



(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



1-0023030

(51)⁷ G10L 21/00, H04S 3/00

(13) B

(21) 1-2014-00277

(22) 15.06.2012

(86) PCT/EP2012/061477 15.06.2012

(87) WO2013/000740 03.01.2013

(30) 11305845.7 30.06.2011 EP

(45) 25.02.2020 383

(43) 26.05.2014 314

(73) Interdigital Madison Patent Holdings (FR)

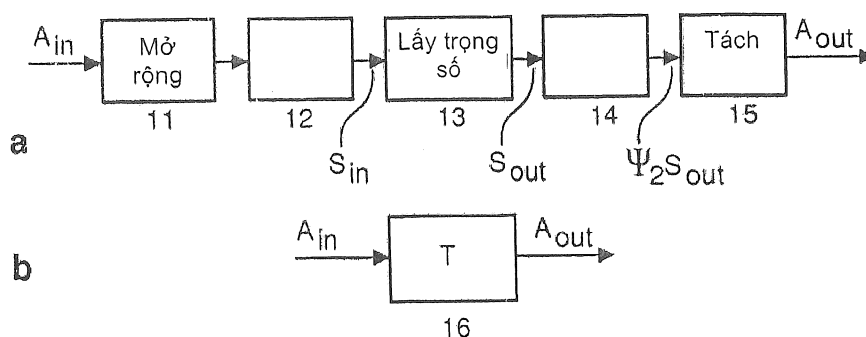
3 rue du Colonel Moll, 75017 Paris, France.

(72) JAX, Peter (DE), BATKE, Johann-Markus (DE)

(74) Công ty Luật TNHH T&G (TGVN)

(54) PHƯƠNG PHÁP VÀ THIẾT BỊ ĐỂ THAY ĐỔI CÁC VỊ TRÍ TƯƠNG ĐỐI CỦA CÁC ĐỐI TƯỢNG ÂM THANH NẰM TRONG PHÉP BIỂU DIỄN AMBISONICS BẬC CAO HƠN

(57) Sáng chế đề cập đến phương pháp và thiết bị để thay đổi các vị trí tương đối của các đối tượng âm thanh nằm trong phép biểu diễn Ambisonics bậc cao hơn (HOA). Trong đó HOA là phép biểu diễn của các trường âm thanh không gian hỗ trợ việc thu, quản lý, ghi, truyền và phát lại của các môi trường âm thanh phức tạp với độ phân giải không ưu việt, cả trong không gian hai chiều (2D) và ba chiều (3D). Trường âm thanh được dự đoán tại và xung quanh điểm tham chiếu trong không gian bởi chuỗi Fourier-Bessel. Sáng chế ứng dụng kỹ thuật làm cong không gian (12, 13, 14, 16) để thay đổi nội dung không gian và/hoặc tái tạo thông tin trường âm thanh mà được thu hoặc được tạo ra như phép biểu diễn HOA. Các đặc tính làm cong khác nhau có thể sử dụng cho các trường âm thanh 2D và 3D. Việc làm cong được thực hiện trong miền không gian mà không cần thực hiện việc phân tích hoặc phân giải môi trường. Các hệ số của phép biểu diễn HOA đầu vào với bậc định sẵn được giải mã thành các trọng số hoặc các tín hiệu đầu vào của các loa phóng thanh (ảo) có vị trí cách đều nhau.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến phương pháp và thiết bị để thay đổi các vị trí tương đối của các đối tượng âm thanh được chứa trong phép biểu diễn Ambisonics bậc cao hơn hai chiều hoặc ba chiều của môi trường âm thanh.

Tình trạng kỹ thuật của sáng chế

Ambisonics bậc cao hơn (HOA) là phép biểu diễn của các trường âm thanh không gian mà hỗ trợ việc thu, quản lý, ghi, truyền và phát lại của môi trường âm thanh phức tạp với độ phân giải không gian ưu việt, cả trong 2D và 3D. Trường âm thanh được dự đoán tại và xung quanh điểm tham chiếu trong không gian bởi chuỗi Fourier-Bessel.

Có tồn tại chỉ một số kỹ thuật hạn chế để quản lý bố trí không gian của môi trường âm thanh thu được bằng các kỹ thuật HOA. Về nguyên tắc, có hai cách:

A) Phân tích môi trường âm thanh thành các đối tượng âm thanh riêng biệt và thông tin vị trí được kết hợp, ví dụ thông qua DirAC, và cấu thành môi trường mới với các tham số vị trí được quản lý. Nhược điểm đó là việc phân tích môi trường có thể có lỗi và phức tạp là bắt buộc.

B) Nội dung của phép biểu diễn HOA có thể được thay đổi thông qua biến đổi tuyến tính của các vector HOA. Ở đây, chỉ việc quay, phản chiếu, và sự nhân mạnh của các chiều phía trước/phía sau được đề xuất. Tất cả điều này đã được biết đến, các kỹ thuật cải biến dựa trên biến đổi giữ cố định vị trí tương đối của các đối tượng trong môi trường âm thanh.

Để quản lý hoặc thay đổi các nội dung của môi trường, việc làm cong không gian được đề xuất, bao gồm quay và phản chiếu của các trường âm thanh HOA, và thay đổi ảnh hưởng của các chiều cụ thể:

Ấn phẩm: G.J. Barton, M.A. Gerzon, "Ambisonic Decoders for HDTV", AES Convention, 1992;

Ấn phẩm: J. Daniel, "Représentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimédia", PhD thesis, Université de Paris 6, 2001, Paris, France;

Ấn phẩm: M. Chapman, Ph. Cotterell, "Towards a Comprehensive Account of Valid Ambisonic Transformations", Ambisonics Symposium, 2009, Graz, Austria.

Ấn phẩm H. Pomberger, F. Zotter, "Warping of 3D Ambisonic Recordings", Ambisonics Symposium 2011, 2-3 June 2011, Lexington, KY, USA mô tả sự thay đổi vị trí tương đối của các đối tượng âm thanh nằm trong phép biểu diễn Ambisonics bậc cao hơn của môi trường âm thanh. Việc làm cong tập trung vào góc cực được đề xuất, và nhấn mạnh rằng việc làm cong dựa trên góc cực là thích đáng.

Bản chất kỹ thuật của sáng chế

Vấn đề được giải quyết bởi sáng chế là để hỗ trợ việc thay đổi các vị trí tương đối của các đối tượng âm thanh nằm trong môi trường âm thanh dựa trên HOA, mà không cần thiết phân tích thành phần của môi trường. Vấn đề này được giải quyết bởi phương pháp được bộc lộ trong điểm 1 yêu cầu bảo hộ. Thiết bị mà sử dụng phương pháp này được bộc lộ trong điểm 2 yêu cầu bảo hộ.

Sáng chế sử dụng việc làm cong không gian để thay đổi nội dung không gian và/hoặc sự tái tạo của thông tin trường âm thanh mà thu được hoặc được tạo ra như phép biểu diễn Ambisonics bậc cao hơn. Việc làm cong không gian trong miền HOA thể hiện cả, phương pháp đa bước hoặc, hiệu quả về mặt tính toán hơn, phép nhân ma trận tuyến tính đơn bước. Các đặc tính làm cong khác nhau là có thể sử dụng cho các trường âm thanh 2D và 3D.

Việc làm cong được thực hiện trong miền không gian mà không thực hiện việc phân tích hoặc phân giải môi trường. Các hệ số HOA đầu vào với bậc đưa ra được giải mã thành các trọng số hoặc các tín hiệu đầu vào của các loa phóng thanh (ảo) được đặt tại vị trí cách đều nhau.

Xử lý làm cong không gian theo sáng chế có một vài ưu điểm:

- rất linh hoạt do các mức độ tự do trong việc tham số hóa;
- có thể được triển khai theo cách rất hiệu quả, tức là với độ phức tạp tương đối thấp;
- không yêu cầu bất kỳ phân tích hay phân giải môi trường.

Về nguyên tắc, phương pháp theo sáng chế là thích hợp để thay đổi các vị trí tương đối của các đối tượng âm thanh nằm trong phép biểu diễn Ambisonics bậc cao

hơn HOA hai chiều hoặc ba chiều của môi trường âm thanh, trong đó vector đầu vào $\mathbf{A}_{in}$ với chiều O_{in} xác định các hệ số của chuỗi Fourier của tín hiệu đầu vào và vector đầu ra $\mathbf{A}_{out}$ với chiều O_{out} xác định các hệ số của chuỗi Fourier của tín hiệu đầu ra được thay đổi tương ứng, phương pháp nêu trên bao gồm các bước:

- giải mã vector đầu vào $\mathbf{A}_{in}$ nêu trên của các hệ số HOA đầu vào thành các tín hiệu đầu vào $\mathbf{s}_{in}$ trong miền không gian đối với các vị trí loa phóng thanh đặt tại vị trí cách đều nhau sử dụng nghịch đảo Ψ_1^{-1} của ma trận chế độ Ψ_1 bằng cách tính toán $\mathbf{s}_{in} = \Psi_1^{-1} \mathbf{A}_{in}$;

- làm cong và mã hóa trong miền không gian các tín hiệu đầu vào $\mathbf{s}_{in}$ nêu trên thành vector đầu ra $\mathbf{A}_{out}$ nêu trên của các hệ số HOA đầu ra được áp dụng bằng cách tính toán $\mathbf{A}_{out} = \Psi_2 \mathbf{s}_{in}$, trong đó các vector chế độ của ma trận chế độ Ψ_2 được thay đổi theo hàm làm cong $f(\phi)$ bởi đó các góc của các vị trí loa phóng thanh gốc được ánh xạ một-một thành các góc mục tiêu của các vị trí loa phóng thanh mục tiêu trong vector đầu ra $\mathbf{A}_{out}$ nêu trên.

