



(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)



1-0026219

(51)<sup>7</sup> H04S 3/02; G10L 19/00

(13) B

(21) 1-2016-03926

(22) 25/03/2011

(62) 1-2012-03174

(86) PCT/EP2011/054644 25/03/2011

(87) WO2011/117399 29/09/2011

(30) 10305316.1 26/03/2010 EP

(45) 25/11/2020 392

(43) 25/01/2013 298A

(73) DOLBY INTERNATIONAL AB (SE)

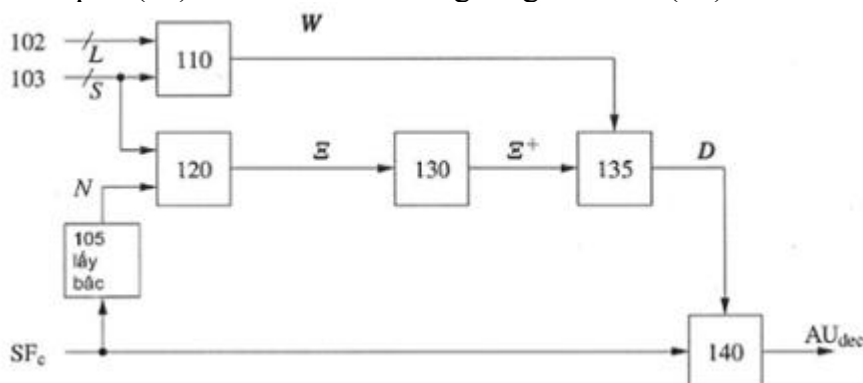
Apollo Building, 3E, Herikerbergweg 1-35, 1101 CN Amsterdam Zuidoost,  
Netherlands

(72) BATKE, Johann-Markus (DE); KEILER, Florian (DE); BOEHM, Johannes (DE).

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

(54) PHƯƠNG PHÁP VÀ THIẾT BỊ GIẢI MÃ BIỂU DIỄN TRƯỜNG ÂM THANH ĐỂ PHÁT LẠI ÂM THANH

(57) Sáng chế đề cập đến phương pháp và thiết bị giải mã biểu diễn trường âm thanh để phát lại âm thanh. Tín hiệu trường âm thanh như, ví dụ, Ambisonics mang biểu diễn của trường âm thanh mong muốn. Định dạng Ambisonics dựa vào sự phân tích hàm điều hòa cầu của trường âm thanh, và Ambisonics bậc cao hơn (Higher Order Ambisonics - HOA) sử dụng hàm điều hòa cầu ít nhất bậc 2. Tuy nhiên, cài đặt loa thường được sử dụng không theo quy tắc và dẫn đến những vấn đề về thiết kế của bộ giải mã. Phương pháp cho sự giải mã cải thiện biểu diễn trường âm thanh để phát lại âm thanh bao gồm tính (110) hàm quét (W) bằng cách sử dụng phương pháp hình học dựa vào vị trí của nhiều loa và nhiều hướng nguồn, tính (120) ma trận chế độ ( $\Xi$ ) từ vị trí của loa, tính (130) ma trận giả nghịch đảo ( $\Xi^+$ ) và giải mã (140) biểu diễn trường âm thanh. Sự giải mã dựa vào ma trận giải mã (D) mà thu được từ hàm quét (W) và ma trận chế độ giả nghịch đảo ( $\Xi^+$ ).



### **Lĩnh vực kỹ thuật được đề cập**

Sáng chế đề cập đến phương pháp và thiết bị để giải mã trường âm thanh, và cụ thể là giải mã biểu diễn âm thanh định dạng Ambisonics (tái tạo âm thanh chất lượng cao) để phát lại âm thanh.

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

Mục đích của mục này là giới thiệu cho người đọc các khía cạnh khác nhau của lĩnh vực kỹ thuật có thể liên quan đến các khía cạnh khác nhau của sáng chế mà được mô tả và/hoặc yêu cầu bảo hộ dưới đây. Tin rằng phần thảo luận này là hữu ích để cung cấp cho người đọc thông tin cơ sở để hiểu rõ hơn các khía cạnh khác nhau của sáng chế. Theo đó, cần hiểu rằng phần này phải được đọc trong ngữ cảnh của sáng chế này, và không phải là sự thừa nhận về tình trạng kỹ thuật đã biết, trừ khi nguồn tài liệu được nêu rõ ràng.

Sự định vị hóa chính xác là mục tiêu chính đối với hệ thống phát lại âm thanh không gian bất kỳ. Hệ thống phát lại này thường có tính ứng dụng cao cho các hệ thống hội thảo, trò chơi hoặc các môi trường ảo khác có sử dụng âm thanh 3D. Các lớp âm thanh 3D có thể được tổng hợp hoặc thu dưới dạng trường âm thanh tự nhiên. Tín hiệu của trường âm thanh như Ambisonics mang biểu diễn của trường âm thanh mong muốn. Định dạng Ambisonics dựa trên sự phân tích hàm điều hòa cầu của trường âm thanh. Trong khi định dạng Ambisonics cơ sở hoặc định dạng B sử dụng hàm điều hòa cầu có bậc không và bậc một thì Ambisonics bậc cao hơn (HOA) còn sử dụng thêm hàm điều hòa cầu ít nhất là bậc hai. Cần quy trình giải mã để thu được các tín hiệu loa riêng rẽ. Để tổng hợp các lớp âm thanh, cần hàm quét mà chỉ sự sắp xếp loa không gian để đạt được sự định vị hóa không gian của nguồn âm thanh cho trước. Nếu trường âm thanh tự nhiên phải được ghi lại, cần các mảng micro tích hợp để bắt thông tin không gian. Phương pháp Ambisonics đã biết là công cụ rất thích hợp để hoàn thành quy trình này. Tín hiệu được định dạng Ambisonics mang biểu diễn của trường âm thanh thích hợp. Cần quy trình giải mã để thu được tín hiệu loa riêng

rẽ từ các tín hiệu được định dạng Ambisonics này. Ngoài ra, do trong trường hợp này hàm quét có thể được lấy đạo hàm từ hàm giải mã nên hàm quét là vấn đề mấu chốt để mô tả nhiệm vụ của sự định vị hóa không gian. Trong bản mô tả này, sự sắp xếp không gian của các loa được gọi là cài đặt loa.

Cài đặt loa thường được sử dụng là cài đặt chế độ âm thanh nổi; chế độ này sử dụng hai loa, cài đặt âm thanh vòm chuẩn sử dụng năm loa, và mở rộng cài đặt âm thanh vòm sử dụng nhiều hơn năm loa. Các cài đặt này đã được biết đến rộng rãi. Tuy nhiên, chúng bị hạn chế ở hai chiều (2D), ví dụ không phát lại thông tin chiều cao.

Các cài đặt loa để phát lại 3 chiều (3D) được mô tả, ví dụ, trong "Wide listening area with exceptional spatial sound quality of a 22.2 multichannel sound system", K. Hamasaki, T. Nishiguchi, R. Okumaura, và Y. Nakayama trong tạp chí Audio Engineering Society Preprints, Vienna, Austria, 5/2007, đây là đề xuất liên quan đến TV độ nét cao NHK với định dạng 22.2 hoặc sự sắp xếp 2+2+2 của Dabringhaus (mdg-musikproduktion dabringhaus und grimm, [www.mdg.de](http://www.mdg.de)) và cài đặt 10.2 trong "Sound for Film and Television", T. Holman trong 2nd ed. Boston: Focal Press, 2002. Một trong vài hệ thống đã biết liên quan đến kỹ thuật phát lại và quét trong không gian là phương pháp quét biên độ dựa vào vector (VBAP) trong "Virtual sound source positioning using vector base amplitude panning," Journal of Audio Engineering Society, vol. 45, no. 6, pp. 456-466, 6/1997, ở đây là Pulkki. VBAP (Vector Based Amplitude Panning) đã được sử dụng bởi Pulkki để phát lại nguồn âm thanh ảo với cài đặt loa tùy ý. Cần có hai loa để đặt nguồn ảo trong mặt phẳng 2D và cần có 3 loa để đặt nguồn ảo trong mặt phẳng 3D. Đối với mỗi nguồn ảo, tín hiệu đơn âm với độ khuếch đại khác nhau (tùy thuộc vào vị trí của nguồn ảo) được cấp đến các loa đã chọn từ cài đặt đầy đủ. Sau đó, tín hiệu loa cho tất cả nguồn ảo được tổng hợp. VBAP áp dụng phương pháp hình học để tính độ khuếch đại của các tín hiệu loa cho sự quét giữa các loa.

