



(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)



1-0025103

(51)⁷ G10L 19/008; H04S 7/00; H04S 3/00

(13) B

(21) 1-2015-04000

(22) 17/03/2014

(86) PCT/EP2014/055313 17/03/2014

(87) WO2014/147029 25/09/2014

(30) 13305352.0 22/03/2013 EP

(45) 25/08/2020 389

(43) 25/12/2015 333A

(73) DOLBY INTERNATIONAL AB (SE)

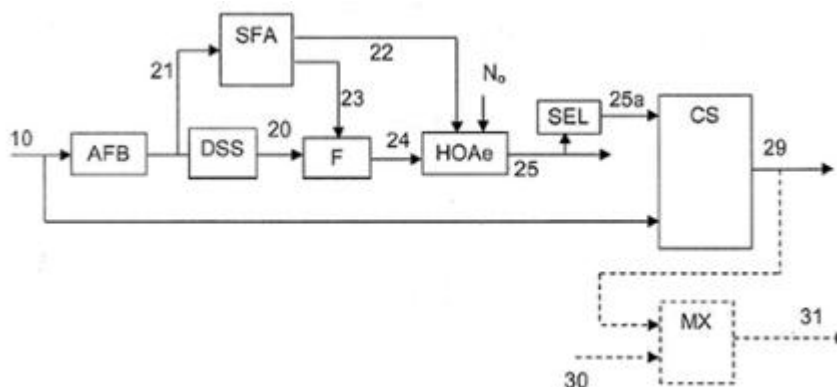
Apollo Building, 3E, Herikerbergweg 1-35, 1101 CN Amsterdam Zuidoost,
Netherlands

(72) BOEHM, Johannes (DE).

(74) Công ty TNHH Quốc tế D & N (D&N INTERNATIONAL CO.,LTD.)

(54) PHƯƠNG PHÁP VÀ THIẾT BỊ NÂNG CAO HƯỚNG TÍNH CỦA TÍN HIỆU ĐẦU VÀO LÀ TÍN HIỆU AMBISONIC BẬC 1 VÀ CÓ CÁC HỆ SỐ BẬC 0 VÀ BẬC 1

(57) Sáng chế đề cập đến phương pháp và thiết bị nâng cao hướng tính của tín hiệu đầu vào là tín hiệu Ambisonic bậc 1 và có các hệ số bậc 0 và bậc 1. Các bản ghi từ các micro mà cung cấp các tín hiệu Ambisonic bậc 1, được gọi là các tín hiệu định dạng B, cung cấp sự nhận biết giới hạn về hướng tính âm thanh. Các nguồn âm thanh được cảm nhận rộng hơn so với thực tế, đặc biệt là đối với các vị trí nghe lệch tâm, và các nguồn âm thanh thường được định vị là đi từ các vị trí loa gần nhất. Trong phương pháp và thiết bị nâng cao hướng tính của các tín hiệu Ambisonic bậc 1, thông tin hướng tính bổ sung (22,23) được trích (SFA) từ tín hiệu đầu vào Ambisonic bậc thấp hơn (10). Thông tin hướng tính bổ sung được sử dụng để đánh giá các hệ số Ambisonic bậc cao hơn (25a), mà sau đó được kết hợp (CS) với các hệ số của tín hiệu đầu vào. Do đó, hướng tính của tín hiệu Ambisonic được nâng cao, mà dẫn đến nâng cao độ chính xác của việc định vị nguồn không gian khi tín hiệu Ambisonic được giải mã thành các tín hiệu loa. Tín hiệu đầu ra thu được có năng lượng cao hơn so với tín hiệu đầu vào.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến lĩnh vực xử lý tín hiệu âm thanh Ambisonic và kỹ thuật âm thanh.

Tình trạng kỹ thuật của sáng chế

Ambisonic là kỹ thuật mô tả cảnh âm thanh về mặt áp suất âm thanh, và thực hiện việc ghi, sản xuất, truyền và phát lại các cảnh âm thanh phức với độ phân giải không gian ưu việt, cả trong 2D và 3D. Trong Ambisonic, cảnh âm thanh không gian được mô tả bởi các hệ số $A_n^m(k)$ của chuỗi Fourier-Bessel. Các dãy micrô cung cấp các tín hiệu Ambisonic bậc 1 được gọi là các tín hiệu định dạng B được biết đến. Tuy nhiên, việc giải mã và kết xuất các tín hiệu Ambisonic bậc 1 tới các bố trí loa đối với âm thanh vòm 2D hoặc 3D chỉ có sự nhận biết giới hạn về hướng tính âm thanh. Các nguồn âm thanh thường được cảm nhận là rộng hơn so với thực tế. Đặc biệt đối với các vị trí nghe lệch tâm, các nguồn âm thanh thường được định vị là đi từ các vị trí loa gần nhất, thay vì vị trí ảo dự định của chúng giữa các loa. Các tín hiệu Ambisonic (định dạng B) bậc 1 bao gồm bốn hệ số của mô tả chuỗi Fourier-Bessel của áp suất âm thanh, mà tạo thành dạng biểu diễn trường âm thanh 3D. Đó là kênh W (trộn mono, hoặc bậc 0) và các kênh X,Y,Z (bậc 1). Các tín hiệu bậc cao hơn sử dụng nhiều hệ số hơn, mà gia tăng độ chính xác của việc định vị nguồn không gian khi các hệ số được giải mã thành các tín hiệu loa. Tuy nhiên, các tín hiệu bậc cao hơn này không được chứa trong các tín hiệu định dạng B được đưa ra bởi các dãy micrô.

Mã hóa âm thanh có hướng (Directional Audio Coding - DirAC) là kỹ thuật đã biết [5,9] để biểu diễn hoặc tái tạo các tín hiệu âm thanh. Kỹ thuật này sử dụng bộ giải mã định dạng B mà tách biệt âm thanh trực tiếp với âm thanh khuếch tán, sau đó sử dụng quét biên độ dựa trên vector (Vector-Based Amplitude Panning - VBAP) để khuếch đại có chọn lọc âm thanh trực tiếp trong miền tần số, và sau khi lọc tổng hợp cuối cùng cung cấp các tín hiệu loa tại đầu ra của nó.

Fig.1a) thể hiện cấu trúc của việc giải mã định dạng B dựa trên DirAC. Các tín hiệu định dạng B 10 là các tín hiệu miền thời gian, và được lọc trong băng lọc phân tích AFB_D thành K băng tần số 11. Khối phân tích trường âm thanh SFA_D đánh giá sự đánh giá khuếch tán $\Psi(f_k)$ 13 và các hướng đi đến (directions-of-arrival - DoA) 12. DoA là phương vị $\phi(f_k)$ và độ nghiêng $\theta(f_k)$ của các hướng tới nguồn tại tần số trung cụ thể của băng k . Bộ giải mã Ambisonic bậc 1 AmbD kết xuất các tín hiệu Ambisonic tới L tín hiệu loa 14. Khối tách biệt trực tiếp-khuếch tán DDS tách biệt các tín hiệu Ambisonic bậc 1 thành L tín hiệu âm thanh trực tiếp 15 và L tín hiệu âm thanh khuếch tán 16, nhờ sử dụng bộ lọc mà được xác định từ đánh giá khuếch tán 13. L tín hiệu âm thanh khuếch tán 16 thu được bằng cách nhân đầu ra 14 của bộ giải mã AmbD với $\sqrt{\Psi(f_k)}$, mà thu được từ đánh giá khuếch tán 13. Các tín hiệu có hướng thu được từ việc nhân với $\sqrt{1 - \Psi(f_k)}$. Các tín hiệu âm thanh trực tiếp 15 còn được xử lý nhờ sử dụng kỹ thuật gọi là quét biên độ dựa trên vector (Vector Base Amplitude Panning - VBAP) [8]. Trong bộ VBAP VP, giá trị khuếch đại đối với mỗi tín hiệu loa (trong mỗi băng tần số) được nhân để quét âm thanh trực tiếp tới các hướng mong muốn, theo DoA 12 và các vị trí của các loa. Các tín hiệu khuếch tán 16 được giải tương quan bởi việc lọc giải tương quan DF, và các tín hiệu khuếch tán được giải tương quan 17 được bổ sung vào các tín hiệu âm thanh trực tiếp thu được từ bộ VPAB VP. Băng lọc tổng hợp SFB_D kết hợp các băng tần số tới tín hiệu miền thời gian 19, mà có thể được tái tạo bởi L loa. Các bộ lọc làm trơn (không được thể hiện trên Fig.1) để tích hợp theo thời gian được áp dụng để tính toán đánh giá khuếch tán $\Psi(f)$ 13 và để làm trơn các giá trị khuếch đại mà thu được bởi VBAP.

Fig.1b) thể hiện các chi tiết của khối phân tích trường âm thanh SFA_D . Các tín hiệu định dạng B biểu diễn trường âm thanh trong miền tần số tại gốc (vị trí quan sát, $r = 0$). Cường độ âm thanh mô tả việc truyền tải động năng hoặc thế năng trong trường âm thanh. Trong trường âm thanh, không phải tất cả sự di chuyển cục bộ của năng lượng âm thanh tương ứng với truyền tải mạng. Cường độ hoạt động I_a (cường độ âm thanh được lấy trung bình theo thời gian, $DoA \sim I_a$) là tốc độ của truyền tải năng lượng thuần có hướng – năng lượng trên thời gian đơn vị đối với ba hướng Cartesian. Cường độ hoạt động 11a của tín hiệu định dạng B 11 thu được trong khối phân tích cường độ hoạt động AIA_D , và được cấp tới khối phân tích khuếch tán

DAB_D và khối phân tích DoA $DOAAB_D$, mà xuất ra DoA 12 và đánh giá khuếch tán 13, một cách tương ứng. Chi tiết hơn về DirAC được mô tả trong [9], lý thuyết cơ bản trong [5].

Bản chất kỹ thuật của sáng chế

Có mong muốn nâng cao hướng tính của các tín hiệu Ambisonic bậc 1, như các bản ghi micro định dạng B. Việc nâng cao hướng tính này được mong muốn để có bản phát lại thực tế hơn, hoặc để trộn âm thanh được ghi thực với nội dung bậc cao hơn khác, ví dụ để lồng âm thanh phim mà nhằm mục đích được phát lại cho các thiết lập loa khác nhau. Một vấn đề cần được giải quyết bởi sáng chế là nâng cao hướng tính của các tín hiệu Ambisonic bậc 1 hoặc các tín hiệu định dạng B, ngay cả nếu các hệ số bậc cao hơn đối với các tín hiệu này không khả dụng.

Theo sáng chế, các vấn đề này và các vấn đề khác có thể được giải quyết bằng cách khuếch đại có chọn lọc các thành phần âm thanh trực tiếp, trong khi các thành phần âm thanh khuếch tán không được thay đổi. Khi khuếch đại có chọn lọc âm thanh trực tiếp, có ưu điểm là thu được tín hiệu có định dạng Ambisonic với bậc được gia tăng, do nó có thể dễ dàng được kết hợp với các tín hiệu có định dạng Ambisonic khác. Với sáng chế, có thể gia tăng bậc của tín hiệu Ambisonic bậc 1, nhờ đó chỉ các thành phần âm thanh trực tiếp được xem xét. Điều này lần nữa tạo ra tín hiệu có định dạng Ambisonic, nhưng với bậc cao hơn (tức là ít nhất bậc 2). Về nguyên tắc, phương pháp nâng cao hướng tính của tín hiệu Ambisonic bậc 1 được bộc lộ thu được các hệ số bậc cao hơn từ thông tin hệ số bậc 1 và bổ sung các hệ số bậc cao hơn thu được vào tín hiệu Ambisonic. Do đó, thông tin hệ số bậc 1 (tức là các hệ số bậc 0 và bậc 1) của tín hiệu Ambisonic bậc 1 được duy trì hiệu quả (ngoại trừ việc tạo định dạng lại, theo một phương án).

Nói cách khác, thông tin hướng tính bổ sung được trích từ tín hiệu Ambisonic bậc thấp hơn, và thông tin hướng tính bổ sung được sử dụng để đánh giá các hệ số bậc cao hơn. Theo cách này, hướng tính của tín hiệu Ambisonic được nâng cao, mà dẫn đến nâng cao độ chính xác của việc định vị nguồn không gian khi tín hiệu Ambisonic được giải mã thành các tín hiệu loa. Một hiệu quả của sáng chế đó là tín hiệu đầu ra thu được có năng lượng cao hơn so với tín hiệu đầu vào.

Sáng chế đề cập đến phương pháp nâng cao hướng tính của của tín hiệu đầu vào mà là tín hiệu Ambisonic bậc 1 và có các hệ số bậc 0 và bậc 1, như được xác định trong điểm 1.

Sáng chế cũng đề cập đến thiết bị nâng cao hướng tính của tín hiệu Ambisonic bậc 1 có các hệ số bậc 0 và bậc 1, như được xác định trong điểm 9.