Về nguyên tắc thiết bị theo sáng chế là thích hợp để thay đổi các vị trí tương đối của các đối tượng âm thanh nằm trong phép biểu diễn Ambisonics bậc cao hơn HOA hai chiều hoặc ba chiều của môi trường âm thanh, trong đó vector đầu vào $\mathbf{A}_{in}$ với chiều O_{in} xác định các hệ số của chuỗi Fourier của tín hiệu đầu vào và vector đầu ra $\mathbf{A}_{out}$ với chiều O_{out} xác định các hệ số của chuỗi Fourier của tín hiệu đầu ra được thay đổi tương ứng, thiết bị nêu trên bao gồm:

- phương tiện được áp dụng để giải mã vector đầu vào $\mathbf{A}_{in}$ nêu trên của các hệ số HOA đầu vào thành các tín hiệu đầu vào $\mathbf{s}_{in}$ trong miền không gian đối với các vị trí loa phóng thanh đặt tại vị trí cách đều nhau sử dụng nghịch đảo Ψ_1^{-1} của ma trận chế độ Ψ_1 bằng cách tính toán $\mathbf{s}_{in} = \Psi_1^{-1} \mathbf{A}_{in}$;

- phương tiện được áp dụng để làm cong và mã hóa trong miền không gian các tín hiệu đầu vào $\mathbf{s}_{in}$ nêu trên thành vector đầu ra $\mathbf{A}_{out}$ nêu trên của các hệ số HOA đầu ra được áp dụng bằng cách tính toán $\mathbf{A}_{out} = \Psi_2 \mathbf{s}_{in}$, trong đó các vector chế độ của ma trận chế độ Ψ_2 được thay đổi theo hàm làm cong $f(\phi)$ bởi đó các góc của các vị trí loa phóng thanh gốc được ánh xạ một-một thành các góc mục tiêu của các vị trí loa phóng

thanh mục tiêu trong vector đầu ra $\mathbf{A}_{\text{out}}$ nêu trên.

Các phương án bổ sung hiệu quả của sáng chế được mô tả trong các điểm yêu cầu bảo hộ phụ thuộc tương ứng.

Mô tả vắn tắt các hình vẽ

Các phương án lấy làm ví dụ của sáng chế được mô tả có viện dẫn tới các hình vẽ kèm theo, trong đó:

Fig.1 là hình vẽ thể hiện nguyên tắc làm cong trong miền không gian;

Fig.2 là hình vẽ thể hiện ví dụ của việc làm cong không gian với $N_{\text{in}} = 3$, $N_{\text{out}} = 12$ và hàm làm cong $f(\phi) = \phi + 2 \operatorname{atan}\left(\frac{a \sin \phi}{1 - a \cos \phi}\right)$ with $a = -0,4$;

Fig.3 là hình vẽ thể hiện các méo dạng ma trận đối với các hàm làm cong khác nhau và các bậc "bên trong" N_{warp} .

Mô tả chi tiết sáng chế

Trong phần sau đây, nhằm có thể hiểu được, việc làm cong không gian theo sáng chế được mô tả đối với thiết lập hai chiều, phép biểu diễn HOA dựa trên các hàm điều hòa tròn, và giả thiết rằng trường âm thanh được bảo diễn bao gồm chỉ các sóng âm thanh phẳng. Sau đó phần mô tả được mở rộng tới các trường hợp ba chiều, dựa trên các hàm điều hòa cầu.

Chú thích

Trong lý thuyết Ambisonics, trường âm thanh tại và xung quanh điểm cụ thể trong không gian được mô tả bởi chuỗi Fourier-Bessel bị chặn. Nói chung, điểm tham chiếu được giả thiết là tại gốc của hệ tọa độ được lựa chọn.

Đối với ứng dụng ba chiều sử dụng các tọa độ cầu, chuỗi Fourier với các hệ số A_n^m cho tất cả các chỉ số được xác định $n = 0, 1, \dots, N$ và $m = -n, \dots, n$ mô tả áp suất của trường âm thanh tại góc phương vị ϕ , độ nghiêng θ và khoảng cách r từ gốc:

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

trong đó, k là số sóng và $j_n(kr) Y_n^m(\phi, \theta)$ là hàm hạch của chuỗi Fourier-Bessel mà liên quan mật thiết tới hàm điều hòa cầu đối với chiều được xác định bởi θ và ϕ . Nhằm

thuận tiện, trong các hệ số HOA sau đây A_n^m được sử dụng với sự xác định $A_n^m = C_n^m j_n(kr)$. Đối với bậc cụ thể N số lượng hệ số trong chuỗi Fourier-Bessel là $O = (N + 1)^2$.

Đối với ứng dụng hai chiều sử dụng các tọa độ tròn, các hàm hạch phụ thuộc vào chỉ góc phương vị ϕ . Tất cả các hệ số với $m \neq n$ có giá trị là không và có thể được bỏ qua. Do đó, số lượng hệ số HOA được làm giảm xuống chỉ $O = 2N + 1$. Ngoài ra, độ nghiêng $\theta = \pi/2$ là cố định. Lưu ý rằng đối với trường hợp 2D và đối với phân phối đều hoàn hảo của các đối tượng âm thanh trên đường tròn, tức là với $\phi_i = i \frac{2\pi}{O}$, các vector chế độ trong ma trận chế độ Ψ là đồng nhất với các hàm hạch của biến đổi Fourier rời rạc DFT đã biết.

Các quy ước khác nhau tồn tại đối với sự xác định của các hàm hạch mà cũng dẫn đến các xác định khác nhau của các hệ số Ambisonics A_n^m . Tuy nhiên, định nghĩa chính xác không đóng vai trò đối với phần mô tả và các đặc tính cơ bản của các kỹ thuật làm cong không gian được mô tả trong đơn này.

"Tín hiệu" HOA bao gồm vector $\mathbf{A}$ của các hệ số Ambisonics đối với mỗi thời điểm. Đối với thiết lập hai chiều tức là thiết lập tròn, thành phần và thứ tự điển hình của vector hệ số là

$$\mathbf{A}_{2D} = (A_N^{-N}, A_{N-1}^{-N+1}, \dots, A_1^{-1}, A_0^0, A_1^1, \dots, A_N^N)^T. \quad (2)$$

Đối với thiết lập ba chiều, cầu, thứ tự thường dùng của các hệ số là khác:

$$\mathbf{A}_{3D} = (A_0^0, A_1^{-1}, A_1^0, A_1^1, A_2^{-2}, \dots, A_N^N)^T. \quad (3)$$

Việc mã hóa của các phép biểu diễn HOA thao tác theo cách tuyến tính và do đó các hệ số HOA đối với nhiều đối tượng âm thanh riêng biệt có thể được cộng lại để thu được các hệ số HOA của trường âm thanh cuối cùng.

Mã hóa phẳng

Mã hóa phẳng của các đối tượng âm thanh từ một vài chiều có thể được thực hiện trực tiếp trong đại số vector. "Mã hóa" nghĩa là bước để thu được vector của các hệ số HOA $\mathbf{A}(k, l)$ tại thời điểm l và số sóng k từ thông tin về các phân góp áp suất $s_i(k, l)$ của các đối tượng âm thanh riêng biệt ($i = 0 \dots M - 1$) tại cùng thời điểm l , cộng với

các chiều ϕ_i và θ_i mà từ đó các sóng âm thanh đang đi đến tại gốc của hệ tọa độ

$$\mathbf{A}(k, l) = \Psi \cdot \mathbf{s}(k, l) . \quad (4)$$

Nếu thiết lập hai chiều và thành phần của các vector HOA như được xác định trong phương trình (2) được giả thiết, ma trận chế độ Ψ được xây dựng từ các vector chế độ $\mathbf{Y}(\phi) = (Y_N^{-N}, \dots, Y_0^0, \dots, Y_N^N)^T$. Cột thứ i của ma trận chế độ Ψ chứa vector chế độ theo chiều ϕ_i của đối tượng âm thanh thứ i

$$\Psi = (\mathbf{Y}(\phi_0), \mathbf{Y}(\phi_1), \dots, \mathbf{Y}(\phi_{M-1})) . \quad (5)$$

Như được xác định nêu trên, việc mã hóa của phép biểu diễn HOA có thể được hiểu như là việc biến đổi tần số-không gian vì các tín hiệu đầu vào (các đối tượng âm thanh) được phân phối theo không gian. Biến đổi này bởi ma trận Ψ có thể được nghịch đảo mà không tổn hao thông tin nếu chỉ số lượng đối tượng âm thanh là đồng nhất với số lượng hệ số HOA, tức là nếu $M = O$, và nếu các chiều ϕ_i được trải rộng hợp lý xung quanh đường tròn đơn vị. Về mặt toán học, các điều kiện đối với khả năng nghịch đảo đó là ma trận chế độ Ψ phải là vuông ($O \times O$) và khả nghịch.