Ví dụ về cài đặt loa 3D mang tính minh họa được xem xét và đề xuất trong bản mô tả này có 16 loa được bố trí như trên Fig.2. Cách bố trí này được chọn nhờ lý do thiết thực, có bốn cột, trên mỗi cột có ba loa và các loa phụ ở giữa các cột. Chi tiết hơn, tám loa được phân bố đều trên vòng tròn quanh đầu của người nghe, tạo thành

các góc 45 độ. Bốn loa phụ được đặt tại đỉnh và đáy, tạo thành các góc phương vị 90 độ. Đối với Ambisonics, cài đặt này không theo quy tắc và dẫn đến những vấn đề về thiết kế bộ giải mã, như được nêu trong "An ambisonics format for flexible playback layouts," bởi H. Pomberger và F. Zotter trong Proceedings of the 1st Ambisonics Symposium, Graz, Austria, 7/2009.

Giải mã Ambisonics thông thường, như được mô tả trong "Three-dimensional surround sound systems based on spherical harmonics" bởi M. Poletti trong J. Audio Eng. Soc, vol. 53, no.11, pp. 1004-1025, 11/2005, sử dụng quy trình so khớp chế độ thông thường đã biết. Các chế độ này được mô tả bằng các vector chế độ mà chứa các giá trị của hàm điều hòa cầu đối với hướng tới phân biệt. Sự kết hợp của tất cả các hướng đưa ra bởi các loa riêng rẽ tạo ra ma trận chế độ của cài đặt loa, vì vậy ma trận chế độ thể hiện vị trí loa. Để phát lại chế độ của tín hiệu nguồn phân biệt, chế độ của loa được tính trọng số theo cách trong đó các chế độ chồng lên nhau của các loa riêng rẽ kết hợp thành một chế độ mong muốn. Để đạt được trọng số cần thiết, cần tính biểu diễn ma trận nghịch đảo của ma trận chế độ loa. Về sự giải mã tín hiệu, trọng số tạo thành tín hiệu dẫn động của các loa, và ma trận chế độ loa nghịch đảo được gọi là "ma trận giải mã", mà được ứng dụng để giải mã biểu diễn tín hiệu định dạng Ambisonics. Đặc biệt, đối với nhiều cài đặt loa, ví dụ, cài đặt được thể hiện trên Fig.2, khó thu được ma trận nghịch đảo của ma trận chế độ.

Như đã nêu trên đây, cài đặt loa được thường dùng bị hạn chế ở 2D, tức là không tái tạo thông tin độ cao. Sự giải mã biểu diễn trường âm thanh thành cài đặt loa với sự phân bố không gian không theo quy tắc toán học dẫn đến các vấn đề về định vị hóa và màu sắc với những kỹ thuật đã biết thông thường. Để giải mã tín hiệu Ambisonics, ma trận giải mã (tức là, ma trận của hệ số giải mã) được sử dụng. Trong kỹ thuật giải mã tín hiệu Ambisonics thông thường, và cụ thể là tín hiệu HOA, ít nhất xảy ra hai vấn đề. Thứ nhất, để giải mã chính xác cần phải biết hướng của nguồn tín hiệu để thu được ma trận giải mã. Thứ hai, sự ánh xạ đến cài đặt loa đã có là sai một cách hệ thống do vấn đề toán học dưới đây: sự giải mã chính xác về toán học sẽ tạo ra không chỉ biên độ loa dương mà còn tạo ra một số biên độ loa âm. Tuy nhiên, các biên độ này được phát lại sai thành các tín hiệu dương, do đó dẫn đến các vấn đề nêu trên.

## Bản chất kỹ thuật của sáng chế

Sáng chế mô tả phương pháp giải mã biểu diễn trường âm thanh cho các phân bố không gian không theo quy tắc với những đặc tính định vị hóa và màu sắc được cải thiện nhiều. Sáng chế thể hiện cách khác để thu được ma trận giải mã cho dữ liệu trường âm thanh, tức là ở định dạng Ambisonics, và sử dụng quy trình theo cách ước lượng hệ thống. Xét tập hợp hướng tới có thể có, hàm quét liên quan đến loa thích hợp được tính toán. Hàm quét chọn làm đầu ra của quy trình giải mã Ambisonics. Tín hiệu đầu vào cần thiết là ma trận chế độ của tất cả các hướng xét. Do đó, như được nêu dưới đây, ma trận giải mã được thu bằng cách nhân ma trận trọng số với phiên bản nghịch đảo của ma trận chế độ chứa tín hiệu đầu vào.

Liên quan đến vấn đề thứ hai nêu trên, đã phát hiện ra rằng cũng có thể thu được ma trận giải mã từ ma trận nghịch đảo của ma trận chế độ, mà thể hiện vị trí của loa, và chức năng tính trọng số phụ thuộc vị trí ("hàm quét")  $W$ . Theo một khía cạnh của sáng chế hàm quét  $W$  này có thể được lấy đạo hàm bằng cách sử dụng phương pháp khác với phương pháp thường được sử dụng. Ưu tiên sử dụng phương pháp hình học đơn giản. Phương pháp này không đòi hỏi biết hướng của nguồn tín hiệu, do đó giải quyết vấn đề thứ nhất nêu trên. Phương pháp như vậy được gọi là "Quét biên độ dựa trên Vector" (Vector-Based Amplitude Panning" (VBAP)). Theo sáng chế, VBAP được sử dụng để tính hàm quét yêu cầu, hàm quét này sau đó lại được sử dụng để tính ma trận giải mã Ambisonics. Vấn đề khác xảy ra trong đó cần ma trận nghịch đảo của ma trận chế độ (mà thể hiện cài đặt loa). Tuy nhiên, khó thu được ma trận nghịch đảo chính xác, điều này dẫn đến sự phát lại âm thanh sai. Do đó, theo khía cạnh khác, để thu được ma trận giải mã, cần tính ma trận chế độ giả nghịch đảo mà dễ thu được hơn nhiều.

Sáng chế sử dụng phương pháp hai bước. Bước thứ nhất là lấy đạo hàm hàm quét mà phụ thuộc vào cài đặt loa được sử dụng để phát lại. Trong bước thứ hai, ma trận giải mã Ambisonics được tính từ các hàm quét này cho tất cả các loa.

Ưu điểm của sáng chế là ở chỗ không cần mô tả thông số của nguồn âm thanh; thay vào đó, có thể sử dụng mô tả trường âm thanh như Ambisonics.

Theo sáng chế, phương pháp giải mã biểu diễn trường âm thanh để phát lại âm thanh bao gồm các bước tính, đối với mỗi loa trong số nhiều loa, hàm quét sử dụng phương pháp hình học dựa vào vị trí của loa và nhiều hướng nguồn, tính ma trận chế độ từ các hướng nguồn, tính ma trận chế độ giả nghịch đảo của ma trận chế độ, và giải mã biểu diễn trường âm thanh, trong đó sự giải mã này dựa vào ma trận đã giải mã mà thu được từ ít nhất hàm quét và ma trận chế độ giả nghịch đảo.