Ngoài ra, sáng chế đề cập đến vật ghi đọc được bởi máy tính lưu trữ trên đó các lệnh đọc được bởi máy tính mà, khi được thực hiện trên máy tính, làm cho máy tính thực hiện phương pháp nâng cao hướng tính của tín hiệu Ambisonic bậc 1 có các hệ số bậc 0 và bậc 1 như được xác định trong điểm 1.

Lưu ý rằng các tín hiệu Ambisonic có bậc bất kỳ nói chung bao gồm các hệ số không chỉ có bậc định trước, mà còn các hệ số có tất cả bậc thấp hơn, ngay cả nếu không được đề cập cụ thể. Ví dụ, tín hiệu HOA bậc 2 bao gồm các hệ số không chỉ có bậc 2, mà còn có bậc 0 và bậc 1.

Các phương án có lợi của sáng chế được bộc lộ trong các điểm yêu cầu bảo hộ phụ thuộc, phần mô tả sau đây và các hình vẽ.

Mô tả vắn tắt các hình vẽ

Các phương án làm ví dụ của sáng chế được mô tả có viện dẫn đến các hình vẽ kèm theo, mà thể hiện trên

Fig.1a) cấu trúc của bộ giải mã định dạng B dựa trên DirAC đã biết;

Fig.1b) cấu trúc tổng quát của khối phân tích trường âm thanh đã biết;

Fig.2 cấu trúc của thiết bị theo phương án tổng quát của sáng chế;

Fig.3 cấu trúc của thiết bị theo phương án mà sử dụng việc kết hợp trong miền thời gian;

Fig.4 cấu trúc của thiết bị theo phương án thứ nhất mà sử dụng việc kết hợp trong miền tần số;

Fig.5 cấu trúc của thiết bị theo phương án thứ hai mà sử dụng việc kết hợp trong miền tần số;

Fig.6 lưu đồ của phương pháp theo sáng chế; và

Fig.7 lưu đồ chi tiết của bước kết hợp.

Mô tả chi tiết sáng chế

Fig.2 thể hiện cấu trúc của thiết bị theo phương án tổng quát của sáng chế. Tín hiệu đầu vào Ambisonic bậc 1 miền thời gian 10 (như tín hiệu định dạng B) được lọc trong băng lọc phân tích (Analysis Filter Bank- AFB), trong đó thu được bốn kênh miền tần số 21. Đó là dạng biểu diễn miền tần số của tín hiệu đầu vào 10: một trong các kênh miền tần số biểu diễn các hệ số bậc 0 (tức là kênh W), và ba kênh miền tần số khác biểu diễn các hệ số bậc 1 (các kênh X,Y,Z).

Bộ tách âm thanh trực tiếp DSS tách âm thanh trực tiếp (tức là âm thanh có hướng) 20 trong bốn kênh miền tần số 21 từ âm thanh khuếch tán. Theo một phương án, bộ tách âm thanh trực tiếp DSS đơn giản là lựa chọn kênh W và sử dụng nó như là âm thanh trực tiếp 20. Ngoài ra, bộ phân tích trường âm thanh SFA thực hiện việc phân tích trường âm thanh của bốn kênh miền tần số, thu được các hướng nguồn Θ, Φ 22 và đánh giá khuếch tán Ψ 23 đối với mọi băng tần số của các kênh tần số. Theo một phương án, bộ phân tích trường âm thanh SFA bao gồm bộ phân tích hướng đi đến (Direction of Arrival - DoA) để thu được thông tin hướng 22.

Âm thanh trực tiếp 20 thu được bởi bộ tách âm thanh trực tiếp DSS sau đó được lọc trong bộ lọc F, nhờ đó các thành phần khuếch tán được làm tắt dần và do đó âm thanh có hướng được khuếch đại (một cách tương đối) có chọn lọc. Bộ lọc F sử dụng đánh giá khuếch tán Ψ 23 cho việc khuếch đại có chọn lọc; về nguyên tắc, nó nhân âm thanh trực tiếp 20 với $\sqrt{2(1 - \Psi(f))}$ để thu được âm thanh trực tiếp được khuếch đại có chọn lọc 24. Âm thanh trực tiếp được khuếch đại có chọn lọc 24 sau đó được mã hóa Ambisonic trong bộ mã hóa HOA HOAe, trong đó thu được tín hiệu HOA 25 có bậc được xác định trước N_0 ($N_0 > 1$, tức là ít nhất bậc 2). Bộ mã hóa HOA HOAe sử dụng các hướng nguồn Θ, Φ 22 cho việc mã hóa. Nó có thể sử dụng định dạng Ambisonic mà có bậc 0 và các hệ số bậc 1 theo định dạng B. Thay vào đó, nó cũng có thể sử dụng định dạng Ambisonic khác. Các định dạng Ambisonic khác thường có bậc hệ số tuần tự được xác định mà khác với bậc tuần tự của định dạng B, hoặc việc định tỷ lệ hệ số mà khác với việc định tỷ lệ hệ số của định dạng B, hoặc cả hai.

Bộ lựa chọn SEL lựa chọn các phần được xác định của tín hiệu HOA 25, và các phần được lựa chọn 25a sau đó được kết hợp trong bộ kết hợp và tổng hợp CS với

tín hiệu định dạng B gốc. Các phần được lựa chọn 25a là các phần bậc cao hơn của tín hiệu HOA 25, tức là các phần (các hệ số, theo một phương án) có ít nhất bậc 2. Bộ kết hợp và tổng hợp CS cung cấp các tín hiệu miền thời gian đầu ra của nó 29 (ở định dạng HOA), mà có thể được sử dụng để kết xuất các tín hiệu loa. Bộ kết hợp và tổng hợp CS bao gồm bộ lọc tổng hợp SF để lọc các tín hiệu có định dạng Ambisonic và thu được các tín hiệu miền thời gian.

Fig.2 cũng thể hiện bộ trộn MX bổ sung tùy chọn, trong đó tín hiệu đầu ra HOA 29 thu được có thể được trộn với tín hiệu HOA đầu vào 30 khác có bậc cao hơn. Tín hiệu đầu vào HOA 30 khác cũng có thể có định dạng Ambisonic khác so với tín hiệu đầu vào 10, do bộ thích ứng định dạng HOA HFA được mô tả dưới đây. Bộ trộn MX tạo ra tín hiệu HOA 31 mà bao gồm hỗn hợp tín hiệu đầu ra HOA 29 thu được (tức là tín hiệu đầu vào định dạng B được nâng cao) và tín hiệu đầu vào HOA 30.

Hai loại phương án cơ bản của bộ kết hợp và tổng hợp CS được mô tả sau đây: Theo một loại phương án, bộ kết hợp và tổng hợp CS kết hợp các phần được lựa chọn 25a với tín hiệu định dạng B gốc 10 trong miền thời gian. Do đó, nó thực hiện tổng hợp chỉ các phần được lựa chọn 25a trong miền thời gian. Theo loại phương án khác, bộ kết hợp và tổng hợp CS kết hợp các phần được lựa chọn 25a với tín hiệu định dạng B gốc 10 trong miền tần số, và thực hiện việc tổng hợp trong miền thời gian sau đó.

Fig.3 thể hiện phương án thuộc loại thứ nhất. Trong phương án này, bộ kết hợp và tổng hợp CS tổng hợp chỉ các hệ số bậc cao hơn được lựa chọn 25a của tín hiệu HOA 25 trong Bảng lọc tổng hợp SFB để thu được tín hiệu miền thời gian được tổng hợp 26. Bộ kết hợp miền thời gian CB_t kết hợp tín hiệu miền thời gian được tổng hợp 26 với tín hiệu đầu vào trong miền thời gian, để thu được tín hiệu đầu ra miền thời gian 29. Theo một phương án, bộ thích ứng định dạng HOA miền thời gian HFA_t làm thích ứng định dạng của tín hiệu đầu vào miền thời gian theo định dạng mà bộ mã hóa HOA HOA_e sử dụng. Điều này làm đơn giản hóa việc kết hợp tín hiệu HOA miền thời gian 28 thu được với tín hiệu miền thời gian được tổng hợp 26 trong bộ kết hợp miền thời gian CB_t . Trong một vài phương án, ví dụ trong đó bộ mã hóa HOA HOA_e sử dụng định dạng mà có thể tương thích với tín hiệu đầu vào HOA, bộ thích

ứng định dạng HOA HFA_t có thể không được yêu cầu. Bộ thích ứng định dạng HOA HFA_t có thể sắp xếp lại và/hoặc định tỷ lệ lại các hệ số của tín hiệu HOA.

Băng lọc phân tích AFB thu được các băng tần số khác nhau, ví dụ bằng cách thực hiện FFT (Fast Fourier Transform – Biến đổi Fourier nhanh). Điều này tạo ra độ trễ thời gian. Theo một phương án, bộ bù độ trễ DC của tín hiệu đầu vào miền thời gian bù độ trễ băng lọc, ví dụ của băng lọc phân tích AFB, bộ lọc khuếch đại có chọn lọc F v.v.. Trong khi trong phương án được thể hiện việc bù độ trễ được thực hiện trước khi làm thích ứng định dạng HOA HFA , trong phương án khác nó có thể được thực hiện sau khi làm thích ứng định dạng HOA. Trong phương án khác, việc bù độ trễ được thực hiện trong hai bước, với một bộ bù độ trễ trước khi làm thích ứng định dạng và một bộ bù độ trễ khác sau khi làm thích ứng định dạng.

Các hình vẽ Fig.4 và Fig.5 thể hiện các phương án mà sử dụng loại thứ hai của bộ kết hợp và tổng hợp CS. Trong phương án này, bộ kết hợp và tổng hợp CS thu các hệ số Ambisonic bậc 0 và bậc 1 miền tần số của tín hiệu đầu vào, như thu được từ băng lọc phân tích. Đây có thể là Băng lọc phân tích AFB' riêng biệt, như trong phương án được thể hiện trên Fig.4, hoặc nó có thể là Băng lọc phân tích AFB được đề cập trên đây, như trong phương án được thể hiện trên Fig.5. Trong trường hợp thứ hai, bốn kênh miền tần số 21 được cấp bởi băng lọc phân tích AFB được đưa vào trực tiếp bộ kết hợp và tổng hợp CS. Bộ kết hợp miền tần số CB_f kết hợp các hệ số bậc cao hơn được lựa chọn 25a của tín hiệu HOA 25 với các hệ số Ambisonic bậc 0 hoặc bậc 1 của tín hiệu đầu vào trong miền tần số. Băng lọc tổng hợp SFB' tổng hợp các hệ số Ambisonic được kết hợp, trong đó thu được tín hiệu đầu ra miền thời gian 29. Theo một phương án, việc làm thích ứng định dạng HOA miền tần số HFA_f tùy chọn được thực hiện trên các hệ số Ambisonic bậc 0 hoặc bậc 1 của tín hiệu đầu vào, trước khi kết hợp chúng với các hệ số bậc cao hơn được lựa chọn của tín hiệu HOA 25. Bộ thích ứng định dạng HOA HFA_f có thể sắp xếp lại và/hoặc định tỷ lệ lại các hệ số của tín hiệu HOA. Như được đề cập trên đây, bộ thích ứng định dạng HOA HFA_f có thể không được yêu cầu trong một vài phương án. Ngoài ra, cũng như được đề cập trên đây, việc bù độ trễ (không được thể hiện) có thể được sử dụng theo một phương án đối với độ trễ bất kỳ được chèn trong chuỗi xử lý (ví dụ bộ lọc khuếch đại lựa chọn F,

bộ mã hóa HOA HOAe). Tuy nhiên, điều này sẽ không thường xuyên được yêu cầu, do độ trễ được chèn bởi băng lọc phân tích AFB, AFB' không cần được bù.

Bộ kết hợp miền thời gian CB_t là bộ kết hợp mà hoạt động trong miền thời gian, trong khi bộ kết hợp miền tần số CB_f là bộ kết hợp mà hoạt động trong miền tần số. Cả hai loại bộ kết hợp thêm các hệ số thu được của các phần được lựa chọn 25a vào các hệ số (có thể được tạo định dạng lại) của tín hiệu đầu vào 10.