Giải mã phẳng

Bằng cách giải mã, các tín hiệu điều khiển của các loa phóng thanh thực và ảo thu được mà phải được áp dụng để chơi lại chính xác trường âm thanh mong muốn như được mô tả bởi các hệ số HOA đầu vào. Việc giải mã này phụ thuộc vào số M và các vị trí của các loa phóng thanh. Ba trường hợp quan trọng sau đây phải được phân biệt (lưu ý: các trường hợp này được đơn giản hóa trong đó chúng được xác định thông qua "số lượng loa phóng thanh", giả thiết rằng chúng được thiết lập theo cách hình học hợp lý. Chính xác hơn, việc xác định được thực hiện thông qua hạng của ma trận chế độ của việc thiết lập loa phóng thanh mục tiêu). Trong các quy tắc giải mã ví dụ được thể hiện dưới đây, quy tắc giải mã khớp chế độ được áp dụng, nhưng các quy tắc giải mã khác có thể được sử dụng mà có thể dẫn đến các quy tắc giải mã khác nhau đối với ba trường hợp.

- Trường hợp được xác định trên: Số lượng các loa phóng thanh là lớn hơn số lượng hệ số HOA, tức là $M > O$. Trong trường hợp này, không có nghiệm duy nhất đối

với vấn đề giải mã tồn tại, nhưng dải của các nghiệm cho phép tồn tại mà nằm trong không gian con chiều $M - O$ của không gian chiều M của tất cả các nghiệm có khả năng. Một cách điển hình, nghịch đảo giả của ma trận chế độ Ψ của thiết lập loa phóng thanh cụ thể được sử dụng để xác định các tín hiệu loa phóng thanh $\mathbf{s}$, $\mathbf{s} = \Psi^T (\Psi \Psi^T)^{-1} \mathbf{A}$. (6)

Nghiệm này phân phát các tín hiệu loa phóng thanh với tổng công suất phát lại cực tiểu $\mathbf{s}^T \mathbf{s}$ (xem ví dụ ẩn phẩm: L.L.Scharf, "Statistical Signal Processing. Detection, Estimation, và Time Series Analysis", Addison-Wesley Publishing Company, Reading, Massachusetts, 1990). Đối với các thiết lập thông thường của các loa phóng thanh (mà có thể đạt được dễ dàng trong trường hợp 2D) phép toán ma trận sinh ra ma trận đơn vị, và quy tắc giải mã từ phương trình (6) làm đơn giản thành $\mathbf{s} = \Psi^T \mathbf{A}$.

- Trường hợp được xác định: Số lượng các loa phóng thanh là bằng số lượng hệ số HOA. Chính xác một nghiệm duy nhất đối với vấn đề giải mã tồn tại, mà được xác định bởi nghịch đảo Ψ^{-1} của ma trận chế độ Ψ : $\mathbf{s} = \Psi^{-1} \mathbf{A}$. (7)

- Trường hợp được xác định dưới: Số lượng M của các loa phóng thanh là thấp hơn số lượng O của các hệ số HOA. Do đó, vấn đề toán học của giải mã trường âm thanh được xác định dưới và không có nghiệm duy nhất, chính xác tồn tại. Thay vì đó, việc tối ưu số học phải được sử dụng để xác định các tín hiệu loa phóng thanh có thể khớp nhất với trường âm thanh mong muốn.

Việc chính quy hóa có thể được áp dụng để thu được nghiệm ổn định, ví dụ bởi công thức

$$\mathbf{s} = \Psi^T (\Psi \Psi^T + \lambda \mathbf{I})^{-1} \mathbf{A}, \quad (8)$$

trong đó, $\mathbf{I}$ ký hiệu ma trận đơn vị và hệ số vô hướng λ xác định lượng chính quy hóa. Như là ví dụ, λ có thể được thiết lập là giá trị trung bình của các giá trị đặc trưng của $\Psi \Psi^T$.

Các mẫu chùm cuối cùng có thể gần điểm tối ưu do nói chung các mẫu chùm thu được bằng phương pháp này là định hướng thái quá, và nhiều thông tin âm thanh sẽ bị miêu tả không đúng mức.

Đối với tất cả các ví dụ giải mã được mô tả nêu trên giả thiết được đưa ra rằng các loa phóng thanh phát ra các sóng phẳng. Các loa phóng thanh thực tế có các đặc tính chơi lại khác nhau, các đặc tính mà quy tắc giải mã sẽ cần quan tâm đến.

Làm cong cơ bản

Nguyên tắc của việc làm cong không gian theo sáng chế như được minh họa trên Fig.1a. Việc làm cong được thực hiện trong miền không gian. Do đó, đầu tiên các hệ số HOA đầu vào $\mathbf{A}_{in}$ với bậc N_{in} và chiều O_{in} được giải mã trong bước/tầng 12 thành các trọng số hoặc các tín hiệu đầu vào $\mathbf{s}_{in}$ đối với các loa phóng thanh (ảo) đặt tại vị trí cách đều nhau. Đối với bước giải mã này có ưu điểm để áp dụng bộ giải mã được xác định, tức là bộ giải mã mà để số lượng O_{warp} của các loa phóng thanh ảo là bằng hoặc lớn hơn số lượng hệ số HOA O_{in} . Đối với trường hợp sau (nhiều loa phóng thanh hơn so với các hệ số HOA), bậc hoặc chiều của vector $\mathbf{A}_{in}$ của các hệ số HOA có thể được mở rộng dễ dàng bằng cách bổ sung trong bước/tầng 11 các hệ số không đối với các bậc cao hơn. Chiều của vector mục tiêu $\mathbf{s}_{in}$ sẽ được ký hiệu bởi O_{warp} trong phần sau đây.

$$\text{Quy tắc giải mã là } \mathbf{s}_{in} = \Psi_1^{-1} \mathbf{A}_{in} . \quad (9)$$

Các vị trí ảo của các tín hiệu loa phóng thanh sẽ là cách đều nhau, ví dụ $\phi_i = i \cdot 2\pi/O_{warp}$ đối với trường hợp hai chiều. Nhờ đó đảm bảo rằng ma trận chế độ Ψ_1 có điều kiện tốt để xác định ma trận giải mã .

Tiếp theo, các vị trí của các loa phóng thanh ảo được thay đổi trong xử lý "làm cong" theo các đặc tính làm cong mong muốn. Xử lý làm cong trong bước/tầng 14 được kết hợp với việc mã hóa vector mục tiêu $\mathbf{s}_{in}$ (hoặc $\mathbf{s}_{out}$, một cách tương ứng) sử dụng ma trận chế độ Ψ_2 , kết quả thu được vector $\mathbf{A}_{out}$ của các hệ số HOA được làm cong với chiều O_{warp} hoặc, sau đó bước xử lý tiếp theo được mô tả dưới đây, với chiều O_{out} . Về nguyên tắc, các đặc tính làm cong có thể được xác định đầy đủ bằng việc ánh xạ một-một của các góc nguồn tới các góc mục tiêu, tức là đối với mỗi góc nguồn $\phi_{in}=0...2\pi$ và có thể $\theta_{in}=0...2\pi$ góc mục tiêu được xác định, bằng cách đó đối với trường hợp 2D

$$\phi_{out} = f(\phi_{in}) \quad (10)$$

và đối với trường hợp 3D

$$\phi_{\text{out}} = f_{\phi}(\phi_{\text{in}}, \theta_{\text{in}}) \quad (11)$$

$$\theta_{\text{out}} = f_{\theta}(\phi_{\text{in}}, \theta_{\text{in}}) . \quad (12)$$

Về khái niệm, việc định hướng lại (ảo) này có thể được so sánh với việc di chuyển vật lý các loa phóng thanh tới các vị trí mới.

Một vấn đề sẽ được sinh ra bởi thủ tục này đó là khoảng cách giữa các loa phóng thanh lân cận tại các góc nào đó bị thay đổi theo gradien của hàm làm cong $f(\phi)$ (điều này được mô tả đối với trường hợp 2D trong phần sau đây): nếu gradien của $f(\phi)$ là lớn hơn một, không gian góc giống nhau trong trường âm thanh được làm cong sẽ được chiếm giữ bởi ít "loa phóng thanh" hơn so với trong trường âm thanh gốc, và ngược lại. Nói cách khác, mật độ D_s của các loa phóng thanh thao tác theo $D_s(\phi) = \frac{1}{\frac{df(\phi)}{d\phi}}$.

(13)

Mặt khác, điều này nghĩa là việc làm cong không gian làm thay đổi độ cân bằng âm thanh xung quanh người nghe. Các vùng trong đó mật độ loa phóng thanh được tăng lên, tức là để cho $D_s(\phi) > 1$, sẽ trở nên trội hơn, và các vùng trong đó $D_s(\phi) < 1$ sẽ trở nên ít trội hơn.

Như là sự lựa chọn, phụ thuộc vào các yêu cầu của ứng dụng, việc thay đổi nêu trên của mật độ loa phóng thanh có thể được thực hiện bằng cách áp dụng hàm tăng ích $g(\phi)$ đối với các tín hiệu đầu ra loa phóng thanh ảo $\mathbf{s}_{\text{in}}$ trong bước/tầng lấy trọng số 13, kết quả thu được tín hiệu $\mathbf{s}_{\text{out}}$. Về nguyên tắc, bất kỳ hàm trọng số $g(\phi)$ có thể được chỉ rõ. Một biến thể hiệu quả đặc biệt được xác định theo kinh nghiệm là tỷ lệ với đạo hàm của hàm làm cong $f(\phi)$: $g(\phi) = \frac{1}{D_s(\phi)} = \frac{df(\phi)}{d\phi}$.