Theo khía cạnh khác, thiết bị giải mã biểu diễn trường âm thanh để phát lại âm thanh bao gồm phương tiện tính thứ nhất để tính, đối với mỗi trong số nhiều loa, hàm quét sử dụng phương pháp hình học dựa vào vị trí của loa và nhiều hướng nguồn, phương tiện tính thứ hai để tính ma trận chế độ từ các hướng nguồn, phương tiện tính thứ ba để tính ma trận chế độ giả nghịch đảo của ma trận chế độ, và phương tiện giải mã để giải mã biểu diễn trường âm thanh, trong đó sự giải mã này dựa vào ma trận giải mã và phương tiện giải mã sử dụng ít nhất hàm quét và ma trận chế độ giả nghịch đảo để thu được ma trận giải mã. Phương tiện tính thứ nhất, thứ hai và thứ ba có thể là bộ xử lý đơn hoặc hai hoặc nhiều bộ xử lý riêng rẽ.

Theo một khía cạnh khác, vật ghi đọc được bằng máy tính lưu trữ các lệnh thực thi được để khiến máy tính thực hiện phương pháp giải mã biểu diễn trường âm thanh để phát lại âm thanh bao gồm các bước tính, đối với mỗi trong số nhiều loa, hàm quét sử dụng phương pháp hình học dựa vào vị trí của loa và nhiều hướng nguồn, tính ma trận chế độ từ các hướng nguồn, tính ma trận chế độ giả nghịch đảo của ma trận chế độ, và giải mã biểu diễn trường âm thanh, trong đó sự giải mã dựa vào ma trận giải mã mà thu được từ ít nhất hàm quét và ma trận chế độ giả nghịch đảo.

Phương án có lợi của sáng chế được bộc lộ trong các điểm yêu cầu bảo hộ phụ thuộc, phân mô tả và hình vẽ dưới đây.

### **Mô tả vắn tắt các hình vẽ**

Các phương án ví dụ của sáng chế được mô tả với sự tham khảo các hình vẽ đi kèm, trong đó:

Fig.1 là lưu đồ của phương pháp;

Fig.2 là cài đặt 3D ví dụ với 16 loa.

Fig.3 là đồ thị chùm tia thu được từ việc giải mã bằng cách sử dụng phép so khớp chế độ không theo quy tắc;

Fig.4 là đồ thị chùm tia thu được từ việc giải mã bằng cách sử dụng ma trận chế độ theo quy tắc;

Fig.5 là đồ thị chùm tia thu được từ việc giải mã bằng cách sử dụng ma trận giải mã được lấy đạo hàm từ VBAP;

Fig.6 là kết quả của kiểm tra nghe; và

Fig.7 là sơ đồ khối của thiết bị.

### Mô tả chi tiết sáng chế

Như được thể hiện trên Fig.1, phương pháp giải mã biểu diễn trường âm thanh  $SF_c$  để phát lại âm thanh bao gồm các bước tính 110, đối với mỗi trong số nhiều loa, hàm quét  $W$  bằng cách sử dụng phương pháp hình học dựa vào vị trí 102 của các loa ( $L$  là số loa) và nhiều hướng nguồn 103 ( $S$  là số hướng nguồn), tính 120 ma trận chế độ  $\Xi$  từ các hướng nguồn và bậc  $N$  cho trước của biểu diễn trường âm thanh, tính 130 ma trận chế độ giả nghịch đảo  $\Xi^+$  của ma trận chế độ  $\Xi$ , và giải mã 135,140 biểu diễn trường âm thanh  $SF_C$ , trong đó thu được dữ liệu âm thanh đã giải mã  $AU_{dec}$ . Sự giải mã này dựa trên cơ sở ma trận giải mã  $D$  mà thu được 135 từ ít nhất hàm quét  $W$  và ma trận chế độ giả nghịch đảo  $\Xi^+$ . Trong một phương án, ma trận chế độ giả nghịch đảo thu được theo  $\Xi^+ = \Xi^H [\Xi \Xi^H]^{-1}$ . Bậc  $N$  của biểu diễn trường âm thanh có thể được xác định trước, hoặc có thể được khai căn 105 từ tín hiệu đầu vào  $SF_C$ .

Như được thể hiện trên Fig.7, thiết bị để giải mã biểu diễn trường âm thanh để phát lại âm thanh bao gồm phương tiện tính thứ nhất 210 để tính, đối với mỗi loa trong số nhiều loa, hàm quét  $W$  bằng cách sử dụng phương pháp hình học dựa vào vị trí 102 của loa và nhiều hướng nguồn 103, phương tiện tính thứ hai 220 để tính ma

trận chế độ  $\Xi$  từ các hướng nguồn, phương tiện tính thứ ba 230 để tính ma trận chế độ giả nghịch đảo  $\Xi^+$  của ma trận chế độ  $\Xi$ , và phương tiện giải mã 240 để giải mã biểu diễn trường âm thanh. Sự giải mã dựa trên cơ sở ma trận giải mã D mà thu được từ ít nhất hàm quét W và ma trận chế độ giả nghịch đảo  $\Xi^+$  bằng phương tiện tính ma trận giải mã 235 (ví dụ, bộ nhân). Phương tiện giải mã 240 sử dụng ma trận giải mã D để thu được tín hiệu âm thanh đã giải mã  $AU_{dec}$ . Phương tiện tính thứ nhất, thứ hai và thứ ba 220,230,240 có thể là bộ xử lý đơn hoặc hai hoặc nhiều bộ xử lý riêng rẽ. Bậc N của trường âm thanh biểu diễn có thể được xác định trước, hoặc có thể thu được bằng phương tiện 205 để khai căn bậc từ tín hiệu đầu vào  $SF_C$ .

Cài đặt loa 3D đặc biệt hữu ích có 16 loa. Như được thể hiện trên Fig.2, có bốn cột, trên mỗi cột có ba loa và các loa phụ ở giữa các cột. Tám loa được phân bố đều trên vòng tròn quanh đầu của người nghe, tạo thành các góc 45 độ. Bốn loa phụ được đặt tại đỉnh và đáy, tạo thành các góc phương vị 90 độ. Đối với Ambisonics, cài đặt này không theo quy tắc và thường dẫn đến các vấn đề về thiết kế của bộ giải mã.

Phần mô tả dưới đây sẽ mô tả chi tiết phương pháp quét biên độ dựa vào vector (Vector Base Amplitude Panning - VBAP). Trong một phương án, VBAP được sử dụng để đặt nhiều nguồn âm thanh ảo với cài đặt loa tùy ý trong đó giả sử khoảng cách của các loa đến vị trí nghe là như nhau. VBAP sử dụng ba loa để đặt nguồn ảo trong không gian 3D. Đối với mỗi nguồn ảo, một tín hiệu đơn âm với độ khuếch đại khác nhau được nạp vào các loa được sử dụng. Độ khuếch đại đối với các loa khác nhau phụ thuộc vào vị trí của nguồn ảo. VBAP là phương pháp hình học để tính độ khuếch đại của tín hiệu loa để quét giữa các loa. Trong trường hợp 3D, ba loa được sắp xếp theo hình tam giác tạo thành vector cơ sở. Mỗi vector cơ sở được xác định bởi số loa k,m,n và vector vị trí của loa  $l_k, l_m, l_n$  đã cho trong tọa độ Đề Các được bình thường hóa tới chiều dài đơn vị. Vector cơ sở cho các loa k, m, n được xác định bởi:

$$L_{kmn} = \{l_k, l_m, l_n\} \quad (1)$$

Hướng mong muốn  $\Omega = (\theta, \phi)$  của nguồn ảo phải được cho là góc phương vị  $\phi$  và góc nghiêng  $\theta$ . Do đó, vector định vị độ dài đơn vị  $p(\Omega)$  của nguồn ảo trong tọa độ Đề Các được xác định bởi:



$$p(\Omega) = \{\cos\phi \sin\theta, \sin\phi \sin\theta, \cos\theta\}^T \quad (2)$$