Nói chung, thiết bị nâng cao hướng tính của tín hiệu Ambisonic bậc 1 miền thời gian có các hệ số bậc 0 và bậc 1 bao gồm:

băng lọc phân tích AFB để lọc tín hiệu Ambisonic bậc 1, trong đó thu được bốn kênh miền tần số 21 mà là dạng biểu diễn miền tần số của tín hiệu Ambisonic bậc 1, và trong đó một kênh miền tần số 20 trong số các kênh miền tần số biểu diễn các hệ số bậc 0 và ba kênh miền tần số biểu diễn các hệ số bậc 1, bộ phân tích trường âm thanh SFA để thực hiện việc phân tích trường âm thanh của bốn kênh miền tần số, nhờ đó thu được các hướng nguồn $\Theta, \Phi 22$ và đánh giá khuếch tán $\Psi 23$, bộ lọc khuếch đại có chọn lọc F để lọc kênh miền tần số 20 mà có các hệ số bậc 0, trong đó đánh giá khuếch tán $\Psi 23$ được sử dụng và trong đó thu được thành phần âm thanh trực tiếp 24, bộ mã hóa Ambisonic bậc cao HOAe để mã hóa thành phần âm thanh trực tiếp 24 ở định dạng Ambisonic với bậc ít nhất bằng hai được xác định trước, trong đó các hướng nguồn nêu trên $\Theta, \Phi 22$ được sử dụng và trong đó thu được âm thanh trực tiếp được mã hóa 25 ở định dạng Ambisonic có bậc được xác định trước, âm thanh trực tiếp được mã hóa ở định dạng Ambisonic có các hệ số Ambisonic có ít nhất bậc 0, 1, và 2, bộ lựa chọn SEL để lựa chọn, từ âm thanh trực tiếp được mã hóa 25 thu được ở định dạng Ambisonic có bậc được xác định trước, các hệ số Ambisonic 25a có ít nhất bậc 2, và bộ kết hợp và tổng hợp CS để kết hợp các hệ số Ambisonic được lựa chọn có ít nhất bậc 2 của âm thanh trực tiếp được mã hóa 25a với các hệ số Ambisonic của tín hiệu đầu vào Ambisonic bậc 1 10, trong đó thu được dạng biểu diễn miền thời gian của tín hiệu Ambisonic có ít nhất bậc 2 29. Lưu ý rằng các hệ số Ambisonic được lựa chọn 25a có ít nhất bậc 2 không bao gồm các hệ số bậc 0 hoặc 1. Tức là, bộ lựa chọn SEL bỏ qua các hệ số bậc thấp hơn.

Theo một phương án, sáng chế đề cập đến phương pháp nâng cao hướng tính của tín hiệu Ambisonic bậc 1 10 (tức là tín hiệu Ambisonic với chỉ các hệ số bậc 0 và

các hệ số bậc 1). Nói chung, phương pháp này bao gồm các bước tạo ra, trong bộ phân tích trường âm thanh SFA, đánh giá khuếch tán Ψ_{23} và thông tin hướng Θ, Φ_{22} từ tín hiệu Ambisonic bậc 1, tách biệt và khuếch đại có chọn lọc âm thanh trực tiếp 24 từ tín hiệu Ambisonic bậc 1, trong đó bộ lọc F để khuếch đại có chọn lọc sử dụng đánh giá khuếch tán Ψ_{23} , mã hóa âm thanh trực tiếp được khuếch đại có chọn lọc 24 trong bộ mã hóa HOA HOAe, trong đó thông tin hướng Θ, Φ_{22} được sử dụng và thu được tín hiệu HOA 25 có ít nhất bậc 2, lựa chọn phần bậc cao hơn của tín hiệu HOA25, trong đó phần bậc cao hơn được lựa chọn bao gồm chỉ các hệ số có bậc cao hơn bậc 1 (tức là không bao gồm các hệ số bậc 0, và không bao gồm các hệ số bậc 1), và kết hợp các hệ số bậc cao hơn được lựa chọn của tín hiệu HOA 25 với tín hiệu Ambisonic bậc 1 đầu vào trong bộ kết hợp và tổng hợp CS, trong đó thu được dạng biểu diễn miền thời gian 29 của tín hiệu Ambisonic bậc cao hơn (tức là tín hiệu Ambisonic có ít nhất bậc 2).

Theo một phương án, bước kết hợp các hệ số bậc cao hơn được lựa chọn của tín hiệu HOA 25 với tín hiệu Ambisonic bậc 1 đầu vào 10 bao gồm thu các hệ số Ambisonic bậc 0 và bậc 1 miền tần số của tín hiệu đầu vào từ bảng lọc phân tích AFB, kết hợp các hệ số bậc cao hơn (tức là bậc 2 hoặc cao hơn) được lựa chọn 25a của tín hiệu HOA 25 với các hệ số Ambisonic bậc 0 hoặc bậc 1 của tín hiệu đầu vào trong miền tần số, và tổng hợp các hệ số Ambisonic được kết hợp trong Bảng lọc tổng hợp SFB để thu được tín hiệu đầu ra miền thời gian 29.

Theo một phương án, phương pháp này còn bao gồm bước thực hiện việc làm thích ứng định dạng HOA miền tần số HFA_f trên các hệ số Ambisonic bậc 0 hoặc bậc 1 của tín hiệu đầu vào trước khi kết hợp chúng với các hệ số bậc cao hơn được lựa chọn của tín hiệu HOA 25.

Theo phương án khác, bước kết hợp các hệ số bậc cao hơn được lựa chọn 25a của tín hiệu HOA 25 với tín hiệu Ambisonic bậc 1 đầu vào 10 bao gồm các bước tổng hợp chỉ các hệ số bậc cao hơn được lựa chọn 25a của tín hiệu HOA 25 trong Bảng lọc tổng hợp SFB để thu được tín hiệu miền thời gian được tổng hợp 26, và kết hợp tín hiệu miền thời gian được tổng hợp thu được với tín hiệu đầu vào trong miền thời gian để thu được tín hiệu đầu ra miền thời gian 29. Theo một phương án, việc làm thích ứng định dạng HOA miền thời gian HFA_t của tín hiệu đầu vào miền thời gian được

thực hiện trước khi kết hợp. Theo phương án khác, việc bù độ trễ DC của tín hiệu đầu vào miền thời gian để bù độ trễ băng lọc được thực hiện trước bước kết hợp.

Các hệ số bậc cao hơn thu được bằng cách lọc tín hiệu đầu vào Ambisonic bậc 1 10 trong băng lọc phân tích AFB, thực hiện phân tích hướng đi đến (Direction of Arrival - DoA) của tín hiệu được lọc, nhờ đó thu được đánh giá khuếch tán Ψ 23 và các hướng Θ, Φ 22, lọc kênh W (các hệ số bậc 0) nhờ sử dụng đánh giá khuếch tán Ψ 23, nhờ đó âm thanh trực tiếp $S(f)$ 20 được tách biệt, và mã hóa âm thanh trực tiếp $S(f)$ 20 ở định dạng Ambisonic trong bộ mã hóa Ambisonic bậc cao HOAe. Từ tín hiệu HOA 25 thu được, chỉ các hệ số bậc cao hơn được sử dụng, được kết hợp với các hệ số bậc thấp hơn của tín hiệu đầu vào, và kết quả tín hiệu đầu ra Ambisonic 29 được tổng hợp.

Nói chung, bước kết hợp các hệ số bậc cao hơn được lựa chọn 25a của tín hiệu HOA 25 với tín hiệu Ambisonic bậc 1 đầu vào 10 bao gồm bổ sung các hệ số tương ứng của chúng, tức là tín hiệu đầu ra 29 bao gồm tất cả hệ số của tín hiệu đầu vào 10 và các hệ số bổ sung, cụ thể là các hệ số bậc cao hơn của phần được lựa chọn 25a.

Fig.6 thể hiện lưu đồ của phương pháp theo một phương án của sáng chế. Phương pháp 60 để nâng cao hướng tính của tín hiệu đầu vào 10 (tín hiệu Ambisonic bậc 1 có các hệ số bậc 0 và bậc 1) bao gồm các bước:

lọc s1 tín hiệu đầu vào 10, trong đó thu được bốn kênh miền tần số 21, một trong số chúng là kênh W Ambisonic 20,

thực hiện s2 phân tích trường âm thanh SFA của bốn kênh miền tần số 21, nhờ đó thu được các hướng nguồn 22 và đánh giá khuếch tán 23,

lựa chọn và lọc s3 kênh W Ambisonic miền tần số 20, trong đó đánh giá khuếch tán 23 được sử dụng và trong đó thu được thành phần âm thanh trực tiếp 24 của tín hiệu đầu vào 10,

mã hóa s4 trong bộ mã hóa Ambisonic bậc cao HOAe thành phần âm thanh trực tiếp 24 ở định dạng Ambisonic với bậc định trước N_o , trong đó các hướng nguồn nêu trên 22 được sử dụng và trong đó thu được âm thanh trực tiếp được mã hóa 25 ở định dạng Ambisonic có bậc được xác định trước N_o ,

lựa chọn s5, từ âm thanh trực tiếp được mã hóa thu được ở định dạng Ambisonic 25, các phần được xác định 25a bao gồm các hệ số Ambisonic có ít nhất bậc 2 (tức là bậc 2 hoặc bậc cao hơn, bỏ qua các bậc thấp hơn), và

kết hợp s6 tín hiệu biểu diễn các hệ số Ambisonic có ít nhất bậc 2 của các phần được lựa chọn của âm thanh trực tiếp được mã hóa 25a với tín hiệu biểu diễn tín hiệu đầu vào 10, trong đó thu được tín hiệu Ambisonic có ít nhất bậc 2 29.

Bốn kênh miền tần số 21 thu được ở bước lọc s1 là dạng biểu diễn miền tần số của tín hiệu Ambisonic bậc 1, trong đó một kênh miền tần số thứ nhất (kênh W) 20 trong số các kênh miền tần số 21 biểu diễn các hệ số bậc 0, trong khi ba kênh miền tần số còn lại 21 (các kênh X,Y,Z) biểu diễn các hệ số bậc 1.

Ở bước mã hóa s4, bộ mã hóa Ambisonic bậc cao HOAe mã hóa thành phần âm thanh trực tiếp 24 ở định dạng Ambisonic với bậc định trước N_0 , nhờ sử dụng các hướng nguồn nêu trên Θ, Φ , trong đó bậc định trước N_0 ít nhất là bậc hai và âm thanh trực tiếp được mã hóa ở định dạng Ambisonic có bậc được xác định trước có các hệ số Ambisonic có ít nhất bậc 2.

Fig.7a) thể hiện phương án trong đó bước kết hợp s6 sử dụng bốn kênh miền tần số 21 như là dạng biểu diễn của tín hiệu đầu vào 10 (tương ứng với thiết bị được thể hiện trên các Fig.4,5). Bước này bao gồm các bước kết hợp s61 trong bộ kết hợp miền tần số CB_f các hệ số Ambisonic của tín hiệu Ambisonic bậc 1 10, được biểu diễn bởi các hệ số của bốn kênh miền tần số 21,21',28, với các hệ số tần số được lựa chọn 25a của tín hiệu Ambisonic bậc cao hơn nâng cao 25 có ít nhất bậc 2, trong đó tín hiệu 37 thu được mà là dạng biểu diễn miền tần số của tín hiệu Ambisonic có ít nhất bậc 2 và có hướng tính được nâng cao so với tín hiệu đầu vào Ambisonic bậc 1 10, và

lọc s64 trong Bảng lọc tổng hợp SFB' tín hiệu thu được 37, trong đó thu được dạng biểu diễn miền thời gian của tín hiệu Ambisonic bậc cao hơn nâng cao mà có các hệ số có ít nhất bậc 2.

Fig.7 b) thể hiện phương án trong đó bước kết hợp s6 sử dụng các hệ số miền thời gian của tín hiệu đầu vào 10 (tương ứng với thiết bị được thể hiện trên Fig.3). Bước này bao gồm các bước lọc s62 trong Bảng lọc tổng hợp SFB các hệ số Ambisonic được lựa chọn có ít nhất bậc 2 25a từ âm thanh trực tiếp được mã hóa 25,

trong đó dạng biểu diễn miền thời gian của tín hiệu Ambisonic bậc cao hơn nâng cao 26 thu được mà có các hệ số có ít nhất bậc 2, và

kết hợp s65 trong bộ kết hợp miền thời gian CB_t các hệ số Ambisonic của tín hiệu Ambisonic bậc 1 10 (hoặc các hệ số khác biểu diễn các hệ số Ambisonic của tín hiệu Ambisonic bậc 1 10, do định dạng HOA thực tế có thể được làm thích ứng) với dạng biểu diễn miền thời gian của tín hiệu Ambisonic bậc cao hơn nâng cao nêu trên có ít nhất bậc 2 26, trong đó thu được dạng biểu diễn miền thời gian của tín hiệu Ambisonic có ít nhất bậc 2 29 mà có hướng tính được nâng cao so với tín hiệu Ambisonic bậc 1 10.