(14)

Với hàm trọng số cụ thể này, dưới giả thiết của bậc đầu vào và bậc đầu ra cao hợp lý (xem phần Làm thế nào để thiết lập các bậc HOA dưới đây), biên độ của hàm quét tại góc được làm cong cụ thể $f(\phi)$ được giữ bằng với hàm quét gốc tại góc gốc ϕ . Do đó, độ cân bằng (biên độ) âm thanh đồng nhất trên góc mở đạt được.

Ngoài hàm trọng số ví dụ nêu trên, các hàm trọng số khác có thể được sử dụng, ví dụ để thu được công suất cân bằng trên góc mở.

Cuối cùng, trong bước/tầng 14 các tín hiệu loa phóng thanh ảo được lấy trọng số

được làm cong và được mã hóa lại với ma trận chế độ Ψ_2 bằng cách thực hiện $\Psi_2 \mathbf{s}_{\text{out}}$. Ψ_2 bao gồm các vector chế độ khác nhau so với Ψ_1 , theo hàm làm cong $f(\phi)$. Kết quả là phép biểu diễn HOA chiều $\mathbf{O}_{\text{warp}}$ của trường âm thanh được làm cong.

Nếu bậc hoặc chiều của phép biểu diễn HOA mục tiêu thấp hơn so với bậc của ma trận mã hóa Ψ_2 (xem phần Làm thế nào để thiết lập các bậc HOA sau đây), một vài (tức là một phần) của hệ số được làm cong phải được loại bỏ (được tách) trong bước/tầng 15. Nói chung, thao tác tách này có thể được mô tả bằng thao tác tạo cửa sổ: vector được mã hóa $\Psi_2 \mathbf{s}_{\text{out}}$ được nhân với vector cửa sổ $\mathbf{w}$ mà bao gồm các hệ số không đối với các bậc cao nhất mà sẽ bị loại bỏ, mà phép nhân có thể được xem như là biểu diễn trọng số. Trong trường hợp đơn giản nhất, cửa sổ hình chữ nhật có thể được áp dụng, tuy nhiên, các cửa sổ phức tạp hơn có thể được sử dụng như được mô tả trong phần 3 của ấn phẩm: M.A. Poletti, "A Unified Theory of Horizontal Holographic Sound Systems", Journal of the Audio Engineering Society, 48(12), pp.1155-1182, 2000, hoặc các cửa sổ “trong pha” hoặc “cực đại r_E ” từ phần 3.3.2 của luận án tiến sỹ nêu trên của J. Daniel.

Các hàm làm cong đối với 3D

Khái niệm của hàm làm cong $f(\phi)$ và hàm trọng số được kết hợp $g(\phi)$ được mô tả nêu trên đối với trường hợp hai chiều. Phần sau đây là phần mở rộng đối với trường hợp ba chiều mà là phức tạp hơn vì chiều cao hơn và vì hình học cầu phải được áp dụng. Hai trường hợp được đơn giản hóa được đưa ra, cả hai trong số chúng cho phép chỉ rõ việc làm cong không gian mong muốn bởi các hàm làm cong một chiều $f(\phi)$ hoặc $f(\theta)$.

Trong việc làm cong không gian theo các kinh độ, việc làm cong không gian được thực hiện như là hàm của chỉ góc phương vị ϕ . Trường hợp này là hoàn toàn tương tự như trường hợp hai chiều được đưa ra nêu trên. Hàm làm cong được xác định hoàn toàn bởi

$$\theta_{\text{out}} = f_{\theta}(\theta_{\text{in}}, \phi_{\text{in}}) \stackrel{!}{=} \theta_{\text{in}} \quad (15)$$

$$\phi_{\text{out}} = f_{\phi}(\theta_{\text{in}}, \phi_{\text{in}}) \stackrel{!}{=} f_{\phi}(\phi_{\text{in}}) . \quad (16)$$

Nhờ đó, các hàm làm cong tương tự có thể được áp dụng như đối với trường hợp hai chiều. Việc làm cong không gian có ảnh hưởng lớn nhất đối với các đối tượng âm

thanh trên đường xích đạo, trong khi nó có ảnh hưởng ít nhất đối với các đối tượng âm thanh tại các cực của hình cầu.

Mật độ của các đối tượng âm thanh (được làm cong) trên hình cầu chỉ phụ thuộc vào góc phương vị. Do đó hàm trọng số đối với mật độ cố định là $g(\theta) = \frac{df_{\phi}(\phi)}{d\phi}$.
(17)

Sự định hướng tự do của các đặc tính làm cong cụ thể trong không gian là khả thi bằng cách quay (ảo) hình cầu trước khi áp dụng việc làm cong và quay ngược ra sau.

Trong việc làm cong không gian dọc theo các vĩ tuyến, việc làm cong không gian được cho phép chỉ dọc theo các kinh tuyến. Hàm làm cong được xác định bởi

$$\theta_{\text{out}} = f_{\theta}(\theta_{\text{in}}, \phi_{\text{in}}) \stackrel{!}{=} f_{\theta}(\theta_{\text{in}}) \quad (18)$$

$$\phi_{\text{out}} = f_{\phi}(\theta_{\text{in}}, \phi_{\text{in}}) \stackrel{!}{=} \phi_{\text{in}} . \quad (19)$$

Đặc tính quan trọng của hàm làm cong này trên hình cầu đó là, mặc dù góc phương vị được giữ không đổi, khoảng cách góc của hai điểm theo chiều phương vị có thể thay đổi do sự thay đổi của góc nghiêng. Lý do đó là khoảng cách góc giữa hai kinh tuyến là lớn nhất tại đường xích đạo, nhưng nó tiến tới không tại hai cực. Thực tế này phải được tính đến đối với hàm trọng số.

Khoảng cách góc c của hai điểm A và B có thể được xác định bởi quy tắc cosin của hình học cầu, tham chiếu ấn phẩm: Eq.(3.188c) trong I.N. Bronstein, K.A. Semendjajew, G. Musiol, H. Mühlig, "Taschenbuch der Mathematik", Verlag Harri Deutsch, Thun, Frankfurt/Main, 5th edition, 2000:

$$\cos c = \cos \theta_A \cos \theta_B + \sin \theta_A \sin \theta_B \cos \phi_{AB} , \quad (20)$$

trong đó, ϕ_{AB} ký hiệu góc phương vị giữa hai điểm A và B. Liên quan đến khoảng cách góc giữa hai điểm tại cùng độ nghiêng θ , phương trình này đơn giản hóa thành

$$c = \arccos[(\cos \theta_A)^2 + (\sin \theta_A)^2 \cos \phi_{\epsilon}] . \quad (21)$$

Công thức này có thể được áp dụng để thu được khoảng cách góc giữa điểm trong không gian và điểm khác mà là bởi góc phương vị nhỏ ϕ_{ϵ} . “Nhỏ” nghĩa là đủ nhỏ để có thể sử dụng được trong các ứng dụng thực tế mà không phải là không, về lý thuyết, giá

trị giới hạn $\phi_\varepsilon \rightarrow 0$. Tỷ lệ giữa các khoảng cách góc này trước và sau khi làm cong đưa ra hệ số mà mật độ của các đối tượng âm thanh theo chiều ϕ thay đổi:

$$\frac{c_{\text{out}}}{c_{\text{in}}} = \frac{\arccos((\cos \theta_{\text{out}})^2 + (\sin \theta_{\text{out}})^2 \cos \phi_\varepsilon)}{\arccos((\cos \theta_{\text{in}})^2 + (\sin \theta_{\text{in}})^2 \cos \phi_\varepsilon)}. \quad (22)$$

Cuối cùng, hàm trọng số là tích của hai hàm trọng số theo chiều ϕ và theo chiều θ

$$g(\theta, \phi) = \frac{df_\theta(\theta)}{d\theta} \cdot \frac{\arccos((\cos f_\theta(\theta_{\text{in}}))^2 + (\sin f_\theta(\theta_{\text{in}}))^2 \cos \phi_\varepsilon)}{\arccos((\cos \theta_{\text{in}})^2 + (\sin \theta_{\text{in}})^2 \cos \phi_\varepsilon)}. \quad (23)$$

Một lần nữa, như trong trường hợp trước đó, sự định hướng tự do của các đặc tính làm cong cụ thể trong không gian là khả thi bằng cách quay.

Xử lý đơn bước

Các bước được đưa ra có viện dẫn đến Fig.1a, tức là mở rộng bậc, giải mã, lấy trọng số, làm cong+mã hóa và tách theo thứ tự, là các thao tác tuyến tính cần thiết. Do đó, các chuỗi thao tác này có thể được thay thế bằng phép nhân của các hệ số HOA đầu vào với ma trận đơn trong bước/tầng 16 như được thể hiện trên Fig.1b. Bỏ qua thao tác mở rộng và tách, ma trận biến đổi $O_{\text{warp}} \times O_{\text{warp}}$ toàn phần $\mathbf{T}$ được xác định như là

$$\mathbf{T} = \text{diag}(\mathbf{w}) \Psi_2 \text{diag}(\mathbf{g}) \Psi_1^{-1}, \quad (24)$$

trong đó, $\text{diag}(\cdot)$ thể hiện ma trận chéo mà có các giá trị của argumen vector của nó như là các thành phần của đường chéo chính, $\mathbf{g}$ là hàm trọng số, và $\mathbf{w}$ là vector của số để thực hiện việc tách được mô tả nêu trên, tức là, từ hai hàm trọng số để thực hiện việc tách và bản thân việc tách các hệ số được thực hiện trong bước/tầng 15, vector của số $\mathbf{w}$ trong phương trình (24) chỉ đóng vai trò cho việc lấy trọng số.