Vị trí nguồn ảo có thể được biểu diễn bằng vector cơ sở và hệ số khuếch đại  $g(\Omega) = (\tilde{g}_k, \tilde{g}_m, \tilde{g}_n)^T$  bằng:

$$p(\Omega) = L_{kmn} g(\Omega) = \tilde{g}_k l_k + \tilde{g}_m l_m + \tilde{g}_n l_n \quad (3)$$

Bằng cách nghịch đảo ma trận vector, hệ số khuếch đại yêu cầu có thể được tính bằng:

$$g(\Omega) = L_{kmn}^{-1} p(\Omega) \quad (4)$$

Vector cơ sở cần sử dụng được xác định theo tài liệu của Pulkki: Đầu tiên, độ khuếch đại được tính theo Pulkki đối với tất cả các vector cơ sở. Sau đó, đối với mỗi vector cơ sở, cực tiểu trên hệ số khuếch đại được đánh giá bằng  $\tilde{g}_{min} = \min\{\tilde{g}_k, \tilde{g}_m, \tilde{g}_n\}$ . Cuối cùng, vector cơ sở trong đó  $\tilde{g}_{min}$  có giá trị cao nhất được sử dụng. Hệ số khuếch đại thu được phải không âm. Tùy thuộc vào phòng hòa âm, hệ số khuếch đại có thể được chuẩn hóa để bảo toàn năng lượng.

Định dạng Ambisonics được mô tả dưới đây là một định dạng trường âm thanh mang tính minh họa. Biểu diễn Ambisonics là phương pháp mô tả trường âm thanh sử dụng phép tính xấp xỉ toán học của trường âm thanh trong một vị trí. Bằng cách sử dụng hệ tọa độ hình cầu, áp lực tại điểm  $r = (r, \theta, \phi)$  trong không gian được mô tả bằng biến đổi Fourier hình cầu:

$$p(r, k) = \sum_{n=0}^{\infty} \sum_{m=-n}^n A_n^m(k) j_n(kr) Y_n^m(\theta, \phi) \quad (5)$$

trong đó  $k$  là số sóng. Thông thường  $n$  chạy đến bậc xác định  $M$ . Hệ số  $A_n^m(k)$  của cấp số mô tả trường âm thanh (giả sử các nguồn nằm ngoài vùng giá trị),  $j_n(kr)$  là hàm Bessel hình cầu cấp một và  $Y_n^m(\theta, \phi)$  chỉ hàm điều hòa cầu. Hệ số  $A_n^m(k)$  được coi là hệ số Ambisonics trong ngữ cảnh này. Hàm điều hòa cầu  $Y_{mn}(\theta, \phi)$  chỉ phụ thuộc vào góc nghiêng và góc phương vị và mô tả hàm trên hình cầu đơn vị.

Để đơn giản hóa, sóng phẳng được giả thiết là để phát lại trường âm thanh. Hệ số Ambisonics mô tả sóng phẳng là nguồn âm thanh từ hướng  $\Omega_s$ .

$$A_{n,\text{plane}}^m(\Omega_s) = 4\pi i^n Y_n^m(\Omega_s)^* \quad (6)$$

Sự phụ thuộc của chúng vào số sóng  $k$  giảm tới độ phụ thuộc hướng trong trường hợp đặc biệt này. Đối với bậc giới hạn  $M$ , các hệ số tạo thành vector  $A$  mà có thể được sắp xếp là

$$A(\Omega_s) = [A_0^0 \ A_1^{-1} \ A_1^0 \ A_1^1 \ \dots \ A_M^M]^T \quad (7)$$

với  $O = (M + 1)^2$  phần tử. Sự sắp xếp tương tự được sử dụng cho

hệ số hàm điều hòa cầu thu được cho ra vector  $Y(\Omega_s)^* = [Y_0^0 \ Y_1^{-1} \ Y_1^0 \ Y_1^1 \ \dots \ A_M^M]^H$ .

Chỉ số trên  $H$  chỉ hoán vị số liên hợp phức.

Để tính tín hiệu loa từ biểu diễn Ambisonics của trường âm thanh, phương pháp thường được sử dụng là so khớp chế độ. Ý tưởng cơ bản là biểu diễn mô tả trường âm thanh Ambisonics đã cho  $A(\Omega_s)$  bằng tổng có trọng số của mô tả trường âm thanh của loa  $A(\Omega_l)$

$$A(\Omega_s) = \sum_{l=1}^L w_l A(\Omega_l) \quad (8)$$

trong đó  $\Omega_l$  là hướng của loa,  $w_l$  là trọng số, và  $L$  là số loa. Để lấy đạo hàm hàm quét từ phương trình (8), chúng ta giả sử hướng tới đã biết  $\Omega_s$ . Nếu nguồn và trường âm thanh đều là sóng phẳng, thừa số  $4\pi i^n$  (xem phương trình (6)) có thể được giảm và phương trình (8) chỉ phụ thuộc vào các liên hợp phức của vector điều hòa cầu, hay còn gọi là "chế độ". Sử dụng ký hiệu ma trận, phương trình này được viết dưới dạng:

$$Y(\Omega_s)^* = \Psi w(\Omega_s) \quad (9)$$

trong đó  $\Psi$  là ma trận chế độ của cài đặt loa

$$\Psi = [Y(\Omega_1)^*, Y(\Omega_2)^*, \dots, Y(\Omega_L)^*] \quad (10)$$

với  $O \times L$  phần tử. Để thu được vector trọng số  $w$  mong muốn, các phương pháp khác nhau đã được biết đến. Nếu chọn  $M = 3$ ,  $\Psi$  là góc vuông và có thể nghịch đảo. Do cài đặt loa không theo quy tắc nên ma trận được tỷ lệ hóa kém. Trong trường hợp này, thường chọn ma trận giả nghịch đảo và

$$D = [\Psi^H \Psi]^{-1} \Psi^H \quad (11)$$

thu được ma trận giải mã  $L \times O$   $D$ . Cuối cùng, chúng ta có thể viết

$$w(\Omega_s) = DY(\Omega_s)^* \quad (12)$$

trong đó các trọng số  $w(\Omega_s)$  là giải pháp năng lượng tối thiểu cho phương trình (9). Kết quả từ việc sử dụng ma trận giả nghịch đảo được mô tả dưới đây.

Dưới đây sẽ mô tả liên kết giữa hàm quét và ma trận giải mã Ambisonics. Bắt đầu với Ambisonics, hàm quét cho các loa riêng rẽ có thể được tính bằng cách sử dụng phương trình (12). Giả sử

$$\Xi = [Y(\Omega_1)^*, Y(\Omega_2)^*, \dots, Y(\Omega_S)^*] \quad (13)$$

là ma trận chế độ của  $S$  hướng tín hiệu đầu vào  $(\Omega_s)$ , ví dụ lưới hình cầu với góc nghiêng trong các bước một độ từ  $1 \dots 180^\circ$  và góc phương vị từ  $1 \dots 360^\circ$  tương ứng. Ma trận chế độ này có  $O \times S$  phần tử. Sử dụng phương trình (12), ma trận  $W$  thu được có  $L \times S$  phần tử, hàng  $I$  giữ trọng số quét  $S$  cho loa tương ứng:

$$W = D\Xi \quad (14)$$

Ví dụ, hàm quét của loa đơn 2 được thể hiện dưới dạng đồ thị chùm tia trên Fig.2. Ma trận giải mã  $D$  có bậc  $M = 3$  trong ví dụ này. Như có thể thấy, giá trị hàm quét hoàn toàn không liên quan đến sự định vị hữu hình của loa. Điều này là do sự định vị không theo quy tắc toán học của các loa, tức là không đủ làm sơ đồ lấy mẫu trong không gian đối với bậc được chọn. Do đó, ma trận giải mã được gọi là ma trận chế độ không theo quy tắc. Vấn đề này có thể vượt qua bằng cách quy tắc hóa ma trận

chế độ của loa  $\Psi$  trong phương trình (11). Giải pháp này là nhờ sự phân giải trong không gian của ma trận giải mã, ma trận giải mã này lại có thể được biểu diễn là bậc Ambisonics thấp. Fig.4 thể hiện đồ thị chùm tia ví dụ thu được từ sự giải mã bằng cách sử dụng ma trận chế độ theo quy tắc, và cụ thể bằng cách sử dụng giá trị riêng của ma trận cho sự quy tắc hóa. So sánh với Fig.3, giờ đây hướng của loa theo địa chỉ được nhận rõ.