Phần mô tả sau đây mô tả chi tiết hơn về Ambisonic. Theo lý thuyết Ambisonic, cảnh âm thanh không gian được mô tả bởi các hệ số $A_n^m(k)$ của chuỗi Fourier-Bessel. Đối với âm lượng nguồn tự do, áp suất âm thanh tại vị trí quan sát (r, θ, ϕ) có thể được mô tả như là hàm của hệ tọa độ cầu của nó (bán kính r , độ nghiêng Θ , phương vị Φ và tần số không gian $k = \frac{\omega}{c} = \frac{2\pi f}{c}$ bởi

$$p(r, \theta, \phi, k) = \sum_{n=0}^N \sum_{m=-n}^n A_n^m(k) j_n(kr) Y_n^m(\theta, \phi) \quad (1)$$

trong đó, $A_n^m(k)$ là các hệ số Ambisonic; $j_n(kr)$ là các hàm Bessel cầu thuộc loại thứ nhất mà mô tả phần phụ thuộc bán kính; $Y_n^m(\theta, \phi)$ là các hàm điều hòa cầu (Spherical Harmonics - SH), mà có các giá trị thực trong thực tế. Chúng chịu trách nhiệm về các phần phụ thuộc góc. n là chỉ số bậc Ambisonic, và m là độ. Do bản chất của hàm Bessel, mà chỉ có các giá trị quan trọng đối với kr nhỏ, chuỗi tổng có thể được cắt tại bậc bất kỳ $n=N$ với độ chính xác đủ; đối với khôi phục hoàn hảo theo lý thuyết $N \rightarrow \infty$. Thông tin và chi tiết hơn có thể được xem trong [11], [6], [7], [3], [13]. Các hệ số Ambisonic A_n^m tạo thành tín hiệu Ambisonic; chúng có đơn vị vật lý của áp suất âm thanh (1Pa) và thay đổi theo thời gian. Tín hiệu A_0^0 có thể được xem là phiên bản mono của bản ghi Ambisonic. Các giá trị thực tế của các hệ số Ambisonic được xác định bởi định nghĩa của SH, chính xác hơn là phương pháp chuẩn hóa của nó. Số lượng hệ số A_n^m trong phương trình (1) được đưa ra đối với các dạng biểu diễn 2D bởi $O = 2N + 1$, và đối với các dạng biểu diễn 3D bởi $O = (N+1)^2$.

Thực tế, Ambisonic sử dụng các hàm điều hòa cầu (SH) có giá trị thực. Định nghĩa được đưa ra dưới đây, do có các công thức khác nhau và các loại phương pháp

chuẩn hóa khác nhau đối với SH mà ảnh hưởng tới các thao tác mã hóa và giải mã, tức là các giá trị của các hệ số Ambisonic. Công thức của SH có giá trị thực sử dụng các biểu thức không dấu sẽ được đưa ra sau đây:

$$Y_n^m(\theta, \phi) = \tilde{N}_{n,m} P_{n,|m|}(\cos(\theta)) \phi_m(\phi) \quad (2)$$

với $\tilde{N}_{n,m}$ hệ số chuẩn hóa (xem bảng 1), mà tương ứng với quan hệ trực giao giữa Y_n^m và Y_n^m . Tức là,

$$\begin{aligned} & \int_{\Omega \in S_2} Y_n^m(\Omega) Y_{n'}^{m'}(\Omega)^* d\Omega \\ &= \frac{\tilde{N}_{n,m}}{\sqrt{\frac{(2 - \delta_{0,m})(2n+1)(n-|m|)!}{4\pi(n+|m|)!}}} \frac{\tilde{N}_{n',m'}}{\sqrt{\frac{(2 - \delta_{0,m'})(2n'+1)(n'-|m'|)!}{4\pi(n'+|m'|)!}}} \delta_{nn'} \delta_{mm'} \end{aligned}$$

với delta Kronecker $\delta_{aa'}$ bằng 1 đối với $a = a'$, còn lại thì bằng 0. Trong phần sau đây, phương pháp chuẩn hóa trực giao được sử dụng. $P_{n,|m|}$ là các hàm Legendre được kết hợp, mà mô tả phần phụ thuộc của độ nghiêng $\cos(\theta)$. $P_{n,|m|}: [-1,1] \rightarrow \mathbb{R}, n \geq |m| \geq 0$. $P_{n,|m|}$ có thể được biểu diễn nhờ sử dụng công thức Rodrigues bởi phương trình (3) (tức là, tất cả các định nghĩa được thể hiện ở đây mà không sử dụng pha Condon-Shortley, mà việc bù của nó cho các biến có giá trị thực có thể gây ra sự không rõ ràng), nhưng có các phương pháp tính toán hiệu quả hơn để áp dụng.

$$P_{n,|m|}(x) = \frac{1}{2^n n!} (1-x^2)^{\frac{|m|}{2}} \frac{d^{n+|m|}}{dx^{n+|m|}} (x^2-1)^n \quad (3)$$

Phần phụ thuộc trên phần phương vị ϕ được đưa ra bởi:

$$\phi_m(\phi) = \begin{cases} \cos(m\phi), & m > 0 \\ 1 & m = 0 \\ \sin(|m|\phi) & m < 0 \end{cases} \quad (4)$$

Bảng 1 thể hiện các phương pháp chuẩn hóa chung được sử dụng trong Ambisonic. $\delta_{0,m}$ lấy giá trị 1 đối với $m=0$ và còn lại thì bằng 0. Tên quy ước SN3D, N3D được lấy từ [3].

$\tilde{N}_{n,m}$ phương pháp chuẩn hóa đối với SH thực			
Không được chuẩn hóa	Được chuẩn hóa Schmidt Semi, SN3D	Được chuẩn hóa 4π , N3D, Geodesy 4π	Được chuẩn hóa trực giao
$\sqrt{2 - \delta_{0,m}}$	$\sqrt{(2 - \delta_{0,m}) \frac{(n - m)!}{(n + m)!}}$	$\sqrt{(2 - \delta_{0,m}) \frac{(2n + 1)(n - m)!}{(n + m)!}}$	$\sqrt{(2 - \delta_{0,m}) \frac{(2n + 1)(n - m)!}{4\pi(n + m)!}}$

Bảng 1: Các phương pháp chuẩn hóa chung được sử dụng trong Ambisonic

Các tín hiệu được ghi bởi các micrô như SoundFieldTM được biểu diễn bằng cách sử dụng định dạng B. Kỹ thuật này được mô tả trong [2]. Có bốn tín hiệu định dạng B: Tín hiệu W mang tín hiệu tỷ lệ với áp suất âm thanh được ghi bởi micrô vô hướng, nhưng được định tỷ lệ bởi hệ số $1/\sqrt{2}$. Các tín hiệu X, Y, Z mang các tín hiệu tỷ lệ với các gradient áp suất trong ba hướng Cartesian. Bốn hệ số định dạng B W, X, Y, Z liên quan đến các hệ số HOA bậc 1 sử dụng các phương pháp chuẩn hóa N3D [3], [4] bởi $W = A_{0\ N3D}^0/\sqrt{2}$, $X = A_{1\ N3D}^1/\sqrt{3}$, $Y = A_{1\ N3D}^{-1}/\sqrt{3}$, $Z = A_{1\ N3D}^0/\sqrt{3}$ và liên quan đến các hệ số HOA sử dụng chuẩn hóa SN3D bởi $W = A_{0\ SN3D}^0/\sqrt{2}$, $X = A_{1\ SN3D}^1$, $Y = A_{1\ SN3D}^{-1}$, $Z = A_{1\ SN3D}^0$. Ngoài ra, định dạng B giả thiết mô hình mã hóa sóng phẳng mà hệ số i^n được bỏ qua trong dạng biểu diễn hệ số.

Các tín hiệu HOA cũng có thể được biểu diễn bởi các sóng phẳng. Áp suất âm thanh của sóng phẳng được đưa ra bởi [11]:

$$p(r, \theta, \phi, k) = \sum_{n=0}^N \sum_{m=-n}^n i^n P_{S_0} Y_n^m(\theta_s, \phi_s)^* j_n(kr) Y_n^m(\theta, \phi) \quad (5)$$

Nhờ sử dụng phương pháp chuẩn hóa N3D đối với các hàm điều hòa cầu, và A_n^m sẽ trở thành:

$$A_n^m(f) = 4\pi i^n P_{S_0}(f) Y_n^m(\theta(f)_s, \phi(f)_s)^* \quad (6)$$

trong đó, $P_{S_0}(f)$ là áp suất âm thanh tại gốc của hệ tọa độ tại tần số f . $\theta(f)_s, \phi(f)_s$ là các hướng (độ nghiêng, phương vị) tới nguồn (DoA), và $*$ chỉ báo liên hợp phức. Nhiều định dạng và hệ thống Ambisonic, bao gồm định dạng B và hệ thống micrô

SoundField™, giả thiết mô hình mã hóa và giải mã sóng phẳng và hệ số i^n được bỏ qua. Sau đó A_n^m trở thành:

$$A_n^m(f) = 4\pi P_{s_0}(f) Y_n^m(\Theta_s(f), \phi_s(f))^* \quad (7)$$

Như được đề cập trên đây, Fig 1b) thể hiện các khối của khối phân tích trường âm thanh SFA_D. Về nguyên tắc nó giống như khối phân tích trường âm thanh SFA của sáng chế, ngoại trừ rằng việc xem xét thời gian-tần số được tổng quát hóa được sử dụng ở đây, mà cho phép sử dụng cửa sổ thời gian bất kỳ. Tức là Phân tích trường âm thanh được đơn giản hóa thành các chuẩn hóa theo thời gian khác nhau. Việc tổng quát hóa này cho phép sử dụng băng lọc phức tùy ý. Việc tổng quát hóa khác được sử dụng ở đây đó là trường âm thanh hoạt động được tổng hợp từ sự xếp chồng của các sóng phẳng. Tất cả tham số trường âm thanh là các hàm của tần số, và chúng có thể được tính toán đối với mỗi tần số trung tâm của dải băng lọc k . Phần phụ thuộc của f_k từ k được bỏ qua trong phần mô tả chi tiết sau đây.

Cường độ hoạt động được mô tả tiếp theo.

Cường độ hoạt động $I_a(f)$ được xác định theo (xem [5]):

$$I_a = \text{Re}\{P(f)^* U(f)\} \quad (8)$$

Đơn vị của cường độ hoạt động là $\text{W/m}^2 = \text{N}/(\text{ms})$. $P(f)^*$ là áp suất âm thanh liên hợp phức (theo Pascal $= 1\text{N/m}^2$) và $U(f)$ là vận tốc hạt theo m/s vector trong ba chiều Cartesian. $\text{Re}\{\cdot\}$ ký hiệu phần thực. Các công thức khác của cường độ hoạt động sử dụng hệ số bổ sung là $\frac{1}{2}$, như trong [11], mà sau đó dẫn đến hệ số bổ sung đối với phương trình (13). Tín hiệu định dạng B W tỷ lệ với tín hiệu áp suất âm thanh $P(f)$, và các tín hiệu $\mathbf{X}(f) = [X(f), Y(f), Z(f)]^T$ tỷ lệ với vận tốc âm thanh U .

$$U(f) = \frac{1}{Z_0} (X(f)e_x + Y(f)e_y + Z(f)e_z) \quad (9)$$

trong đó, e_i là các vector đơn vị của các trục tọa độ Cartesian và e_u là chiều vector đơn vị của sóng phẳng lan truyền. Z_0 là trở kháng đặc tính (tích của tốc độ âm thanh và mật độ không khí $Z_0 = \rho_0 c$). Sau đó, cường độ hoạt động I_a có thể được biểu diễn nhờ sử dụng các tín hiệu định dạng B là (xem [5]):

$$I_a(f) = \frac{\sqrt{2}}{Z_0} \text{Re}\{W(f)^* \mathbf{X}(f)\} \quad (10)$$

trong đó, hệ số $\sqrt{2}$ biểu diễn sự định tỷ lệ của hệ số W ở định dạng B; * ký hiệu liên hợp phức. $I_a(f)$, $X(f)$ là các hàm vector của tần số trong các hệ tọa độ Cartesian.

Hướng đi đến (Direction of Arrival) được mô tả tiếp theo.

Vector đơn vị của cường độ hoạt động $e_I(f) = [e_{Ix}(f), e_{Iy}(f), e_{Iz}(f)]^T$ được đưa ra bởi:

$e_I(f) = I_a(f)/||I_a(f)||$ Góc phương vị của DoA được đưa theo radian bởi:

$$\phi(f) = \text{atan2}\left(\frac{e_{Iy}(f)}{e_{Ix}(f)}\right) \quad (11)$$

trong đó, $I_{ai}(f)$ là các thành phần Cartesian của $I_a(f)$ và atan2 là tang ngược bốn góc phần tư. Góc nâng $\theta(f)$ có thể được tính toán bởi:

$$\theta(f) = \text{atan2}\left(\frac{\sqrt{e_{Ix}(f)^2 + e_{Iy}(f)^2}}{e_{Iz}(f)}\right) \quad (12)$$

Sự khuếch tán được mô tả tiếp theo.