Hai áp dụng của các thứ tự trong phương pháp đa bước, tức là mở rộng bậc trước giải mã và tách các hệ số HOA sau khi mã hóa, cũng có thể được tích hợp vào ma trận biến đổi $\mathbf{T}$ bằng cách loại bỏ các cột và/hoặc các dòng tương ứng. Nhờ đó, ma trận có kích cỡ $O_{\text{out}} \times O_{\text{in}}$ thu được mà có thể được áp dụng trực tiếp tới các vector HOA đầu vào. Sau đó, thao tác làm cong trở thành $\mathbf{A}_{\text{out}} = \mathbf{T} \mathbf{A}_{\text{in}}$. (25)

Một cách thuận lợi, do việc làm giảm hiệu quả của các chiều của ma trận biến đổi $\mathbf{T}$ từ $O_{\text{warp}} \times O_{\text{warp}}$ thành $O_{\text{out}} \times O_{\text{in}}$, độ phức tạp tính toán được yêu cầu để thực hiện

xử lý đơn bước theo Fig.1b là thấp hơn đáng kể so với độ phức tạp tính toán được yêu cầu đối với phương pháp đa bước của Fig.1a, mặc dù xử lý đơn bước mang đến các kết quả hoàn toàn tương tự. Cụ thể, nó tránh được các méo dạng mà có thể sinh ra nếu xử lý đa bước được thực hiện với bậc thấp hơn N_{warp} của các tín hiệu tạm thời của nó (xem chi tiết hơn trong phần Làm thế nào để thiết lập các bậc HOA dưới đây).

Tình trạng kỹ thuật: quay và phản chiếu

Các việc quay và phản chiếu của trường âm thanh có thể được xem xét như là phạm trù con “đơn giản” của việc làm cong không gian. Đặc tính đặc biệt của các biến đổi này đó là vị trí tương đối của các đối tượng âm thanh với nhau *không bị thay đổi*. Điều này nghĩa là, đối tượng âm thanh mà nằm tại ví dụ 30° sang bên phải của đối tượng âm thanh khác trong môi trường âm thanh gốc sẽ nằm tại 30° sang bên phải của cùng đối tượng âm thanh trong môi trường âm thanh được quay. Đối với việc phản chiếu, chỉ dấu thay đổi nhưng các khoảng cách góc vẫn là như cũ.

Các thuật toán và ứng dụng để quay và phản chiếu của thông tin trường âm thanh được nghiên cứu và được mô tả ví dụ trong các bài viết nêu trên của Barton/Gerzon và J.Daniel, và trong ấn phẩm: M. Noisternig, A. Sontacchi, Th. Musil, R. Höldrich, "A 3D Ambisonic Based Binaural Sound Reproduction System", Proc. of the AES 24th Intl. Conf. on Multichannel Audio, Banff, Canada, 2003, và trong H. Pomberger, F. Zotter, "An Ambisonics Format for Flexible Playback Layouts", 1st Ambisonics Symposium, Graz, Austria, 2009.

Các phương pháp này được dựa trên các biểu thức phân tích đối với các ma trận quay. Ví dụ, việc quay của trường âm thanh hình tròn (trường hợp 2D) bằng góc bất kỳ α có thể được thực hiện bằng pháp nhân với ma trận làm cong T_α trong đó chỉ tập con của các hệ số là khác không:

$$T_\alpha(\mu, \nu) = \begin{cases} \cos(-\alpha(\mu - (O + 1)/2)) & ; \quad \nu = \mu \\ \sin(-\alpha(\mu - (O + 1)/2)) & ; \quad \nu = N - \mu + 1 \\ 0 & ; \quad \text{cách khác} \end{cases} \quad (26)$$

Như trong ví dụ này, tất cả các ma trận làm cong cho các thao tác quay và/hoặc phản chiếu có các đặc tính đặc biệt mà chỉ các hệ số của cùng bậc n là ảnh hưởng lẫn nhau. Do đó các ma trận làm cong này được bố trí rất rải rác, và bậc đầu ra N_{out} có thể

bằng bậc đầu vào N_{in} mà không làm mất bất kỳ thông tin không gian nào.

Có nhiều ứng dụng được quan tâm, mà việc quay hoặc phản chiếu của thông tin trường âm thanh được yêu cầu. Một ví dụ đó là việc phát lại của các trường âm thanh thông qua các tai nghe bằng hệ thống bám theo đầu. Thay vì nội suy các HRTF (hàm truyền liên quan đến đầu) theo (các) góc quay của đầu, có ưu điểm để quay trước trường âm thanh theo vị trí của đầu và để sử dụng các HRTF được cố định cho việc phát lại thực tế. Xử lý này được mô tả trong bài viết Noisternig/Sontacchi/Musil/Höldrach nêu trên.

Ví dụ khác đã được mô tả trong bài viết Pomberger/Zotter nêu trên trong ngữ cảnh của mã hóa thông tin trường âm thanh. Có thể cường bức vùng không gian mà được mô tả bởi các vectơ HOA đối với các phần cụ thể của hình tròn (trường hợp 2D) hoặc hình cầu. Do sự cường bức, một vài phần của các vectơ HOA sẽ trở thành không. Ý tưởng được đề xuất trong bài viết này đó là sử dụng đặc tính làm giảm-dư thừa này để mã hóa bậc kết hợp thông tin trường âm thanh. Do sự cường bức nêu trên, chỉ có thể thu được đối với các vùng cụ thể trong không gian, thao tác quay được yêu cầu chung để dịch chuyển một phần thông tin được truyền tới vùng mong muốn trong không gian.

Ví dụ thực hiện sáng chế

Fig.2 minh họa ví dụ của việc làm cong không gian trong trường hợp hai chiều (hình tròn). Hàm làm cong được lựa chọn là

$$f(\phi) = \phi + 2 \operatorname{atan}\left(\frac{a \sin \phi}{1 - a \cos \phi}\right) \quad \text{with } a = -0.4, \quad (27)$$

mà tương tự phản hồi pha của bộ lọc toàn thông rời rạc theo thời gian với một tham số có giá trị thực, tham chiếu ấn phẩm: M. Kappelan, "Eigenschaften von Allpass-Ketten und ihre Anwendung bei der nicht-äquidistanten spektralen Analyse und Synthese", luận án tiến sĩ, đại học Aachen (RWTH), Aachen, Đức, 1998.

Hàm làm cong được thể hiện trên Fig.2a. Hàm làm cong cụ thể này $f(\phi)$ được lựa chọn do nó đảm bảo hàm làm cong tuần hoàn trong khi nó cho phép thay đổi lượng méo dạng không gian với một tham số a .

Hàm trọng số tương ứng $g(\phi)$ được thể hiện trên Fig.2b là kết quả được xác định

đối với hàm làm cong cụ thể này.

Fig.2c thể hiện ma trận làm cong biến đổi đơn bước 7×25 T. Các giá trị tuyệt đối theo logarit của các hệ số riêng biệt của ma trận được thể hiện bởi dải màu xám hoặc các kiểu tạo bóng theo dải màu xám hoặc thanh tạo bóng kèm theo. Ma trận ví dụ này được thiết kế đối với bậc HOA đầu vào là $N_{\epsilon} = 3$ và bậc đầu ra là $N_{\text{out}} = 12$. Bậc đầu ra cao hơn được yêu cầu để thu được hầu hết thông tin mà được trải rộng bởi biến đổi từ các hệ số bậc thấp tới các hệ số bậc cao. Nếu bậc đầu ra được làm giảm hơn nữa, độ chính xác của thao tác làm cong sẽ bị giảm do các hệ số khác không của toàn ma trận làm cong sẽ bị bỏ qua (xem phần Làm thế nào để thiết lập các bậc HOA dưới đây cho sự mô tả chi tiết hơn).

Đặc tính rất hữu ích của ma trận làm cong đặc biệt này đó là phần lớn của ma trận này là không. Điều này cho phép tiết kiệm nhiều công suất tính toán khi thực hiện thao tác này, nhưng nó không phải là quy tắc chung mà các phần nào đó của ma trận biến đổi đơn bước là không.

Fig.2d và Fig.2e minh họa các đặc tính làm cong tại ví dụ của các mẫu chùm được tạo ra bởi một vài sóng phẳng. Cả hai hình vẽ là kết quả từ cùng bảy sóng phẳng đầu vào tại các vị trí ϕ là $0, 2/7\pi, 4/7\pi, 6/7\pi, 8/7\pi, 10/7\pi$ và $12/7\pi$, tất cả với biên độ đồng nhất của chúng, và thể hiện bảy phân phối biên độ góc, tức là vector cuối cùng $\mathbf{s}$ của thao tác giải mã thông thường được xác định trên

$$\mathbf{s} = \Psi^{-1} \mathbf{A} \quad , \quad (28)$$

trong đó, vector HOA $\mathbf{A}$ là là biến thể gốc hoặc được làm cong của tập các sóng phẳng. Các số nằm ngoài đường tròn biểu diễn góc ϕ . Số lượng của các loa phóng thanh ảo (ví dụ 360) là cao hơn đáng kể so với số lượng tham số HOA. Phân phối biên độ hoặc mẫu chùm đối với sóng phẳng đến từ chiều phía trước nằm tại $\phi = 0$.