Như đã nêu trong phần giới thiệu, có thể có cách khác để thu được ma trận giải mã  $D$  để phát lại tín hiệu Ambisonics khi hàm quét đã được biết đến. Hàm quét  $W$  được xem là tín hiệu thích hợp trên hướng nguồn ảo  $\Omega$ , và ma trận chế độ  $\Xi$  của các hướng này đóng vai trò là tín hiệu đầu vào. Sau đó, ma trận giải mã có thể được tính bằng cách sử dụng:

$$D = W \Xi^H [\Xi \Xi^H]^{-1} = W \Xi^+ \quad (15)$$

trong đó  $\Xi^H [\Xi \Xi^H]^{-1}$  hoặc đơn giản  $\Xi^+$  là giả nghịch đảo của ma trận chế độ  $\Xi$ . Trong phương pháp mới, chúng ta lấy hàm quét theo  $W$  từ VBAP và tính ma trận giải mã Ambisonics từ đó.

Hàm quét đối với  $W$  được lấy làm giá trị khuếch đại  $g(\Omega)$  được tính bằng cách sử dụng phương trình (4), trong đó  $\Omega$  được chọn theo phương trình (13). Ma trận giải mã tạo thành bằng cách sử dụng phương trình (15) là ma trận giải mã Ambisonics tạo thuận lợi cho hàm quét VBAP. Ví dụ được mô tả trên Fig.5, mà thể hiện đồ thị chùm tia kết quả của sự giải mã bằng cách sử dụng ma trận giải mã được lấy đạo hàm từ VBAP. Có lợi là thùy bên SL nhỏ hơn nhiều so với thùy bên  $SL_{reg}$  của kết quả so khớp chế độ theo quy tắc theo Fig.4. Hơn nữa, đồ thị chùm tia từ VBAP đối với các loa riêng rẽ tuân theo hình học của cài đặt loa do hàm quét VBAP phụ thuộc vào cơ sở vector của hướng được xử lý. Vì vậy, phương pháp mới theo sáng chế cho ra kết quả tốt hơn đối với tất cả các hướng cài đặt loa.

Hướng nguồn 103 có thể được xác định khá tự do. Điều kiện đối với số lượng hướng nguồn  $S$  ít nhất phải là  $(N+1)^2$ . Do đó, cho bậc  $N$  của tín hiệu trường âm thanh  $SF_c$ , có thể xác định  $S$  theo  $S \geq (N+1)^2$ , và phân bố hướng nguồn  $S$  đồng đều trên hình cầu đơn vị. Như đã nêu trên, kết quả có thể là lưới hình cầu với góc nghiêng  $\theta$  chạy

trong các bước cố định  $x$  (ví dụ  $x = 1...5$  hoặc  $x=10, 20$ , v.v.) độ từ  $1...180^\circ$  và góc phương vị  $\phi$  từ  $1...360^\circ$  tương ứng, trong đó mỗi hướng nguồn  $\Omega = (\theta, \phi)$  có thể được cho bởi góc phương vị  $\phi$  và góc nghiêng  $\theta$ .

Tác dụng có lợi đã được khẳng định trong kiểm tra nghe. Để đánh giá sự định vị hóa của nguồn đơn, một nguồn ảo được so sánh với nguồn thực làm đối chiếu. Đối với nguồn thực loa tại vị trí mong muốn được sử dụng. Phương pháp phát lại được sử dụng là VBAP, giải mã so khớp chế độ Ambisonics và giải mã Ambisonics mới được đề xuất sử dụng hàm quét VBAP theo sáng chế. Đối với hai phương pháp sau, đối với mỗi vị trí kiểm tra và mỗi tín hiệu đầu vào kiểm tra, tín hiệu Ambisonics bậc ba được tạo ra. Tín hiệu Ambisonics tổng hợp này sau đó được giải mã bằng cách sử dụng ma trận giải mã tương ứng. Tín hiệu kiểm tra được sử dụng là tín hiệu nhiễu hồng băng rộng và tín hiệu tiếng nói nam giới. Vị trí kiểm tra được đặt trong vùng trước với các hướng

$$\Omega_1 = (76,1^\circ, -23,2^\circ), \Omega_2 = (63,3^\circ, -4,3^\circ) \quad (16)$$

Kiểm tra nghe được thực hiện trong phòng hòa âm với thời gian vọng trung bình xấp xỉ 0,2 giây. Chín người tham gia kiểm tra nghe. Đối tượng kiểm tra được đề nghị phân loại hiệu suất phát lại trong không gian của tất cả các phương pháp phát lại so với tham chiếu. Giá trị phân loại đơn phải thể hiện sự định vị hóa của nguồn ảo và sự thay đổi âm sắc. Fig.5 thể hiện kết quả kiểm tra nghe.

Theo kết quả, sự giải mã so khớp chế độ Ambisonics không theo quy tắc được phân loại kém hơn về mặt cảm nhận so với các phương pháp khác khi kiểm tra. Kết quả này tương ứng với Fig.3. Phương pháp so khớp chế độ Ambisonics đóng vai trò làm mẫu neo trong kiểm tra nghe. Ưu điểm khác là ở chỗ khoảng tin cậy cho tín hiệu tạp nhiễu lớn hơn khoảng tin cậy cho VBAP đối với các phương pháp khác. Giá trị trung bình thể hiện giá trị cao nhất đối với sự giải mã Ambisonics bằng cách sử dụng hàm quét VBAP. Do đó, mặc dù độ phân giải trong không gian được giảm – do sử dụng bậc Ambisonics – nhưng phương pháp này thể hiện ưu điểm so với phương pháp VBAP tham số. So với VBAP, sự giải mã Ambisonics với hàm quét thô và VBAP có ưu điểm ở chỗ không chỉ ba loa được sử dụng kết xuất nguồn ảo. Trong

VBAP, loa đơn có thể là chính nếu vị trí nguồn ảo gần với một trong số vị trí vật lý của loa. Theo báo cáo của hầu hết các đối tượng, sự thay đổi đơn sắc đối với VBAP sử dụng Ambisonics ít hơn đối với VBAP được sử dụng trực tiếp. Vấn đề về sự thay đổi đơn sắc đối với VBAP đã được biết đến từ Pulkki. Ngược lại với VBAP, phương pháp mới đề xuất sử dụng nhiều hơn ba loa để phát lại nguồn ảo, nhưng đáng ngạc nhiên là tạo ra ít màu sắc (coloration) hơn.