Mật độ năng lượng, tức là năng lượng âm thanh trên thể tích đơn vị (theo các đơn vị vật lý $\text{N/m}^2 = \text{kg m/s}^2 \text{ 1/m}^2$), của trường âm thanh được mô tả bởi [5]:

$$E(f) = \frac{\rho_0}{2} (Z_0^{-2} |P(f)|^2 + ||U(f)||^2) \quad (13)$$

trong đó, $||U||$ mô tả chuẩn tắc ma trận 2, độ dài Euclidean của các vector. Đối với các tín hiệu Ambisonic, bậc 1/ định dạng B trở thành:

$$E(f) = \frac{\rho_0}{Z_0} (|W(f)|^2 + \frac{||X(f)||^2}{2}) \quad (14)$$

Trong phần sau đây, sự phụ thuộc tần số trong ký hiệu được giảm cho khả năng đọc tốt hơn. Đánh giá khuếch tán Ψ được xác định như [5]:

$$\Psi = \frac{\mathbb{E}(E) - ||\mathbb{E}(I_a)||/c}{\mathbb{E}(E)} = 1 - \frac{||\mathbb{E}(I_a)||}{c\mathbb{E}(E)} \quad (15)$$

$\mathbb{E}$ là toán tử giá trị kỳ vọng mà có thể được áp dụng sử dụng phép lấy trung bình theo thời gian thu được bởi giá trị trung bình được tạo cửa sổ hoặc bậc 1 bởi bộ lọc IIR. Ψ biểu diễn sự đóng góp của các phần năng lượng không hoạt động của trường âm thanh. Giá trị 1 mô tả trường âm thanh khuếch tán toàn phần (không đóng góp động năng), và giá trị 0 mô tả trường âm thanh hoạt động toàn phần. Nhờ sử dụng các tín hiệu định dạng B, độ khuếch tán có thể được biểu diễn là:

$$\Psi = 1 - \frac{\sqrt{2} ||\mathbb{E}(\text{Re}\{W^* X\})||}{\mathbb{E}(|W|^2 + \frac{||X||^2}{2})} \quad (16)$$

Cách thực hiện khác về đánh giá khuếch tán [1] được đưa ra bởi

$$\Psi = \sqrt{1 - \frac{||\mathbb{E}(\mathbf{I}_a)||}{\mathbb{E}(|\mathbf{I}_a|)}} \quad (17)$$

Lọc trung bình được mô tả tiếp theo.

Đánh giá khuếch tán và các hướng DoA yêu cầu lấy trung bình theo thời gian.

Để tính xấp xỉ kỳ vọng $y \cong \mathbb{E}(x)$, đầu ra bộ lọc làm trơn được xác định bởi [12]:

$$y(n, k) = (1 - g) x(n, k) + g y(n - 1, k) \quad (18)$$

trong đó, $x(n, k)$ là đầu vào và $y(n - 1, k)$ đầu ra bị trễ mẫu (khởi biến đổi) trong băng lọc k . Tham số lọc g được đưa ra bởi: $g = \exp\left(-\frac{1}{f_c \tau}\right)$ trong đó f_c là tốc độ mẫu của băng lọc lấy mẫu con. Đối với các băng lọc dựa trên khối với 50% các cửa sổ xếp chồng, f_c trở thành $f_c = \frac{f_s}{N_{hop}}$, với kích cỡ chặng N_{hop} bằng nửa kích cỡ cửa sổ đối với trường hợp xếp chồng 50% này. Hằng số thời gian τ xác định đặc tính của bộ lấy trung bình. Giá trị nhỏ là thích hợp khi các thay đổi nhanh của các tín hiệu đầu vào cần được theo dõi, giá trị lớn thích hợp đối với lấy trung bình dài hạn.

Có các cách thực hiện khác khác, ví dụ (xem [10]):

$$y(k, n) = a x(k, n) + (1 - a) y(k, n - 1) \quad (19)$$

với $a = \frac{N_{hop}}{\tau f_s}$ và $\tau \geq \frac{N_{hop}}{f_s}$. Ở đây τ được xem trong quan hệ tuyệt đối với f_s .

Lọc thích nghi với tham số chuyển đổi phụ thuộc khối cc và hai hằng số thời gian τ_{max} , τ_{min} có thể được sử dụng cho hằng số thời gian:

$$\tau = \tau_{min} + \frac{cc_{max} - cc}{cc_{max}} (\tau_{max} - \tau_{min}) \quad (20)$$

Trong hầu hết các trường hợp, bản ghi Ambisonic bậc 1 bất kỳ sẽ là tín hiệu định dạng B. Phương pháp theo sáng chế thu được các hệ số Ambisonic bậc cao hơn đối với bản ghi Ambisonic bậc 1 đang tồn tại, trong khi duy trì thông tin hệ số bậc 1. Phân tích hướng đi đến (DoA) được thực hiện để thu được các hướng mạnh nhất trên tần số. Kênh W biểu diễn trộn mono của tất cả các tín hiệu này. Kênh W được lọc sao

cho các phần khuếch tán được loại bỏ trên tần số. Do đó, kênh W được lọc trở thành đánh giá của các âm thanh trực tiếp trên tần số. Các hướng DoA được sử dụng cho việc mã hóa Ambisonic của tín hiệu kênh W được lọc để tạo thành tín hiệu HOA mới có bậc Ambisonic được gán trước $N_order > 1$, với $O = (N_order + 1)^2$ đối với 3D và $O = (2N_order + 1)$ hệ số đối với các áp dụng 2D. Bốn hệ số của bản ghi định dạng B (tức là các tín hiệu bậc 1) được chuyển đổi định dạng thành cùng định dạng như các tín hiệu Ambisonic mới, nếu cần thiết, và được kết hợp với các hệ số mới để tạo thành tín hiệu đầu ra. Các hệ số tín hiệu HOA đầu ra thu được C_n^m được kết hợp từ các hệ số định dạng B được chuyển đổi của các thành phần bậc 0 và bậc 1 và từ các hệ số HOA mới của các thành phần bậc cao hơn.

Xử lý hoặc một phần xử lý được áp dụng trong miền tần số băng lọc của băng lọc phân tích.

Một phương án sử dụng băng lọc phân tích dựa trên FFT. Cửa sổ sin xếp chồng 50% được áp dụng cho 960 mẫu, hoặc theo cách khác cho, ví dụ, 640 hoặc 512 mẫu. Việc chèn số 0 sang trái và phải được sử dụng để thu được độ dài FFT 1024 mẫu. Băng lọc ngược (băng lọc tổng hợp) sử dụng tạo cửa sổ và bổ sung bao phủ để khôi phục khối của 480 (320, 256) mẫu. Băng lọc có thể sử dụng được khác, mà sử dụng các băng thông lọc thích hợp hơn với cảm nhận con người, được mô tả trong ISO/IEC 23003 /2007/2010 (MPEG Surround, SAOC).

Khi sử dụng băng lọc FFT, hai hoặc nhiều băng lọc có thể được kết hợp để thích ứng tốt hơn với cảm nhận con người đặc biệt đối với các tần số cao. Theo một phương án, băng thông của xấp xỉ một phần tư của tiếng quát tháo được sử dụng với độ chi tiết của một băng lọc FFT, và các giá trị trung bình của cường độ hoạt động và năng lượng trên các băng kết hợp được sử dụng. Trong các phương án khác nhau, các tham số trường âm thanh “cường độ hoạt động” và/hoặc “mật độ năng lượng” được sử dụng để thu được các góc DoA và đánh giá khuếch tán.

Theo một phương án, các bộ lọc làm tròn đặc biệt theo phương trình (18) đối với các hướng DoA và các đánh giá khuếch tán được sử dụng; sau đó, việc làm tròn của đánh giá khuếch tán được thực hiện như sau (sự phụ thuộc băng tần số được bỏ qua để cho rõ ràng):

Đánh giá khuếch tán theo phương trình (15) được cho bởi $\Psi = 1 - \frac{||E(I_a)||}{cE(E)}$. Các bộ lọc làm tron của bộ đếm $\tilde{I}_a \cong E(I_a) = [E(I_{ax}), E(I_{ay}), E(I_{az})]^T$ thu được bởi các bộ lọc IIR bậc 1 sử dụng cùng hằng số thời gian đối với ba thành phần. Ngoài ra, các bộ lọc có các hệ số kép đặc trưng bởi hằng số thời gian nhỏ $\tau_{\min}$ và hằng số thời gian lớn $\tau_{\max}$. Việc chuyển đổi giữa các hằng số thời gian được thực hiện phụ thuộc vào sự thay đổi của $||I_a||$ và bộ đếm trạng thái bổ sung cc , trong đó $I_a(n)$ là đầu vào bộ lọc và $\tilde{I}_a(n-1)$ là đầu ra bộ lọc của thao tác trước đó.

Nếu $cc=0$ và $||I_a(n)|| \leq ||\tilde{I}_a(n-1) + \epsilon_1||$ hệ số với hằng số thời gian lớn $\tau_{\max}$ được sử dụng.

Nếu $||I_a(n)|| > ||\tilde{I}_a(n-1) + \epsilon_1||$, hệ số đặc trưng bởi hằng số thời gian nhỏ $\tau_{\min}$ được sử dụng và cc được thiết lập là $cc_{\max}$ lớn hơn 1 (ví dụ $cc_{\max}=10$).

Nếu $cc > 0$ và $||I_a(n)|| \leq ||\tilde{I}_a(n-1) + \epsilon_1||$, thì hằng số thời gian $\tau = \tau_{\min} + \frac{cc_{\max}-cc}{cc_{\max}} (\tau_{\max} - \tau_{\min})$ được sử dụng và cc được làm giảm sau đó (xử lý khối) miễn là nó không bằng không.

ϵ_1 là hằng số dương. Việc làm tron năng lượng E được thực hiện theo cách tương tự, nhờ sử dụng bộ lọc riêng biệt nhưng có cùng cấu trúc bộ lọc thích nghi. Nó đặc trưng bởi $\tau_{\max}, \tau_{\min}$ và bộ đếm trạng thái cc của nó, trong đó các thay đổi $|E(n)|$ được sử dụng để chuyển đổi giữa các hằng số thời gian lớn, nhỏ và được nội suy.

$\phi(f)$, $\theta(f)$ thu được từ vector đơn vị của cường độ hoạt động $e_I(f) = I_a(f)/||I_a(f)||$ bằng cách tạo ra hai tín hiệu phức:

$$a_1 = e_{Ix} + i e_{Iy} \quad (21)$$

và

$$a_2 = \sqrt{e_{Ix}^2 + e_{Iy}^2} + i e_{Iz} \quad (22)$$

trong đó, $i = \sqrt{-1}$ và e_{Ix}, e_{Iy}, e_{Iz} là các thành phần Cartesian của vector đơn vị của cường độ hoạt động. Các tín hiệu a_1, a_2 được lọc nhờ sử dụng bộ lọc bậc nhất IIR thích nghi trên mỗi băng con theo phương trình(18):

$$b_1(n) = (1 - g(\Psi))a_1(n) + g(\Psi) b_1(n-1) \quad (23)$$

và tương tự đối với $b_2(n)$ nhờ sử dụng đầu vào $a_2(n)$ và cùng các tham số lọc $g(\Psi)$ mà phụ thuộc vào độ khuếch tán Ψ . Sự phụ thuộc có thể là tuyến tính: $g(\Psi) = (g_{max} - g_{min})\Psi + g_{min}$ với g_{min} tiến gần 0 và $g_{max} \leq 1$.

