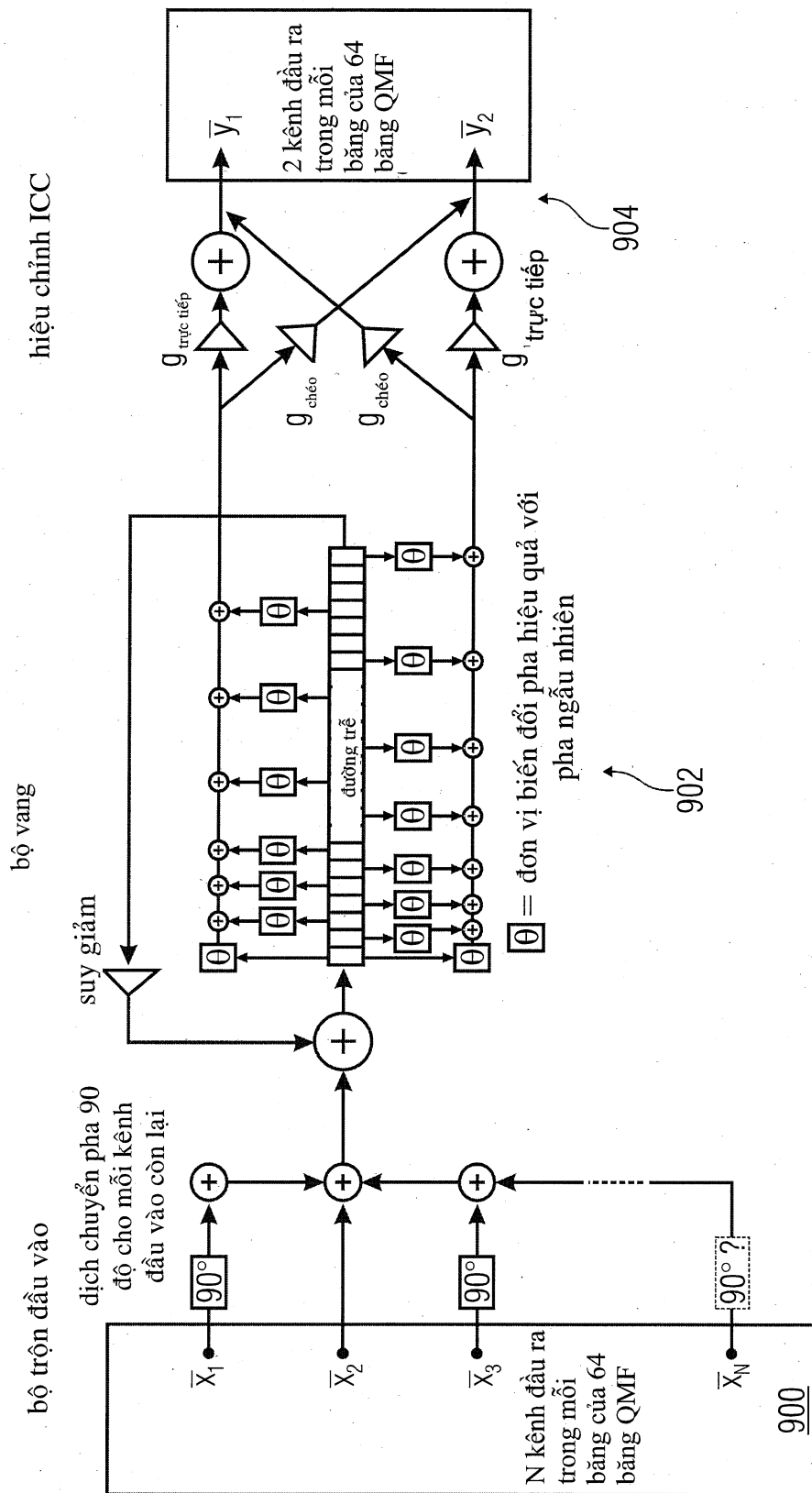


FIG 13