Fig.2d thể hiện phân phối biên độ của phép biểu diễn HOA gốc. Tất cả bảy phân phối có dạng giống nhau và mô tả cùng độ rộng của búp sóng chính. Các cực đại của các búp sóng chính nằm tại các góc $\phi = (0, 2/7\pi, \dots)$ của bảy đối tượng âm thanh ban đầu, như mong muốn. Các búp sóng chính có các độ rộng tương ứng với bậc bị giới hạn $N_{\text{in}} = 3$ của các vector HOA gốc.

Fig.2e thể hiện các phân phối biên độ đối với cùng đối tượng âm thanh, nhưng sau khi thao tác làm cong được thực hiện. Nói chung, các đối tượng này di chuyển tiến lên chiều phía trước là 0 độ và các mẫu chùm được thay đổi: các búp sóng chính xung quanh chiều phía trước $\phi = 0$ trở nên hẹp hơn và được hội tụ hơn, trong khi các búp sóng chính theo chiều phía sau xung quanh 180 độ trở nên rộng hơn đáng kể. Tại các phía này, với ảnh hưởng tối đa tại 90 và 270 độ, các mẫu chùm trở nên bất đối xứng do gradien lớn của hàm trọng số $g(\phi)$ trên Fig.2b đối với các góc này.

Các cải biến có thể được xem xét này (làm hẹp và định hình lại) của các mẫu chùm có thể được tạo ra bởi bậc cao hơn $N_{\text{out}} = 12$ của vectơ HOA được làm cong. Về mặt lý thuyết, độ phân giải của các búp sóng chính theo chiều phía trước được tăng lên bởi hệ số là 2.33, trong khi độ phân giải theo chiều phía sau được làm giảm bởi hệ số là $1/2.33$. Tín hiệu bậc kết hợp được tạo ra với các bậc cục bộ thay đổi trên không gian. Có thể được giả thiết rằng bậc đầu ra tối thiểu là $2.33 \cdot N_{\text{in}} \approx 7$ được yêu cầu để biểu diễn các hệ số HOA được làm cong với độ chính xác thích hợp. Tron phần Làm thế nào để thiết lập các bậc HOA dưới đây, mô tả về bản chất, các bậc cục bộ được chỉ ra chi tiết hơn.

Các đặc tính

Các bước làm cong được đưa ra nêu trên là chung và rất linh hoạt. Ít nhất các thao tác cơ bản sau đây có thể được thực hiện: quay và/hoặc phản chiếu theo các trục và/or mặt phẳng bất kỳ, làm méo dạng không gia với hàm làm cong liên tục, và lấy trọng số của các chiều cụ thể (tạo chùm không gian).

Trong các phần phụ sau đây, các đặc tính của việc làm cong không gian theo sáng chế được làm nổi bật, và các chi tiết này đưa ra chỉ dẫn về những gì có thể và những gì không thể đạt được. Ngoài ra, một vài quy tắc thiết kế được mô tả.

Về nguyên tắc, các tham số sau đây có thể được điều chỉnh với một vài mức độ tự do để thu được các đặc tính làm cong mong muốn:

- Hàm làm cong $f(\theta, \phi)$;
- hàm trọng số $g(\theta, \phi)$;
- Bậc bên trong N_{warp} ;

- bậc đầu ra N_{out} ;
- Tạo cửa sổ của các hệ số đầu ra với vectơ $\mathbf{w}$.

Tính tuyến tính

Các bước biến đổi cơ bản trong xử lý đa bước là tuyến tính theo định nghĩa. Việc ánh xạ phi tuyến tính của các nguồn âm thanh tới các vị trí mới diễn ra trong phần giữa có ảnh hưởng đến sự xác định của ma trận mã hóa, nhưng bản thân ma trận mã hóa lại là tuyến tính. Do đó, thao tác làm cong không gian được kết hợp và phép nhân ma trận với $\mathbf{T}$ cũng là thao tác tuyến tính, tức là

$$\mathbf{T} \mathbf{A}_1 + \mathbf{T} \mathbf{A}_2 = \mathbf{T}(\mathbf{A}_1 + \mathbf{A}_2) . \quad (29)$$

Đặc tính này là cần thiết do nó cho phép xử lý thông tin trường âm thanh phức tạp mà bao gồm các phần góp đồng thời từ các nguồn âm thanh khác nhau.

Tính bất biến-không gian

Nhằm có thể hiểu được (trừ phi hàm làm cong là tuyến tính hoàn hảo với gradient 1 hoặc -1), biến đổi làm cong không gian không phải là bất biến-không gian. Điều này nghĩa là thao tác xử lý khác nhau đối với các đối tượng âm thanh mà được bố trí ban đầu tại các vị trí khác nhau trên nửa hình cầu. Về mặt toán học, đặc tính này là kết quả của tính phi tuyến tính của hàm làm cong $f(\phi)$, tức là $f(\phi + \alpha) \neq f(\phi) + \alpha$ (30)

đối với ít nhất một vài góc bất kỳ $\alpha \in]0 \dots 2\pi[$.

Khả năng nghịch đảo

Một cách điển hình, ma trận biến đổi $\mathbf{T}$ không thể chỉ được nghịch đảo bằng cách nghịch đảo toán học. Một lý do hiển nhiên đó là ma trận $\mathbf{T}$ thường không phải ma trận vuông. Ngay cả ma trận vuông làm cong không gian sẽ không thể nghịch đảo được do thông tin mà một cách điển hình được trải rộng từ các hệ số bậc thấp tới các hệ số bậc cao hơn sẽ bị mất (so với phần Làm thế nào để thiết lập các bậc HOA và ví dụ trong phần ví dụ thực hiện sáng chế), và việc làm mất thông tin trong thao tác có nghĩa rằng thao tác này không thể được nghịch đảo.

Do đó, cách khác được thực hiện để ít nhất nghịch đảo gần đúng thao tác làm cong không gian. Biến đổi làm cong nghịch đảo $\mathbf{T}_{\text{rev}}$ có thể được thực hiện thông qua

hàm nghịch đảo $f_{\text{rev}}(\cdot)$ của hàm làm cong $f(\cdot)$ mà để

$$f_{\text{rev}}(f(\phi)) = \phi. \quad (31)$$

Phụ thuộc vào sự lựa chọn của các bậc HOA, xử lý này là tương đương biến đổi nghịch đảo.

Làm thế nào để thiết lập các bậc HOA

Khía cạnh quan trọng được tính đến khi thực hiện biến đổi làm cong không gian là các bậc HOA. Trong khi, thông thường, bậc N_{in} của các vectơ đầu vào $\mathbf{A}_{\text{in}}$ được xác định trước bởi các cường bức phía ngoài, cả bậc N_{out} của các vectơ đầu ra $\mathbf{A}_{\text{out}}$ và bậc “bên trong” N_{warp} của thao tác làm cong phi tuyến tính thực tế có thể được gán nhiều hoặc ít hơn tùy ý. Tuy nhiên, cả hai bậc N_{in} và N_{warp} này phải được lựa chọn cẩn thận như được giải thích dưới đây.

Bậc “bên trong” N_{warp} :

Bậc “bên trong” N_{warp} xác định độ chính xác của các bước giải mã, làm con mã hóa thực tế trong xử lý làm cong không gian đa bước nêu trên. Một cách điển hình, bậc N_{warp} sẽ lớn hơn đáng kể so với cả bậc đầu vào N_{in} và bậc đầu ra N_{out} . Lý do cho yêu cầu này đó là các méo dạng và các mẫu giả khác sẽ được sinh ra do thao tác làm cong nói chung là, thao tác phi tuyến tính.

Để giải thích thực tế này, Fig.3 thể hiện ví dụ của toàn ma trận làm cong đối với cùng hàm làm cong như được sử dụng cho ví dụ từ Fig.2. Các Fig.3a, Fig.3c và Fig.3e thể hiện các hàm làm cong $f_1(\phi)$, $f_2(\phi)$ và $f_3(\phi)$, một cách tương ứng. Các Fig.3b, Fig.3d và Fig.3f thể hiện các ma trận làm cong $\mathbf{T}_1(\text{dB})$, $\mathbf{T}_2(\text{dB})$ và $\mathbf{T}_3(\text{dB})$, một cách tương ứng. Vì các lý do minh họa, các ma trận làm cong này không được cắt để xác định ma trận làm cong đối với bậc đầu vào N_{in} hoặc bậc đầu ra N_{out} cụ thể. Thay vì đó, các đường chấm của hộp trung tâm trên các Fig.3b, Fig.3d và Fig.3f thể hiện kích cỡ đích $N_{\text{out}} \times N_{\text{in}}$ của kết quả cuối cùng, tức là ma trận biến đổi được cắt. Theo cách này ảnh hưởng của các méo dạng phi tuyến tính đối với ma trận làm cong là có thể nhận thấy rõ ràng. Trong ví dụ này, các bậc mục tiêu được thiết lập tùy ý là $N_{\text{in}} = 30$ và $N_{\text{out}} = 100$.