Các tín hiệu có hướng ϕ, θ có thể được tính toán từ các đầu ra bộ lọc như sau:

$$\phi = \text{atan2}\left(\frac{\text{Im}\{b_1\}}{\text{Re}\{b_1\}}\right) \quad (24)$$

$$\theta = \text{atan2}\left(\frac{\text{Re}\{b_2\}}{\text{Im}\{b_2\}}\right) \quad (25)$$

Các phương án được mô tả nêu trên có viện dẫn tới các hình vẽ từ Fig.2 đến Fig.5 sử dụng bốn bộ lọc phân tích đối với các hệ số bậc 1 được ký hiệu bởi W,X,Y,Z trong trường hợp định dạng B. Phân tích DoA và đánh giá khuếch tán được thực hiện nhờ sử dụng các bộ lọc làm trơn thích nghi được mô tả nêu trên trong K băng tần số với tần số trung tâm f_k . Tín hiệu hệ số W được nhân với $\sqrt{2(1-\Psi(f_k))}$ trong trường hợp của định dạng B, và trong trường hợp của các tín hiệu bậc 1 được chuẩn hóa khác với $\sqrt{(1-\Psi(f_k))}$ trong mỗi băng tần số để thu được tín hiệu S . Các hướng DoA được sử dụng để mã hóa Ambisonic tín hiệu S trong các băng tần số để tạo thành tín hiệu HOA mới có bậc Ambisonic được gán trước $N_order > 1$ với $O = (N_order + 1)^2$ đối với 3D và $O = (2N_order + 1)$ hệ số đối với các áp dụng 2D. O tín hiệu Ambisonic mới được ký hiệu bởi B_n^m . Theo một phương án, bộ mã hóa HOA sử dụng N3d hoặc các hàm điều hòa cầu được chuẩn hóa trực giao bỏ qua hệ số i'' . Phương pháp mã hóa sóng phẳng được sử dụng:

$$\mathbf{B}(f_k) = \mathbf{E}(f_k) S(f_k) \quad (26)$$

trong đó, $\mathbf{B}(f_k)$ là vector đối với mỗi băng tần số k với tần số trung tâm f_k giữ O hệ số Ambisonic $\mathbf{B}(f_k) = [B_0^0(f_k), B_1^{-1}(f_k), B_1^0(f_k), B_1^1(f_k), B_2^{-2}(f_k), \dots]^T$ và $\mathbf{E}$ là vector chế độ có kích cỡ $O \times 1$ giữ các hàm điều hòa cầu có hướng:

$$\mathbf{E}(f_k) = [Y_0^0(\theta_s(f_k), \phi_s(f_k))^*, Y_1^{-1}(\theta_s(f_k), \phi_s(f_k))^*, Y_1^0(\theta_s(f_k), \phi_s(f_k))^*, \dots]^T.$$

Bốn hệ số của tín hiệu đầu vào định dạng B, ví dụ bản ghi, được chuyển đổi định dạng thành cùng định dạng với các tín hiệu Ambisonic mới được tạo ra bởi bộ mã hóa HOA HOAe. Điều này có thể ngụ ý thích ứng với việc chuẩn hóa khác nhau của các hàm điều hòa cầu cũng như việc xem xét tùy chọn của hệ số i'' , mà đôi khi

được chứa trong các hệ số Ambisonic, và việc chuyển đổi 3D-thành-2D để thích ứng với các hàm điều hòa cầu 2D hoặc ngược lại. Các hệ số định dạng B được chuyển đổi và được sắp xếp lại được ký hiệu là $A_0^0, A_1^{-1}, A_1^0, A_1^1$ với quan hệ: $[W, Y, Z, X] \rightarrow [A_0^0, A_1^{-1}, A_1^0, A_1^1]$ và đối với 2D: $[W, Y, X] \rightarrow [A_0^0, A_1^{-1}, A_1^1]$.

Tín hiệu HOA thu được C_n^m được kết hợp từ các tín hiệu định dạng B được chuyển đổi và các hệ số HOA mới B_n^m , với các thành phần bậc 0 và bậc 1 được bỏ qua:

$C_n^m: [A_0^0, A_1^{-1}, A_1^0, A_1^1, B_2^{-2}, B_2^{-1}, B_2^0, B_2^1, B_2^2, \dots]$. Tín hiệu HOA thu được C_n^m có $O = (N_{\text{order}} + 1)^2$ thành phần đối với các áp dụng 3D, hoặc $O = (2N_{\text{order}} + 1)$ thành phần đối với các áp dụng 2D với $C_n^m: [A_0^0, A_1^{-1}, A_1^0, A_1^1, B_2^{-2}, B_2^{-1}, B_2^0, B_2^1, B_2^2, \dots]$. Thủ tục này có thể được xem là trộn lên bậc đối với các tín hiệu Ambisonic.

Phương án được mô tả nêu trên liên quan đến Fig.3 kết hợp các hệ số gốc với các hệ số mới trong miền thời gian và sử dụng $O - 4$ bộ lọc tổng hợp (lưu ý rằng đây là chữ cái “O”, không phải số 0) và độ trễ bổ sung để bù cho độ trễ băng lọc. Các phương án được thể hiện trên các hình vẽ từ Fig.4 đến Fig.5 kết hợp trong miền băng lọc và sử dụng O (không phải số 0) bộ lọc tổng hợp.

Sau khi trộn lên bậc theo sáng chế, các tín hiệu mới C_n^m có thể được sử dụng cho một vài mục đích, ví dụ trộn với nội dung Ambisonic khác của N_order để tạo thành tín hiệu D_n^m , việc giải mã C_n^m hoặc D_n^m để phát lại nhờ sử dụng L loa sử dụng bộ giải mã Ambisonic N_order , truyền và/hoặc lưu trữ C_n^m hoặc D_n^m trong cơ sở dữ liệu, v.v.. Trong một vài trường hợp, ví dụ việc truyền và/hoặc lưu trữ, siêu dữ liệu có thể được sử dụng để chỉ báo xử lý gốc và được thực hiện của tín hiệu Ambisonic.

Trong khi sáng chế thích hợp cho nâng cao hướng tính của bất kỳ tín hiệu Ambisonic bậc thấp hơn thành tín hiệu Ambisonic bậc cao hơn tương ứng, các phương án làm ví dụ được mô tả ở đây sử dụng chỉ các tín hiệu bậc 1 (định dạng B) để được nâng cao ví dụ thành các tín hiệu bậc 2. Tuy nhiên, nguyên tắc tương tự có thể được áp dụng để nâng cao tín hiệu Ambisonic có bậc cho trước thành bậc cao hơn bất kỳ, ví dụ tín hiệu bậc 2 thành tín hiệu bậc 3, tín hiệu bậc 1 thành tín hiệu bậc 4 v.v.. Nói chung, không có ý nghĩa khi tạo ra các hệ số có bậc cao hơn bậc 4.

Một ưu điểm của sáng chế đó là nó cho phép trộn các tín hiệu định dạng B (như ví dụ các bản ghi micro bậc 1) với nội dung bậc cao hơn để nâng cao độ chính xác tái tạo không gian khi giải mã tín hiệu kết hợp.

Trong khi người có hiểu biết trung bình trong lĩnh vực kỹ thuật có thể thực hiện các thao tác bỏ qua, thay thế và thay đổi khác nhau trong thiết bị và phương pháp được mô tả, dưới dạng và chi tiết của các thiết bị được bộc lộ, và trong hoạt động của chúng, điều này rõ ràng nhằm mục đích rằng tất cả các kết hợp của các thành phần mà về cơ bản thực hiện cùng chức năng theo cùng cách để đạt được cùng kết quả đều nằm trong phạm vi của sáng chế. Cần hiểu rằng sáng chế được mô tả dựa vào các ví dụ, và mỗi đặc điểm được bộc lộ trong phần mô tả và (một cách thích hợp) bộ yêu cầu bảo hộ và hình vẽ có thể được đưa ra một cách độc lập hoặc theo cách kết hợp thích hợp bất kỳ. Các đặc điểm, được áp dụng một cách thích hợp trong phần cứng, phần mềm, hoặc kết hợp của cả hai. Các số chỉ dẫn xuất hiện trong bộ yêu cầu bảo hộ chỉ nhằm mục đích minh họa và sẽ không giới hạn phạm vi của yêu cầu bảo hộ.

Danh mục tài liệu trích dẫn

- [1] Jukka Ahonen and Ville Pulkki. Diffuseness estimation using temporal variation of intensity vectors. *2009 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, October 18-21, 2009, New Paltz, NY.
- [2] Peter G. Craven and Michael A. Gerzon. Coincident microphone simulation covering three dimensional space and yielding various directional outputs, 1975.
- [3] Jérôme Daniel. Représentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimédia. PhD thesis, Université Paris 6, 2001.
- [4] Dave Malham. *Space in Music – Music in Space*. PhD thesis, University of York, April 2003.
- [5] Juha Merimaa. *Analysis, Synthesis, and Perception of Spatial Sound - Binaural Localization Modeling and Multichannel Loudspeaker Reproduction*. PhD thesis, Helsinki University of Technology, 2006.
- [6] M. A. Poletti. Three-dimensional surround sound systems based on spherical harmonics. *J. Audio Eng. Soc.*, 53(11):1004–1025, November 2005.
- [7] Mark Poletti. Unified description of Ambisonic using real and complex spherical harmonics. In *Proceedings of the Ambisonic Symposium 2009*, Graz, Austria, June 2009.
- [8] Ville Pulkki. Virtual sound source positioning using vector base amplitude panning. *J. Audio Eng. Soc.*, 45(6):456–466, June 1997.
- [9] Ville Pulkki. Spatial Sound Reproduction with Directional Audio Coding. *J. Audio Eng. Soc.*, 55(6):503–516, June 2007.
- [10] Oliver Thiergart, Giovanni Del Galdo, Magdalena Prus, and Fabian Kuech. Three-dimensional sound field analysis with directional audio coding based on signal adaptive parameter estimators. In *AES 40 TH INTERNATIONAL CONFERENCE, Tokyo, Japan, October 8–10, 2010*.
- [11] Earl G. Williams. *Fourier Acoustics*. Academic Press, 1999.

- [12] Udo Zölzer, editor. *DAFX – Digital Audio Effects*. John Wiley & Sons, 2002.
- [13] Franz Zotter. *Analysis and Synthesis of Sound Radiation with Spherical Arrays*. PhD thesis, Institute of Electronic Music and Acoustics (IEM), 2009.

YÊU CẦU BẢO HỘ

1. Phương pháp nâng cao hướng tính của tín hiệu đầu vào là tín hiệu Ambisonic bậc 1 và có các hệ số bậc 0 và bậc 1, phương pháp này bao gồm các bước:

lọc tín hiệu đầu vào trong băng lọc phân tích, trong đó thu được bốn kênh miền tần số mà là dạng biểu diễn miền tần số của tín hiệu Ambisonic bậc 1, và trong đó một kênh miền tần số thứ nhất trong số các kênh miền tần số biểu diễn các hệ số bậc 0 và ba kênh miền tần số còn lại biểu diễn các hệ số bậc 1;

thực hiện phân tích trường âm thanh của bốn kênh miền tần số, nhờ đó thu được các hướng nguồn và đánh giá khuếch tán;

lọc trong bộ lọc kênh miền tần số thứ nhất mà có các hệ số bậc 0 để xác định âm thanh trực tiếp dựa vào đánh giá khuếch tán;

mã hóa trong bộ mã hóa Ambisonic bậc cao thành phần âm thanh trực tiếp dựa vào các hướng nguồn, trong đó thành phần âm thanh trực tiếp được mã hóa ở định dạng Ambisonic với bậc được xác định trước mà có giá trị ít nhất bằng hai, trong đó thu được âm thanh trực tiếp được mã hóa ở định dạng Ambisonic có bậc được xác định trước, và trong đó âm thanh trực tiếp được mã hóa ở định dạng Ambisonic có bậc được xác định trước bao gồm các hệ số Ambisonic có bậc cao hơn bậc 1;

lựa chọn từ âm thanh trực tiếp được mã hóa thu được ở định dạng Ambisonic có bậc được xác định trước chỉ các hệ số Ambisonic có bậc 2 hoặc bậc cao hơn, trong đó các hệ số bậc 1 và bậc 0 được bỏ qua; và

kết hợp trong bộ kết hợp và tổng hợp tín hiệu biểu diễn các hệ số Ambisonic được lựa chọn có bậc 2 hoặc bậc cao hơn từ âm thanh trực tiếp được mã hóa với tín hiệu đầu vào, trong đó thu được tín hiệu Ambisonic hướng tính nâng cao có ít nhất bậc 2.

2. Phương pháp theo điểm 1, trong đó bước kết hợp các hệ số Ambisonic được lựa chọn có bậc 2 hoặc bậc cao hơn từ âm thanh trực tiếp được mã hóa với tín hiệu đầu vào bao gồm:

kết hợp trong bộ kết hợp miền tần số các hệ số Ambisonic của bốn kênh miền tần số với các hệ số tần số được lựa chọn của các hệ số Ambisonic được lựa chọn có

bậc 2 hoặc bậc cao hơn từ âm thanh trực tiếp được mã hóa, trong đó thu được tín hiệu mà là dạng biểu diễn miền tần số của tín hiệu Ambisonic có ít nhất bậc 2, và

lọc trong Bảng lọc tổng hợp tín hiệu thu được để xác định dạng biểu diễn miền thời gian của tín hiệu Ambisonic bậc cao nâng cao mà có các hệ số có ít nhất bậc 2.

3. Phương pháp theo điểm 1, trong đó bước kết hợp các hệ số Ambisonic được lựa chọn có bậc 2 hoặc bậc cao hơn từ âm thanh trực tiếp được mã hóa với tín hiệu đầu vào bao gồm:

lọc trong Bảng lọc tổng hợp các hệ số Ambisonic được lựa chọn có bậc 2 hoặc bậc cao hơn từ âm thanh trực tiếp được mã hóa, trong đó thu được dạng biểu diễn miền thời gian của tín hiệu Ambisonic bậc cao nâng cao mà bao gồm các hệ số có bậc 2 hoặc bậc cao hơn;

kết hợp trong bộ kết hợp miền thời gian các hệ số Ambisonic biểu diễn tín hiệu đầu vào với dạng biểu diễn miền thời gian của tín hiệu Ambisonic bậc cao nâng cao nêu trên có bậc 2 hoặc bậc cao hơn, trong đó thu được dạng biểu diễn miền thời gian của tín hiệu Ambisonic có ít nhất bậc 2 mà có hướng tính được nâng cao so với tín hiệu đầu vào.