Kết luận là phương pháp mới đề xuất thu được một ma trận giải mã Ambisonics từ hàm quét VBAP được bộc lộ. Đối với các cài đặt loa khác nhau, phương pháp này có lợi so với các ma trận của phương pháp so khớp chế độ. Đặc điểm và các hệ quả đã thảo luận ở trên. Tóm lại, sự giải mã Ambisonics mới đề xuất với hàm quét VBAP tránh các vấn đề thông thường của phương pháp so khớp chế độ đã biết. Kiểm tra nghe cho thấy rằng giải mã Ambisonics từ VBAP có thể tạo ra chất lượng phát lại trong không gian tốt hơn so với sử dụng trực tiếp VBAP. Phương pháp đề xuất chỉ yêu cầu mô tả trường âm thanh trong khi VBAP yêu cầu mô tả tham số của nguồn ảo được kết xuất.

Mặc dù đặc điểm mới cơ bản của sáng chế cũng như các phương án ưu tiên của sáng chế đã được thể hiện, mô tả và chỉ ra, nhưng phải hiểu rằng người có trình độ trong lĩnh vực kỹ thuật có thể thực hiện sự bỏ sót, thay thế và thay đổi trong thiết bị và phương pháp đã mô tả ở dạng và chi tiết của thiết bị đã bộc lộ mà không nằm ngoài ý tưởng của sáng chế. Mục đích rõ ràng là tất cả các cách kết hợp của các thành phần mà thực hiện cùng một chức năng theo cùng một cách để đạt được cùng một kết quả là nằm trong phạm vi bảo hộ của sáng chế. Sự thay thế thành phần từ một phương án đã mô tả cho thành phần khác cũng được mong muốn và dự tính. Phải hiểu rằng sự cải biến chi tiết có thể được thực hiện mà không nằm ngoài phạm vi bảo hộ của sáng chế. Mỗi đặc tính đã bộc lộ trong phần mô tả và (khi thích hợp) yêu cầu bảo hộ cũng như hình vẽ có thể được đưa ra độc lập hoặc theo cách kết hợp phù hợp. Những đặc tính này có thể là được thực hiện trong phần cứng, phần mềm, hoặc sự kết hợp của cả hai.

Các số tham chiếu xuất hiện của yêu cầu bảo hộ chỉ mang tính minh họa và không có tác dụng làm giới hạn phạm vi bảo hộ của các yêu cầu bảo hộ.

## YÊU CẦU BẢO HỘ

1. Phương pháp giải mã biểu diễn trường âm thanh ambisonics để phát lại âm thanh qua nhiều loa, phương pháp này bao gồm các bước:

thu được, với mỗi trong số nhiều loa, hàm quét bằng cách sử dụng phương pháp hình học dựa vào vị trí của các loa và nhiều hướng nguồn;

thu được ma trận chế độ từ các hướng nguồn và bậc của biểu diễn trường âm thanh ambisonics;

thu được ma trận cơ bản từ ma trận chế độ; và

giải mã biểu diễn trường âm thanh ambisonics bằng ma trận giải mã, trong đó ma trận giải mã dựa trên hàm quét và ma trận cơ bản, các hướng nguồn được phân phối đều trên hình cầu đơn vị, và số hướng nguồn là  $S$ , bậc của biểu diễn trường âm thanh ambisonics là  $N$ , và  $S \geq (N+1)^2$ .

2. Phương pháp theo điểm 1, trong đó phương pháp hình học được sử dụng trong bước thu được hàm quét là dựa trên quét biên độ dựa vào vector (Vector Base Amplitude Panning - VBAP).

3. Phương pháp theo điểm 1, trong đó biểu diễn trường âm thanh ambisonics ít nhất là bậc 2.

4. Phương pháp theo điểm 1, trong đó ma trận cơ bản là dựa trên tích của ma trận chế độ và ma trận nghịch đảo.

5. Phương pháp theo điểm 1, trong đó hàm quét được biểu diễn dưới dạng ma trận và ma trận cơ bản là sự quy tắc hóa của ma trận chế độ.

6. Phương pháp theo điểm 1, trong đó hàm quét được biểu diễn dưới dạng các giá trị độ khuếch đại.

7. Thiết bị giải mã biểu diễn trường âm thanh ambisonics để phát lại âm thanh qua nhiều loa, thiết bị này bao gồm:

phương tiện để thu được, với mỗi trong số nhiều loa, hàm quét bằng cách sử dụng phương pháp hình học dựa vào vị trí của các loa và nhiều hướng nguồn;

phương tiện để thu được ma trận chế độ từ các hướng nguồn và bậc của biểu diễn trường âm thanh ambisonics;

phương tiện để thu được ma trận cơ bản từ ma trận chế độ; và

phương tiện để giải mã biểu diễn trường âm thanh ambisonics bằng ma trận giải mã, trong đó ma trận giải mã dựa trên hàm quét và ma trận cơ bản, các hướng nguồn được phân phối đều trên hình cầu đơn vị, và số hướng nguồn là  $S$ , bậc của biểu diễn trường âm thanh ambisonics là  $N$ , và  $S \geq (N+1)^2$ .

8. Thiết bị theo điểm 7, trong đó phương pháp hình học sử dụng trong bước thu được hàm quét là dựa trên quét biên độ dựa vào vector (Vector Base Amplitude Panning - VBAP).

9. Thiết bị theo điểm 7, trong đó biểu diễn trường âm thanh ambisonics ít nhất là bậc 2.

10. Thiết bị theo điểm 7, trong đó ma trận cơ bản là dựa trên tích của ma trận chế độ và ma trận nghịch đảo.

11. Thiết bị theo điểm 7, trong đó hàm quét được biểu diễn dưới dạng ma trận và ma trận cơ bản là sự quy tắc hóa của ma trận chế độ.

12. Thiết bị theo điểm 7, trong đó hàm quét được biểu diễn dưới dạng các giá trị độ khuếch đại.

13. Vật ghi đọc được bằng máy tính lưu trữ trên đó các lệnh thực thi được để khiến cho máy tính thực hiện phương pháp giải mã biểu diễn trường âm thanh ambisonics để phát lại âm thanh, phương pháp này bao gồm các bước:

thu được, với mỗi trong số nhiều loa, hàm quét bằng cách sử dụng phương pháp hình học dựa vào vị trí của các loa và nhiều hướng nguồn;



thu được ma trận chế độ từ các hướng nguồn và bậc của biểu diễn trường âm thanh ambisonics;

thu được ma trận cơ bản từ ma trận chế độ; và

giải mã biểu diễn trường âm thanh ambisonics bằng ma trận giải mã, trong đó ma trận giải mã dựa trên hàm quét và ma trận cơ bản, các hướng nguồn được phân phối đều trên hình cầu đơn vị, và số hướng nguồn là  $S$ , bậc của biểu diễn trường âm thanh ambisonics là  $N$ , và  $S \geq (N+1)^2$ .