Khó khăn cơ bản có thể được thấy trên Fig.3b: rõ ràng rằng do xử lý phi tuyến

tính trong miền không gian các hệ số trong ma trận làm cong được trải rộng xung quanh đường chéo chính - xa hơn từ tâm của ma trận. Tại các khoảng cách rất xa từ tâm, trong ví dụ này tại khoảng $|y| \geq 90$, y là trục dọc, việc trải rộng hệ số tiến tới các biên của toàn ma trận, mà nó dường như trở thành “sự vượt quá”. Điều này tạo ra loại méo dạng đặc biệt mà mở rộng tới phần lớn của ma trận làm cong. Trong các đánh giá thử nghiệm, có nhận thấy rằng các méo dạng này làm suy giảm đáng kể hiệu quả biến đổi, ngay khi các thành phần méo dạng nằm trong vùng mục tiêu của ma trận (được đánh dấu bởi hộp có đường chấm trên hình vẽ).

Đối với ví dụ thứ nhất trên Fig.3b mọi thứ làm việc tốt do bậc “bên trong” của xử lý được lựa chọn là $N_{\text{warp}} = 200$ mà cao hơn đáng kể so với bậc đầu ra $N_{\text{out}} = 100$. Vùng của các méo dạng không mở rộng vào hộp có đường chấm.

Trường hợp khác được thể hiện trên Fig.3d. Bậc bên trong được chỉ rõ là bằng bậc đầu ra, tức là $N_{\text{warp}} = N_{\text{out}} = 100$. Hình vẽ thể hiện rằng việc mở rộng của các méo dạng có tỷ lệ tuyến tính với bậc bên trong. Kết quả đó là các hệ số bậc cao hơn của đầu ra của biến đổi bị ảnh hưởng bởi các phần méo dạng. Ưu điểm của đặc tính tỷ lệ này đó là dường như có thể ngăn ngừa các loại méo dạng phi tuyến tính này bằng cách làm tăng bậc bên trong N_{warp} một cách thích hợp.

Fig.3f thể hiện ví dụ về hàm làm cong tích cực hơn với hệ số lớn hơn $a = 0,7$. Do hàm làm cong tích cực hơn, các méo dạng bây giờ mở rộng vào vùng ma trận mục tiêu ngay cả đối với bậc bên trong là $N_{\text{warp}} = 200$. Đối với trường hợp này, như được mô tả trong đoạn trước đó, bậc bên trong sẽ còn tăng lên đối với ngay cả sự cung cấp vượt mức hơn nữa. Các thử nghiệm đối với hàm làm cong này thể hiện rằng việc làm tăng bậc bên trong tới ví dụ $N = 400$ sẽ loại bỏ các méo dạng phi tuyến tính này.

Tóm lại, thao tác làm cong càng tích cực, bậc bên trong N_{warp} sẽ càng cao. Không tồn tại sự dẫn xuất hình thức của bậc bên trong tối thiểu. Tuy nhiên, nếu nghi ngờ, sự cung cấp vượt mức của bậc “bên trong” là hữu ích do các ảnh hưởng phi tuyến tính được lấy tỷ lệ tuyến tính với kích cỡ của toàn ma trận làm cong. Về nguyên tắc, bậc “bên trong” có thể là cao một cách tùy ý. Cụ thể, nếu ma trận biến đổi đơn bước thu được, bậc bên trong không đóng bất kỳ vai trò đối với độ phức tạp của thao tác làm cong cuối

cùng.

Bậc đầu ra N_{out} :

Để chỉ rõ bậc đầu ra N_{out} của biến đổi làm cong, hai khía cạnh sau đây cần được xem xét:

- Nói chung, bậc đầu ra phải lớn hơn so với bậc đầu vào N_{in} để lưu giữ tất cả thông tin mà được trải rộng tới các hệ số của các bậc khác nhau. Kích cỡ được yêu cầu thực tế cũng phụ thuộc vào các đặc tính của hàm làm cong. Theo kinh nghiệm, hàm làm cong $f(\phi)$ có dải rộng càng nhỏ, bậc đầu ra được yêu cầu càng nhỏ. Nhận thấy rằng trong một vài trường hợp hàm làm cong có thể được lọc thông thấp để giới hạn bậc đầu ra N_{out} được yêu cầu.

Ví dụ có thể được thể hiện trên Fig.3b. Đối với hàm làm cong cụ thể này, bậc đầu ra là $N_{\text{out}} = 100$, như được thể hiện bởi hộp có đường chấm, là đủ để ngăn ngừa tổn hao thông tin. Nếu bậc đầu ra bị làm giảm đáng kể, ví dụ xuống $N_{\text{out}} = 50$, một vài hệ số khác không của ma trận biến đổi sẽ bị sót lại, và sự tổn hao thông tin tương ứng là có thể xảy ra.

- Trong một vài trường hợp, các hệ số HOA đầu ra sẽ được sử dụng cho quy trình hoặc thiết bị mà có thể điều khiển chỉ bậc bị giới hạn. Ví dụ, mục tiêu có thể là việc thiết lập loa phóng thanh với số lượng loa bị giới hạn. Trong các ứng dụng này bậc đầu ra sẽ được chỉ rõ theo các khả năng của hệ thống mục tiêu.

Nếu N_{out} là đủ nhỏ, biến đổi làm cong hiệu quả làm giảm thông tin không gian.

Việc làm giảm của bậc bên trong N_{warp} thành bậc đầu ra N_{out} có thể được thực hiện chỉ bằng cách làm giảm các hệ số bậc cao hơn. Điều này tương ứng với việc áp dụng cửa sổ hình chữ nhật đối với các vector đầu ra HOA. Ngoài ra, các kỹ thuật làm giảm băng thông phức tạp hơn có thể được áp dụng như được mô tả trong bài viết của M.A. Poletti nêu trên hoặc trong bài viết của J. Daniel nêu trên. Nhờ đó, ngay cả nhiều thông tin hơn có thể bị mất so với tạo cửa sổ hình chữ nhật, nhưng các mẫu có tính định hướng ưu việt hơn có thể thu được.

Sáng chế có thể được sử dụng trong các phân khác nhau của chuỗi xử lý âm thanh, ví dụ ghi, sản xuất hậu kỳ, truyền dẫn, phát lại.

YÊU CẦU BẢO HỘ

1. Phương pháp để thay đổi các vị trí tương đối của các đối tượng âm thanh nằm trong phép biểu diễn Ambisonics bậc cao hơn HOA hai chiều hoặc ba chiều của môi trường âm thanh, trong đó vector đầu vào $\mathbf{A}_{in}$ với chiều O_{in} xác định các hệ số của chuỗi Fourier của tín hiệu đầu vào và vector đầu ra $\mathbf{A}_{out}$ với chiều O_{out} xác định các hệ số của chuỗi Fourier của tín hiệu đầu ra được thay đổi tương ứng, phương pháp này bao gồm các bước:

- giải mã (12) vector đầu vào $\mathbf{A}_{in}$ nêu trên của các hệ số HOA đầu vào thành các tín hiệu đầu vào $\mathbf{s}_{in}$ trong miền không gian đối với vị trí ban đầu của nhiều loa phóng thanh sử dụng giả nghịch đảo của ma trận chế độ Ψ_1 bằng cách tính toán:

$$\mathbf{s}_{in} = \Psi^T(\Psi \Psi^T)^{-1} \mathbf{A}_{in};$$

- làm cong và mã hóa (14) trong miền không gian các tín hiệu đầu vào $\mathbf{s}_{in}$ nêu trên thành vector đầu ra $\mathbf{A}_{out}$ nêu trên của các hệ số HOA đầu ra được áp dụng bằng cách tính toán $\mathbf{A}_{out} = \Psi_2 \mathbf{s}_{in}$, trong đó các vector chế độ của ma trận chế độ Ψ_2 được thay đổi so với các vector chế độ của ma trận chế độ Ψ_1 theo hàm làm cong $f(\phi)$ bởi các góc của các vị trí loa phóng thanh gốc được ánh xạ một-một thành các góc mục tiêu của các vị trí loa phóng thanh mục tiêu trong vector đầu ra $\mathbf{A}_{out}$ nêu trên.

2. Phương pháp theo điểm 1, trong đó các tín hiệu đầu vào miền không gian $\mathbf{s}_{in}$ nêu trên được lấy trọng số (13) bởi các hàm tăng ích $g(\phi)$ hoặc $g(\theta, \phi)$ trước bước làm cong và mã hóa (14) nêu trên.

3. Phương pháp theo điểm 2, trong đó trong trường hợp bậc hoặc chiều của các hệ số HOA là thấp hơn bậc hoặc chiều của ma trận chế độ Ψ_2 nêu trên, tín hiệu được làm cong và mã hóa và có thể được lấy trọng số (13) $\Psi_2 \mathbf{s}_{in}$ nêu trên còn được lấy trọng số (15) sử dụng vector cửa sổ $\mathbf{w}$ bao gồm các hệ số không đổi với các bậc cao nhất, để tách (15) một phần của các hệ số được làm cong để đưa ra vector đầu ra $\mathbf{A}_{out}$ nêu trên.

4. Phương pháp theo điểm 3, trong đó bước giải mã (12), lấy trọng số (13) và làm cong và mã hóa (14) nêu trên được thực hiện chung bằng cách sử dụng ma trận biến đổi $\mathbf{T} = \text{diag}(\mathbf{w}) \Psi_2 \text{diag}(\mathbf{g}) \Psi_1^{-1}$ có kích cỡ $O_{warp} \times O_{warp}$, trong đó $\text{diag}(\mathbf{w})$ thể hiện ma trận đường chéo mà có các giá trị của vector cửa sổ $\mathbf{w}$ nêu trên như là các thành phần của

đường chéo chính của nó và $\text{diag}(\mathbf{g})$ thể hiện ma trận đường chéo mà có các giá trị của hàm tăng ích $g(\phi)$ hoặc $g(\theta, \phi)$ nêu trên như là các thành phần của đường chéo chính của nó.