4. Phương pháp theo điểm 1, trong đó ở bước mã hóa thành phần âm thanh trực tiếp ở định dạng Ambisonic với bậc được xác định trước nêu trên,, bộ mã hóa Ambisonic bậc cao sử dụng định dạng B.

5. Phương pháp theo điểm 1,

trong đó ở bước mã hóa thành phần âm thanh trực tiếp ở định dạng Ambisonic với bậc được xác định trước, bộ mã hóa Ambisonic bậc cao sử dụng định dạng Ambisonic khác ngoài định dạng B, phương pháp này còn bao gồm bước:

định dạng lại trong bộ thích ứng định dạng HOA, trước bước kết hợp nêu trên, tín hiệu đầu vào theo định dạng Ambisonic khác ngoài định dạng B để thu được các hệ số Ambisonic được định dạng lại của tín hiệu đầu vào, và trong đó, ở bước kết hợp nêu trên, bộ kết hợp kết hợp các hệ số Ambisonic được định dạng lại của tín hiệu đầu

vào với dạng biểu diễn miền thời gian của tín hiệu Ambisonic bậc cao nâng cao nêu trên có bậc 2 hoặc bậc cao hơn.

6. Phương pháp theo điểm 1, trong đó bước thực hiện phân tích trường âm thanh của bốn kênh miền tần số bao gồm:

thực hiện phân tích cường độ hoạt động của bốn kênh miền tần số, trong đó thu được giá trị biểu diễn cường độ hoạt động;

thực hiện phân tích khuếch tán của bốn kênh miền tần số, trong đó thu được đánh giá khuếch tán nêu trên; và

thực hiện phân tích hướng đi đến của giá trị biểu diễn cường độ hoạt động để thu được các hướng nguồn.

7. Phương pháp theo điểm 1, trong đó phương pháp này còn bao gồm bước trộn tín hiệu Ambisonic nâng cao có ít nhất bậc 2 với tín hiệu đầu vào HOA khác có bậc cao hơn hoặc định dạng Ambisonic khác, trong đó thu được tín hiệu HOA mà bao gồm hỗn hợp của tín hiệu đầu vào và tín hiệu đầu vào HOA khác nêu trên.

8. Phương pháp theo điểm 1, trong đó tín hiệu HOA thu được có $O = (N_{\text{order}} + 1)^2$ thành phần đối với các áp dụng 3D và $O = (2 N_{\text{order}} + 1)$ thành phần đối với các áp dụng 2D, trong đó N_{order} là bậc của bộ mã hóa HOA và tín hiệu HOA thu được có C_n^m hệ số theo C_n^m : $[A_0^0, A_1^{-1}, A_1^0, A_1^1, B_2^{-2}, B_2^{-1}, B_2^0, B_2^1, B_2^2, \dots]$, trong đó A_i^j là các hệ số của tín hiệu đầu vào và B_i^j là các hệ số HOA được lựa chọn từ âm thanh trực tiếp được mã hóa.

9. Thiết bị nâng cao hướng tính của tín hiệu đầu vào là tín hiệu Ambisonic bậc 1 và có các hệ số bậc 0 và bậc 1, thiết bị này bao gồm:

bảng lọc phân tích để lọc tín hiệu đầu vào, trong đó thu được bốn kênh miền tần số mà là dạng biểu diễn miền tần số của tín hiệu Ambisonic bậc 1, và trong đó một kênh miền tần số trong số các kênh miền tần số biểu diễn các hệ số bậc 0 và ba kênh miền tần số biểu diễn các hệ số bậc 1;

bộ phân tích trường âm thanh để thực hiện phân tích trường âm thanh của bốn kênh miền tần số, nhờ đó thu được các hướng nguồn và đánh giá khuếch tán thu được;

bộ lọc để lọc kênh miền tần số mà có các hệ số bậc 0 để xác định âm thanh trực tiếp dựa vào đánh giá khuếch tán;

bộ mã hóa Ambisonic bậc cao để mã hóa thành phần âm thanh trực tiếp dựa vào các hướng nguồn, trong đó thành phần âm thanh trực tiếp được mã hóa ở định dạng Ambisonic với bậc được xác định trước mà có giá trị ít nhất bằng hai, trong đó thu được thành phần âm thanh trực tiếp được mã hóa ở định dạng Ambisonic có bậc được xác định trước, và trong đó âm thanh trực tiếp được mã hóa ở định dạng Ambisonic có bậc được xác định trước bao gồm các hệ số Ambisonic có bậc cao hơn bậc 1;

bộ lựa chọn để lựa chọn, từ âm thanh trực tiếp được mã hóa thu được ở định dạng Ambisonic có bậc được xác định trước, chỉ các hệ số Ambisonic có bậc 2 hoặc bậc cao hơn, trong đó các hệ số bậc 1 và bậc 0 được bỏ qua; và

bộ kết hợp và tổng hợp để kết hợp tín hiệu biểu diễn các hệ số Ambisonic được lựa chọn có bậc 2 hoặc bậc cao hơn từ âm thanh trực tiếp được mã hóa với tín hiệu đầu vào, trong đó thu được tín hiệu Ambisonic hướng tính nâng cao có ít nhất bậc 2.

10. Thiết bị theo điểm 9, trong đó bộ kết hợp và tổng hợp bao gồm:

bộ kết hợp miền tần số để kết hợp các hệ số Ambisonic của bốn kênh miền tần số với các hệ số tần số được lựa chọn của các hệ số Ambisonic được lựa chọn có bậc 2 hoặc bậc cao hơn từ âm thanh trực tiếp được mã hóa, trong đó thu được tín hiệu mà là dạng biểu diễn miền tần số của tín hiệu Ambisonic có ít nhất bậc 2, và

bảng lọc tổng hợp để lọc tín hiệu thu được để xác định dạng biểu diễn miền thời gian của tín hiệu Ambisonic bậc cao nâng cao mà có các hệ số có ít nhất bậc 2.

11. Thiết bị theo điểm 9, trong đó bộ kết hợp và tổng hợp bao gồm:

bảng lọc tổng hợp để lọc các hệ số Ambisonic được lựa chọn có bậc 2 hoặc bậc cao hơn từ âm thanh trực tiếp được mã hóa, trong đó thu được dạng biểu diễn

miền thời gian của tín hiệu Ambisonic bậc cao nâng cao mà bao gồm các hệ số có bậc 2 hoặc bậc cao hơn; và

bộ kết hợp miền thời gian để kết hợp các hệ số Ambisonic biểu diễn tín hiệu đầu vào với dạng biểu diễn miền thời gian của tín hiệu Ambisonic bậc cao nâng cao nêu trên có bậc 2 hoặc bậc cao hơn, trong đó thu được dạng biểu diễn miền thời gian của tín hiệu Ambisonic có ít nhất bậc 2 mà có hướng tính được nâng cao so với tín hiệu đầu vào.

12. Thiết bị theo điểm 9, trong đó bộ mã hóa Ambisonic bậc cao nêu trên, để mã hóa thành phần âm thanh trực tiếp ở định dạng Ambisonic với bậc được xác định trước, sử dụng định dạng B.

13. Thiết bị theo điểm 9, trong đó bộ mã hóa Ambisonic bậc cao, khi mã hóa thành phần âm thanh trực tiếp ở định dạng Ambisonic với bậc được xác định trước, sử dụng định dạng Ambisonic khác ngoài định dạng B, còn bao gồm:

bộ thích ứng định dạng HOA để định dạng lại tín hiệu đầu vào theo định dạng Ambisonic khác ngoài định dạng B để thu được các hệ số Ambisonic được định dạng lại của tín hiệu đầu vào, và trong đó bộ kết hợp kết hợp các hệ số Ambisonic được định dạng lại của tín hiệu đầu vào với dạng biểu diễn miền thời gian của tín hiệu Ambisonic bậc cao nâng cao có bậc 2 hoặc bậc cao hơn.

14. Thiết bị theo điểm 9, trong đó bộ phân tích trường âm thanh bao gồm:

khối phân tích cường độ hoạt động để thực hiện phân tích cường độ hoạt động của bốn kênh miền tần số, trong đó thu được giá trị biểu diễn cường độ hoạt động;

khối phân tích khuếch tán để thực hiện phân tích khuếch tán của bốn kênh miền tần số, trong đó thu được đánh giá khuếch tán nêu trên; và

khối phân tích hướng đi đến để thực hiện phân tích hướng đi đến của giá trị biểu diễn cường độ hoạt động để thu được các hướng nguồn.

15. Thiết bị theo điểm 9, trong đó thiết bị này còn bao gồm bộ trộn để trộn tín hiệu Ambisonic nâng cao có ít nhất bậc 2 với tín hiệu đầu vào HOA khác có bậc cao hơn

hoặc định dạng Ambisonic khác, trong đó thu được tín hiệu HOA mà bao gồm hỗn hợp của tín hiệu đầu vào và tín hiệu đầu vào HOA khác nêu trên.

16. Vật ghi bất biến đọc được bằng máy tính lưu trữ trên đó các lệnh đọc được bằng máy tính mà, khi được thực thi trên máy tính, khiến cho máy tính thực hiện phương pháp nâng cao hướng tính của tín hiệu Ambisonic bậc 1 có các hệ số bậc 0 và bậc 1, phương pháp này bao gồm các bước:

lọc tín hiệu đầu vào trong băng lọc phân tích, trong đó thu được bốn kênh miền tần số mà là dạng biểu diễn miền tần số của tín hiệu Ambisonic bậc 1, và trong đó một kênh miền tần số thứ nhất trong số các kênh miền tần số biểu diễn các hệ số bậc 0 và ba kênh miền tần số còn lại biểu diễn các hệ số bậc 1;

thực hiện phân tích trường âm thanh của bốn kênh miền tần số, nhờ đó thu được các hướng nguồn và đánh giá khuếch tán;

lọc trong bộ lọc kênh miền tần số thứ nhất mà có các hệ số bậc 0 để xác định âm thanh trực tiếp dựa vào đánh giá khuếch tán;

mã hóa trong bộ mã hóa Ambisonic bậc cao thành phần âm thanh trực tiếp dựa vào các hướng nguồn, trong đó thành phần âm thanh trực tiếp được mã hóa ở định dạng Ambisonic với bậc được xác định trước mà có giá trị ít nhất bằng hai, trong đó thu được thành phần âm thanh trực tiếp được mã hóa ở định dạng Ambisonic với bậc được xác định trước, trong đó âm thanh trực tiếp được mã hóa ở định dạng Ambisonic có bậc được xác định trước bao gồm các hệ số Ambisonic có bậc cao hơn bậc 1;

lựa chọn từ âm thanh trực tiếp được mã hóa thu được ở định dạng Ambisonic có bậc được xác định trước các hệ số Ambisonic có bậc 2 hoặc bậc cao hơn, trong đó các hệ số bậc 1 và bậc 0 được bỏ qua; và

kết hợp trong bộ kết hợp và tổng hợp tín hiệu biểu diễn các hệ số Ambisonic được lựa chọn có bậc 2 hoặc bậc cao hơn từ âm thanh trực tiếp được mã hóa với tín hiệu đầu vào, trong đó thu được tín hiệu Ambisonic hướng tính nâng cao có ít nhất bậc 2.

17. Vật ghi đọc được bằng máy tính theo điểm 16, trong đó trong phương pháp này bước kết hợp các hệ số Ambisonic được lựa chọn có bậc 2 hoặc bậc cao hơn từ âm thanh trực tiếp được mã hóa với tín hiệu đầu vào bao gồm:

kết hợp trong bộ kết hợp miền tần số các hệ số Ambisonic của bốn kênh miền tần số với các hệ số tần số được lựa chọn của các hệ số Ambisonic được lựa chọn có bậc 2 hoặc bậc cao hơn từ âm thanh trực tiếp được mã hóa, trong đó thu được tín hiệu mà là dạng biểu diễn miền tần số của tín hiệu Ambisonic có ít nhất bậc 2, và

lọc trong Băng lọc tổng hợp tín hiệu thu được để xác định dạng biểu diễn miền thời gian của tín hiệu Ambisonic bậc cao nâng cao mà có các hệ số có ít nhất bậc 2.