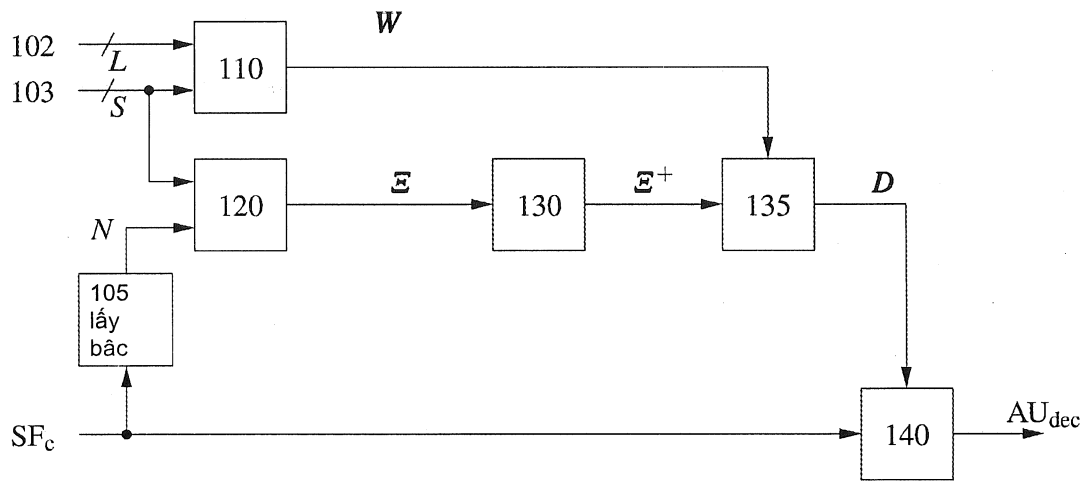


Fig.1

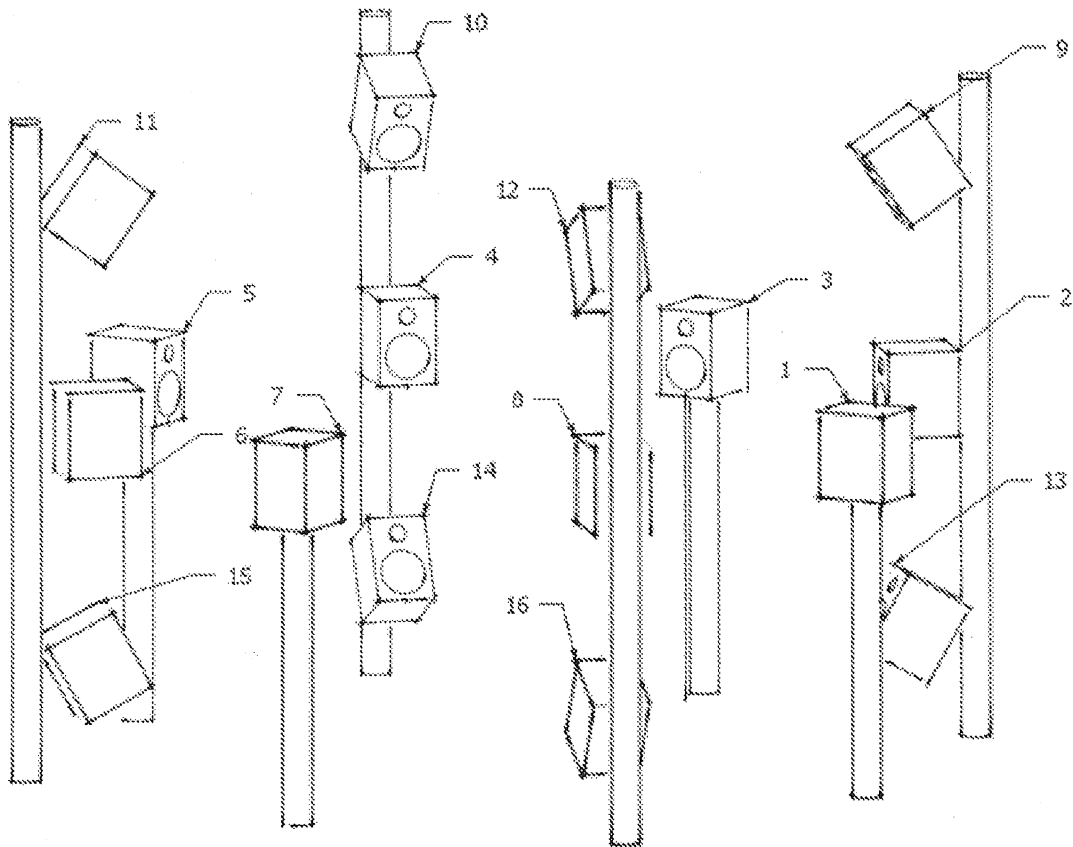


Fig.2

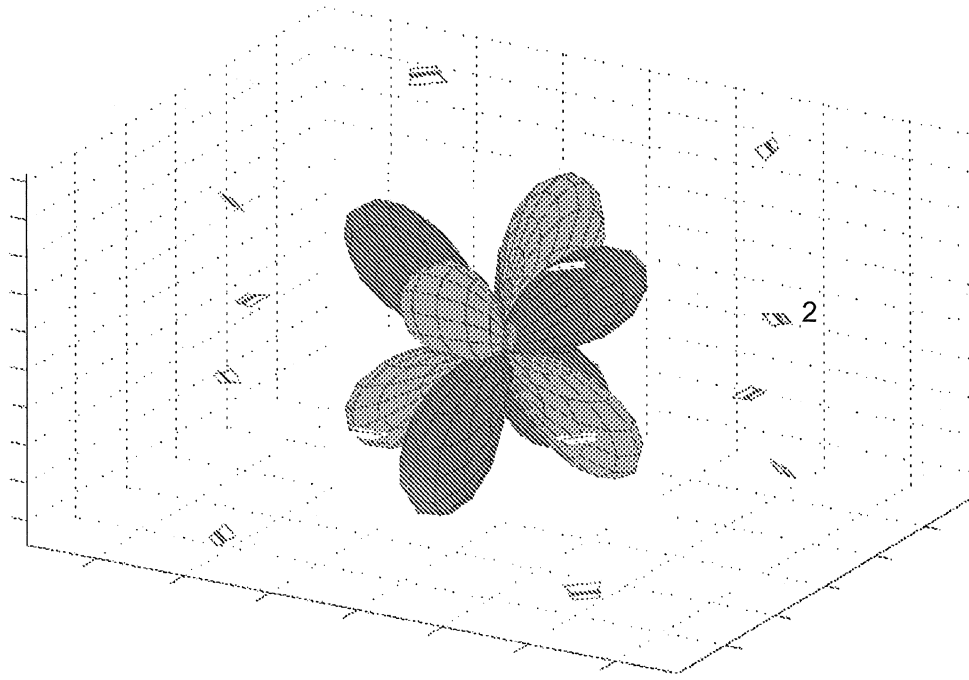


Fig.3

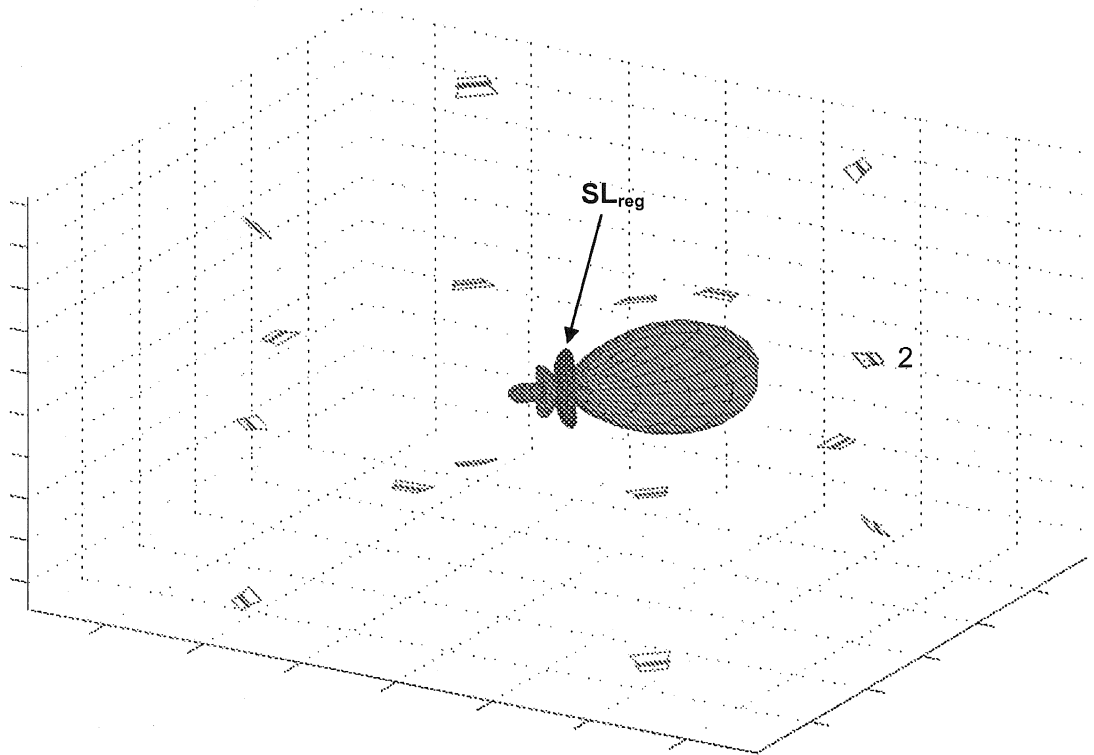