18. Vật ghi đọc được bằng máy tính theo điểm 16, trong đó trong phương pháp này bước kết hợp các hệ số Ambisonic được lựa chọn có bậc 2 hoặc bậc cao hơn từ âm thanh trực tiếp được mã hóa với tín hiệu đầu vào bao gồm:

lọc trong Băng lọc tổng hợp các hệ số Ambisonic được lựa chọn có bậc 2 hoặc bậc cao hơn từ âm thanh trực tiếp được mã hóa, trong đó thu được dạng biểu diễn miền thời gian của tín hiệu Ambisonic bậc cao nâng cao mà bao gồm các hệ số có bậc 2 hoặc bậc cao hơn;

kết hợp trong bộ kết hợp miền thời gian các hệ số Ambisonic biểu diễn tín hiệu đầu vào với dạng biểu diễn miền thời gian của tín hiệu Ambisonic bậc cao nâng cao nêu trên có bậc 2 hoặc bậc cao hơn, trong đó thu được dạng biểu diễn miền thời gian của tín hiệu Ambisonic có ít nhất bậc 2 mà có hướng tính được nâng cao so với tín hiệu đầu vào.

19. Vật ghi đọc được bằng máy tính theo điểm 16, trong đó trong phương pháp này, ở bước mã hóa thành phần âm thanh trực tiếp ở định dạng Ambisonic với bậc được xác định trước nêu trên, bộ mã hóa Ambisonic bậc cao sử dụng định dạng B.

20. Vật ghi đọc được bằng máy tính theo điểm 16, trong đó trong phương pháp này, ở bước mã hóa thành phần âm thanh trực tiếp ở định dạng Ambisonic với bậc được xác định trước, bộ mã hóa Ambisonic bậc cao sử dụng định dạng Ambisonic khác ngoài định dạng B còn bao gồm:

định dạng lại trong bộ thích ứng định dạng HOA, trước bước kết hợp nêu trên, tín hiệu đầu vào theo định dạng Ambisonic khác ngoài định dạng B để thu được các hệ số Ambisonic được định dạng lại của tín hiệu đầu vào, và

trong đó, ở bước kết hợp nêu trên, bộ kết hợp kết hợp các hệ số Ambisonic được định dạng lại của tín hiệu đầu vào với dạng biểu diễn miền thời gian của tín hiệu Ambisonic bậc cao nâng cao nêu trên có bậc 2 hoặc bậc cao hơn.

21. Vật ghi đọc được bằng máy tính theo điểm 16, trong đó trong phương pháp này bước thực hiện phân tích trường âm thanh của bốn kênh miền tần số bao gồm:

thực hiện phân tích cường độ hoạt động của bốn kênh miền tần số, trong đó thu được giá trị biểu diễn cường độ hoạt động;

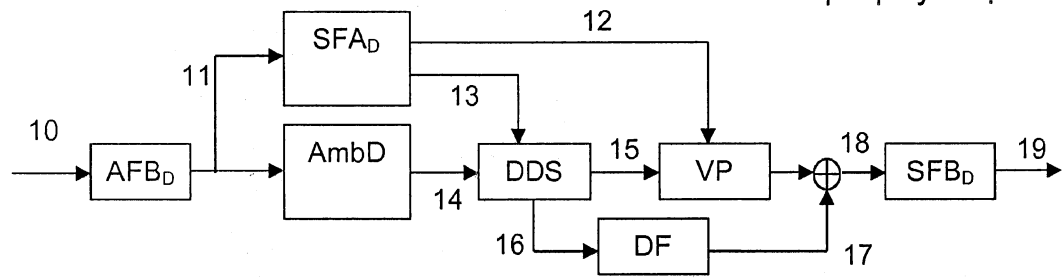
thực hiện phân tích khuếch tán của bốn kênh miền tần số, trong đó thu được đánh giá khuếch tán nêu trên; và

thực hiện phân tích hướng đi đến của giá trị biểu diễn cường độ hoạt động để thu được các hướng nguồn.

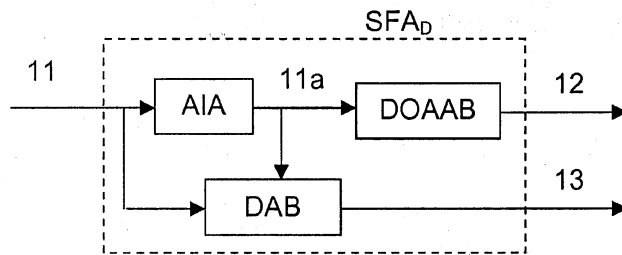
22. Vật ghi đọc được bằng máy tính theo điểm 16, trong đó trong phương pháp này bước trộn tín hiệu Ambisonic nâng cao có ít nhất bậc 2 với tín hiệu đầu vào HOA khác có bậc cao hơn hoặc định dạng Ambisonic khác, trong đó thu được tín hiệu HOA mà bao gồm hỗn hợp của tín hiệu đầu vào và tín hiệu đầu vào HOA khác nêu trên.

23. Vật ghi đọc được bằng máy tính theo điểm 16, trong đó trong phương pháp này tín hiệu HOA thu được có $O = (N_{\text{order}} + 1)^2$ thành phần đối với các áp dụng 3D và $O = (2 N_{\text{order}} + 1)$ thành phần đối với các áp dụng 2D, trong đó N_{order} là bậc của bộ mã hóa HOA và tín hiệu HOA thu được có C_n^m hệ số theo $C_n^m: [A_0^0, A_1^{-1}, A_1^0, A_1^1, B_2^{-2}, B_2^{-1}, B_2^0, B_2^1, B_2^2, \dots]$, trong đó A_i^j là các hệ số của tín hiệu đầu vào và B_i^j là các hệ số HOA được lựa chọn từ âm thanh trực tiếp được mã hóa.

Giải pháp kỹ thuật đã biết



a)



b)

Fig.1

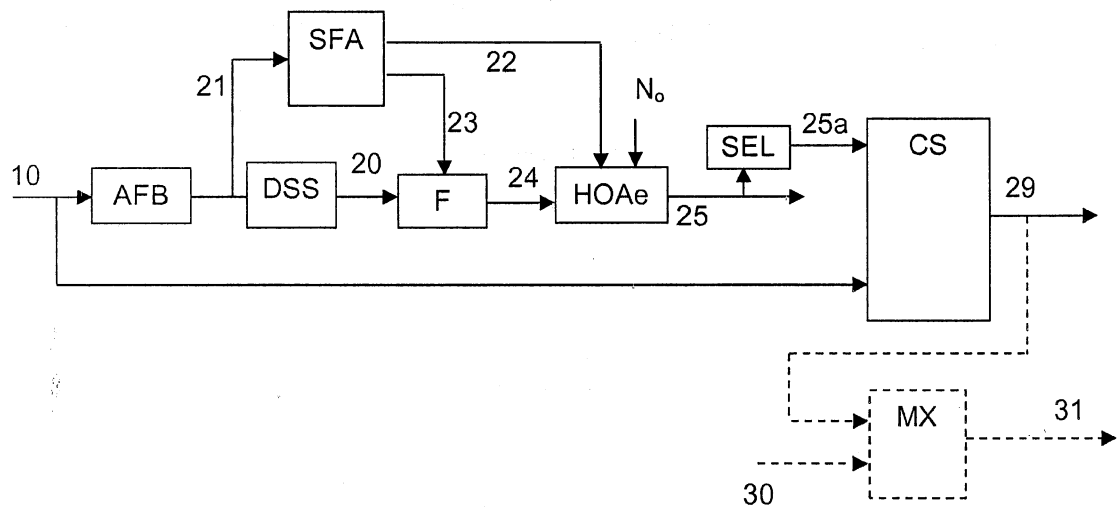


Fig.2

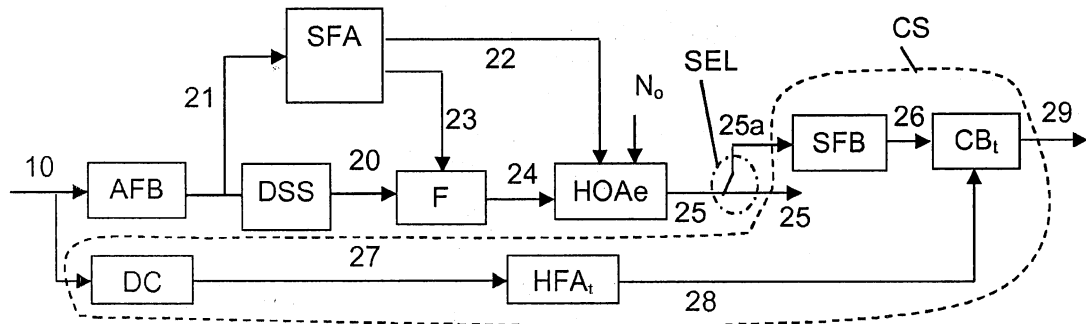


Fig.3

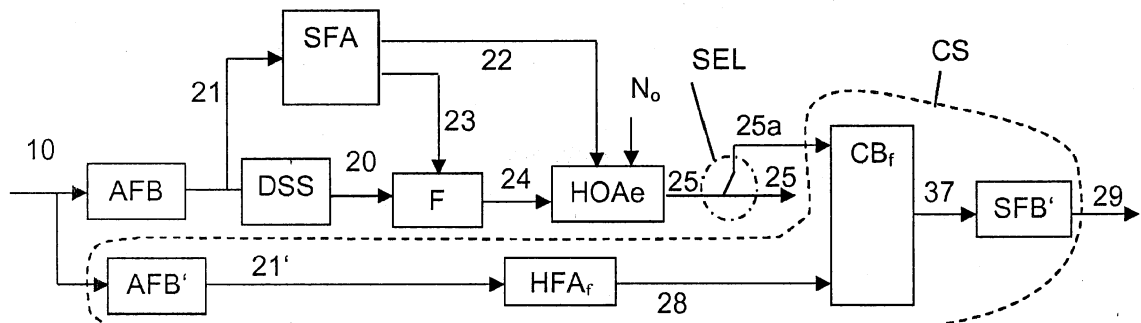


Fig.4

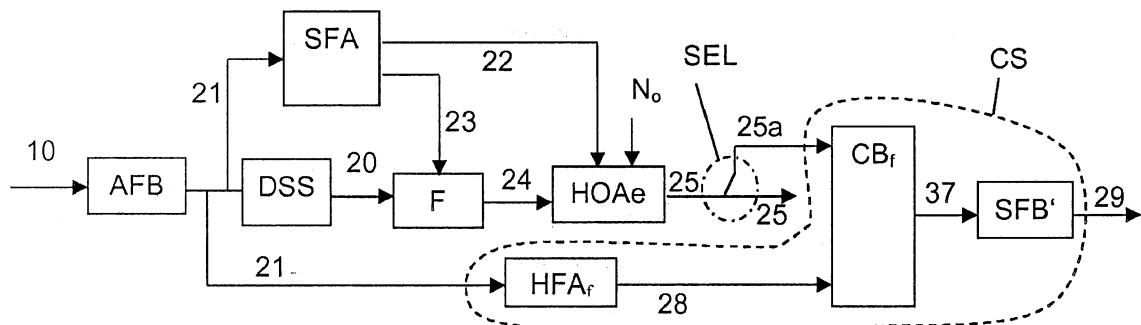


Fig.5

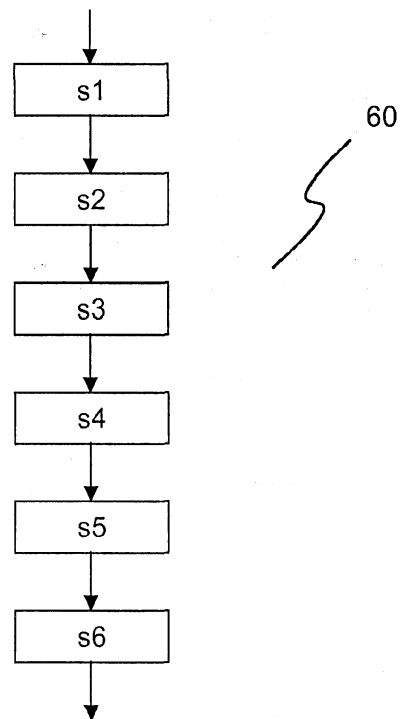


Fig.6

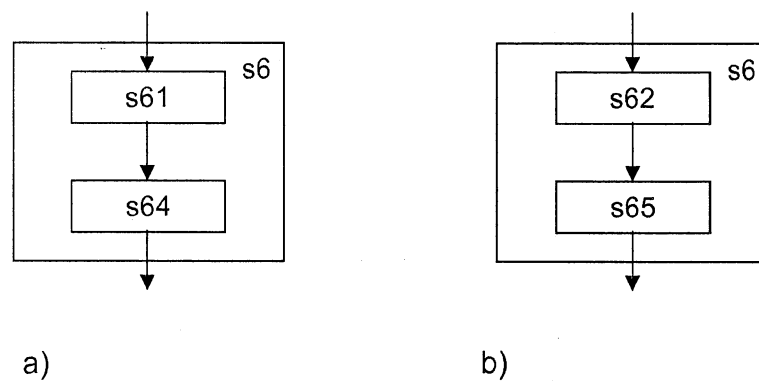


Fig.7