Fig.4

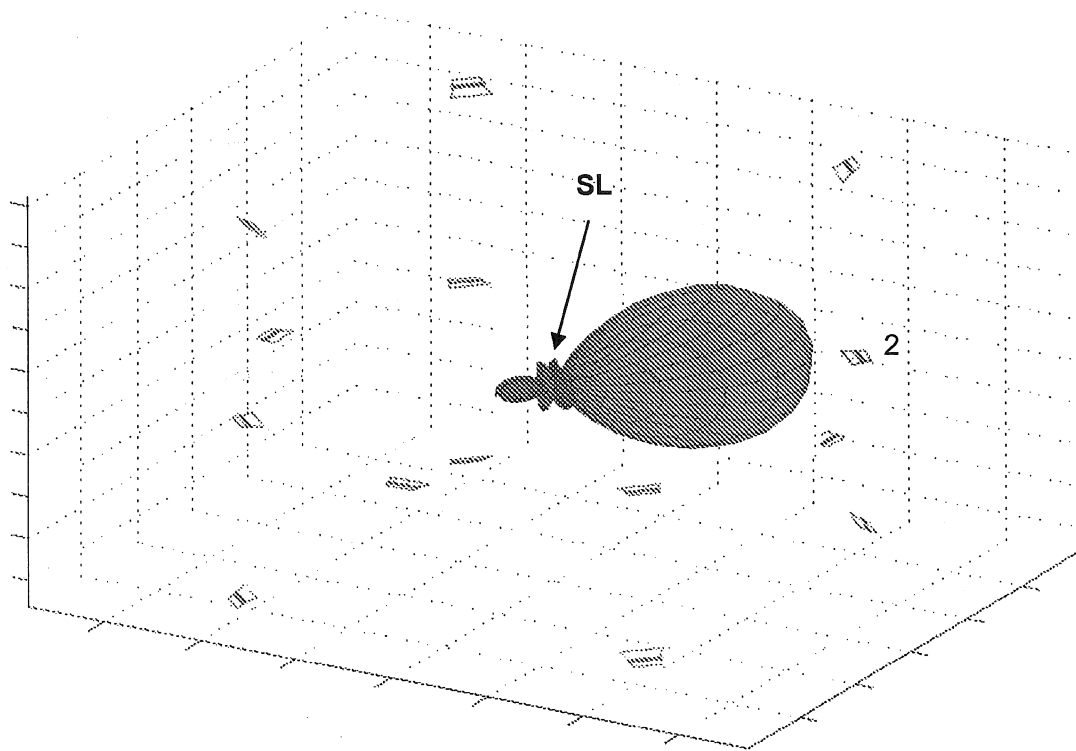


Fig.5

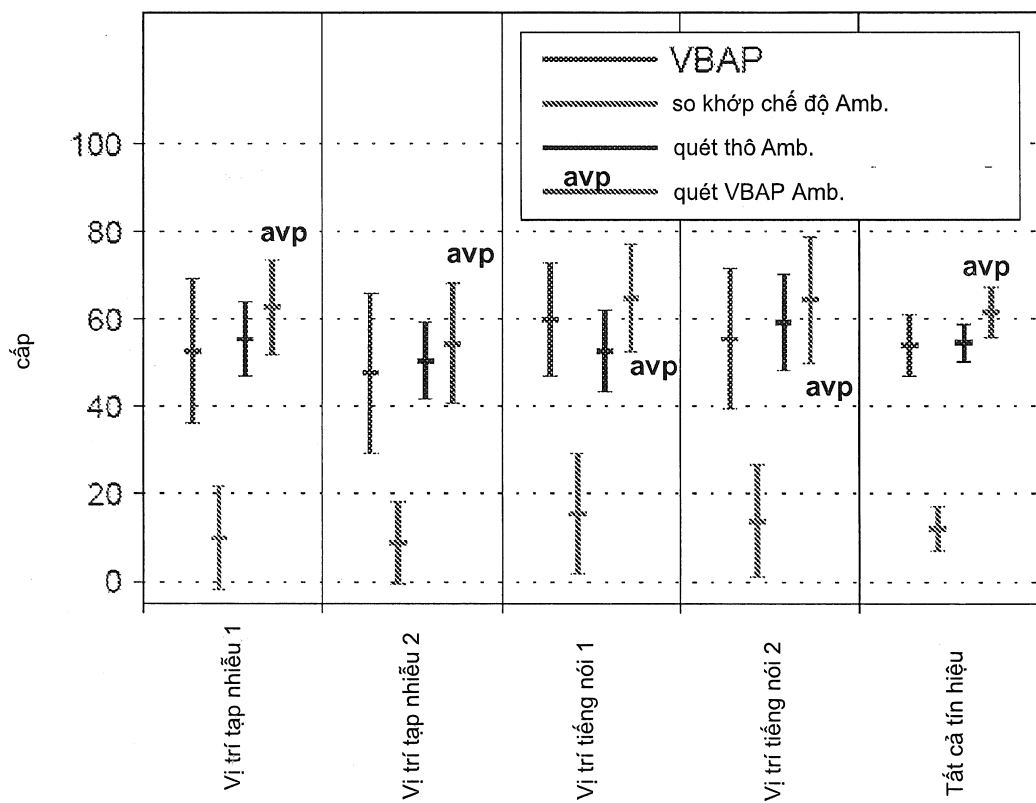


Fig.6

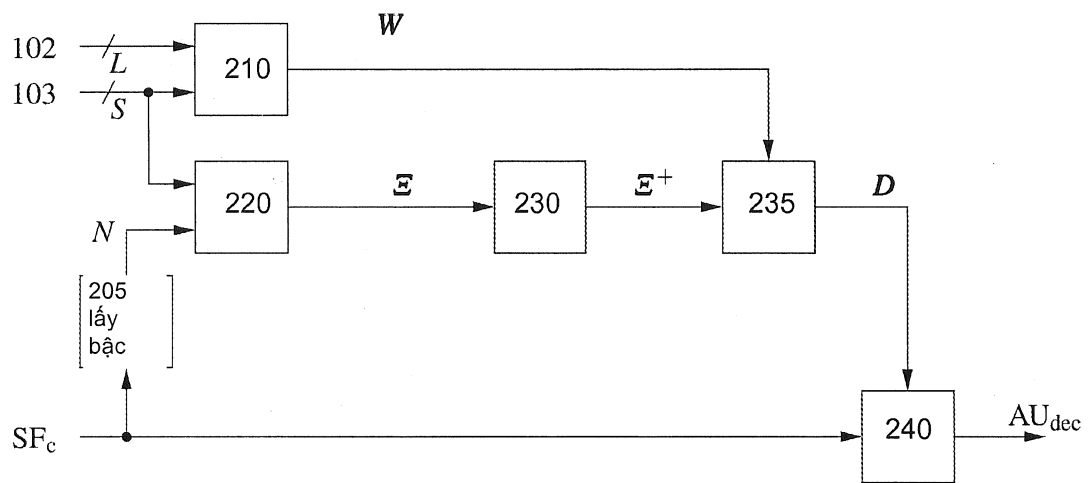


Fig.7