5. Phương pháp theo điểm 4, trong đó để tạo dạng ma trận biến đổi $\mathbf{T}$ nêu trên để có kích cỡ $O_{\text{out}} \times O_{\text{in}}$, các cột và/hoặc dòng tương ứng của ma trận biến đổi $\mathbf{T}$ nêu trên được loại bỏ để thực hiện thao tác làm cong không gian $\mathbf{A}_{\text{out}} = \mathbf{T} \mathbf{A}_{\text{in}}$.

6. Phương pháp theo điểm bất kỳ trong số các điểm từ 1 đến 5, trong đó đối với phép biểu diễn Ambisonics hai chiều, hàm tăng ích là $g(\phi) = \frac{df_{\phi}(\phi)}{d\phi}$,

và đối với phép biểu diễn Ambisonics ba chiều, hàm tăng ích là $g(\theta, \phi) = \frac{df_{\theta}(\theta)}{d\theta}$.

$\frac{\arccos((\cos f_{\theta}(\theta_{\text{in}}))^2 + (\sin f_{\theta}(\theta_{\text{in}}))^2 \cos \phi_{\epsilon})}{\arccos((\cos \theta_{\text{in}})^2 + (\sin \theta_{\text{in}})^2 \cos \phi_{\epsilon})}$ theo chiều ϕ và theo chiều θ , trong đó ϕ là góc phương vị, θ là góc nghiêng và ϕ_{ϵ} là góc phương vị nhỏ.

7. Phương pháp theo điểm bất kỳ trong số các điểm từ 1 đến 6, trong đó trong trường hợp số lượng hoặc chiều O_{warp} của các loa phóng thanh ảo là bằng hoặc lớn hơn số lượng hoặc chiều O_{in} của các hệ số HOA, trước giải mã (12) nêu trên, bậc hoặc chiều của vector đầu vào nêu trên $\mathbf{A}_{\text{in}}$ được mở rộng (11) bằng cách bổ sung (11) các hệ số không đối với các bậc cao hơn.

8. Thiết bị để thay đổi các vị trí tương đối của các đối tượng âm thanh nằm trong phép biểu diễn Ambisonics bậc cao hơn HOA hai chiều hoặc ba chiều của môi trường âm thanh, trong đó vector đầu vào $\mathbf{A}_{\text{in}}$ với chiều O_{in} xác định các hệ số của chuỗi Fourier của tín hiệu đầu vào và vector đầu ra $\mathbf{A}_{\text{out}}$ với chiều O_{out} xác định các hệ số của chuỗi Fourier của tín hiệu đầu ra được thay đổi tương ứng, thiết bị này bao gồm:

- phương tiện (12) được áp dụng để giải mã vector đầu vào $\mathbf{A}_{\text{in}}$ nêu trên của các hệ số HOA đầu vào thành các tín hiệu đầu vào $\mathbf{s}_{\text{in}}$ trong miền không gian đối với các vị trí ban đầu của nhiều loa phóng thanh sử dụng nghịch đảo giả của ma trận chế độ Ψ_1 bằng cách tính toán $\mathbf{s}_{\text{in}} = \Psi^T (\Psi \Psi^T)^{-1} \mathbf{A}_{\text{in}}$;

- phương tiện (14) được áp dụng để làm cong và mã hóa trong miền không gian các tín hiệu đầu vào $\mathbf{s}_{\text{in}}$ nêu trên thành vector đầu ra $\mathbf{A}_{\text{out}}$ nêu trên của các hệ số HOA đầu

ra được áp dụng bằng cách tính toán $\mathbf{A}_{\text{out}} = \Psi_2 \mathbf{s}_{\text{in}}$, trong đó các vector chế độ của ma trận chế độ Ψ_2 được thay đổi theo hàm làm cong $f(\phi)$ bởi đó các góc của các vị trí loa phóng thanh gốc được ánh xạ một-một thành các góc mục tiêu của các vị trí loa phóng thanh mục tiêu trong vector đầu ra $\mathbf{A}_{\text{out}}$ nêu trên.

9. Thiết bị theo điểm 8, thiết bị này bao gồm phương tiện (11) được áp dụng để mở rộng, trước giải mã (12) nêu trên, bậc hoặc chiều của vector đầu vào $\mathbf{A}_{\text{in}}$ nêu trên bằng cách bổ sung các hệ số không đổi với các bậc cao hơn, trong trường hợp số lượng hoặc chiều O_{warp} của các loa phóng thanh ảo là bằng hoặc lớn hơn số lượng hoặc chiều O_{in} của các hệ số HOA.

10. Thiết bị theo điểm 9, thiết bị này bao gồm phương tiện (15) được áp dụng để còn lấy trọng số sử dụng vector cửa sổ $\mathbf{w}$ bao gồm các hệ số không đổi với các bậc cao nhất, tín hiệu được làm cong và mã hóa $\Psi_2 \mathbf{s}_{\text{in}}$ nêu trên, và để tách một phần của các hệ số được làm cong để đưa ra vector đầu ra $\mathbf{A}_{\text{out}}$ nêu trên.

11. Thiết bị theo điểm 10, thiết bị này bao gồm phương tiện (12,13,14,15) được áp dụng để thực hiện chung bước giải mã, lấy trọng số và làm cong/mã hóa nêu trên bằng cách sử dụng ma trận biến đổi $\mathbf{T} = \text{diag}(\mathbf{w})\Psi_2 \text{diag}(\mathbf{g})\Psi_1^{-1}$ có kích cỡ $O_{\text{warp}} \times O_{\text{warp}}$, trong đó $\text{diag}(\mathbf{w})$ thể hiện ma trận đường chéo mà có các giá trị của vector cửa sổ $\mathbf{w}$ nêu trên như là các thành phần của đường chéo chính của nó và $\text{diag}(\mathbf{g})$ thể hiện ma trận đường chéo mà có các giá trị của hàm tăng ích $g(\phi)$ hoặc $g(\theta, \phi)$ nêu trên như là các thành phần của đường chéo chính của nó.

12. Thiết bị theo điểm 11, trong đó để tạo dạng ma trận biến đổi $\mathbf{T}$ nêu trên để có kích cỡ $O_{\text{out}} \times O_{\text{in}}$, trong các phương tiện (12,13,14,15) nêu trên được áp dụng để thực hiện chung bước giải mã, lấy trọng số và làm cong/mã hóa nêu trên, các cột và/hoặc dòng tương ứng của ma trận biến đổi $\mathbf{T}$ nêu trên được loại bỏ để thực hiện thao tác làm cong không gian $\mathbf{A}_{\text{out}} = \mathbf{T} \mathbf{A}_{\text{in}}$.

13. Thiết bị theo điểm 12, thiết bị này bao gồm phương tiện (13) được áp dụng để lấy trọng số các tín hiệu đầu vào miền không gian $\mathbf{s}_{\text{in}}$ nêu trên bởi hàm tăng ích $g(\phi)$ hoặc $g(\theta, \phi)$ trước bước làm cong và mã hóa (14) nêu trên.

14. Thiết bị theo điểm 13, trong đó đối với phép biểu diễn Ambisonics hai chiều, hàm

tăng ích nêu trên là $g(\phi) = \frac{df_\phi(\phi)}{d\phi}$,

và đối với phép biểu diễn Ambisonics ba chiều, hàm tăng ích nêu trên là $g(\theta, \phi) = \frac{df_\theta(\theta)}{d\theta} \cdot \frac{\arccos((\cos f_\theta(\theta_{in}))^2 + (\sin f_\theta(\theta_{in}))^2 \cos \phi_\varepsilon)}{\arccos((\cos \theta_{in})^2 + (\sin \theta_{in})^2 \cos \phi_\varepsilon)}$ theo chiều ϕ và theo chiều θ , trong đó ϕ là góc phương vị, θ là góc nghiêng và ϕ_ε là góc phương vị nhỏ.

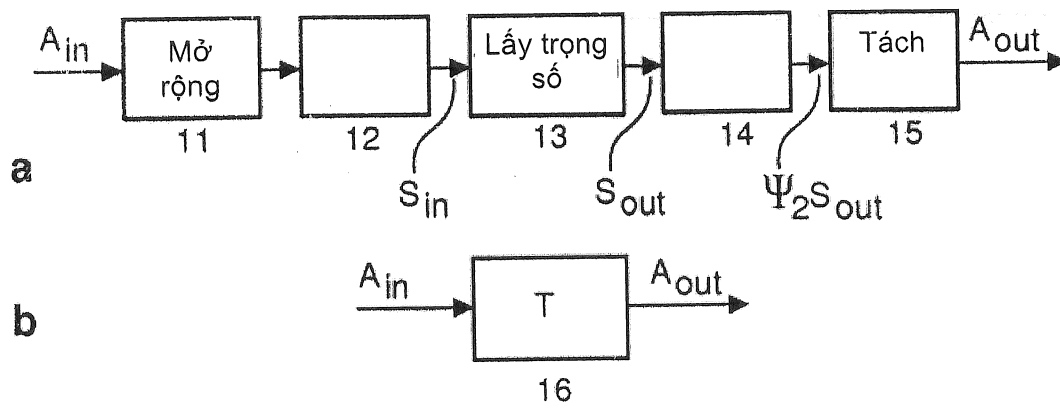


Fig. 1

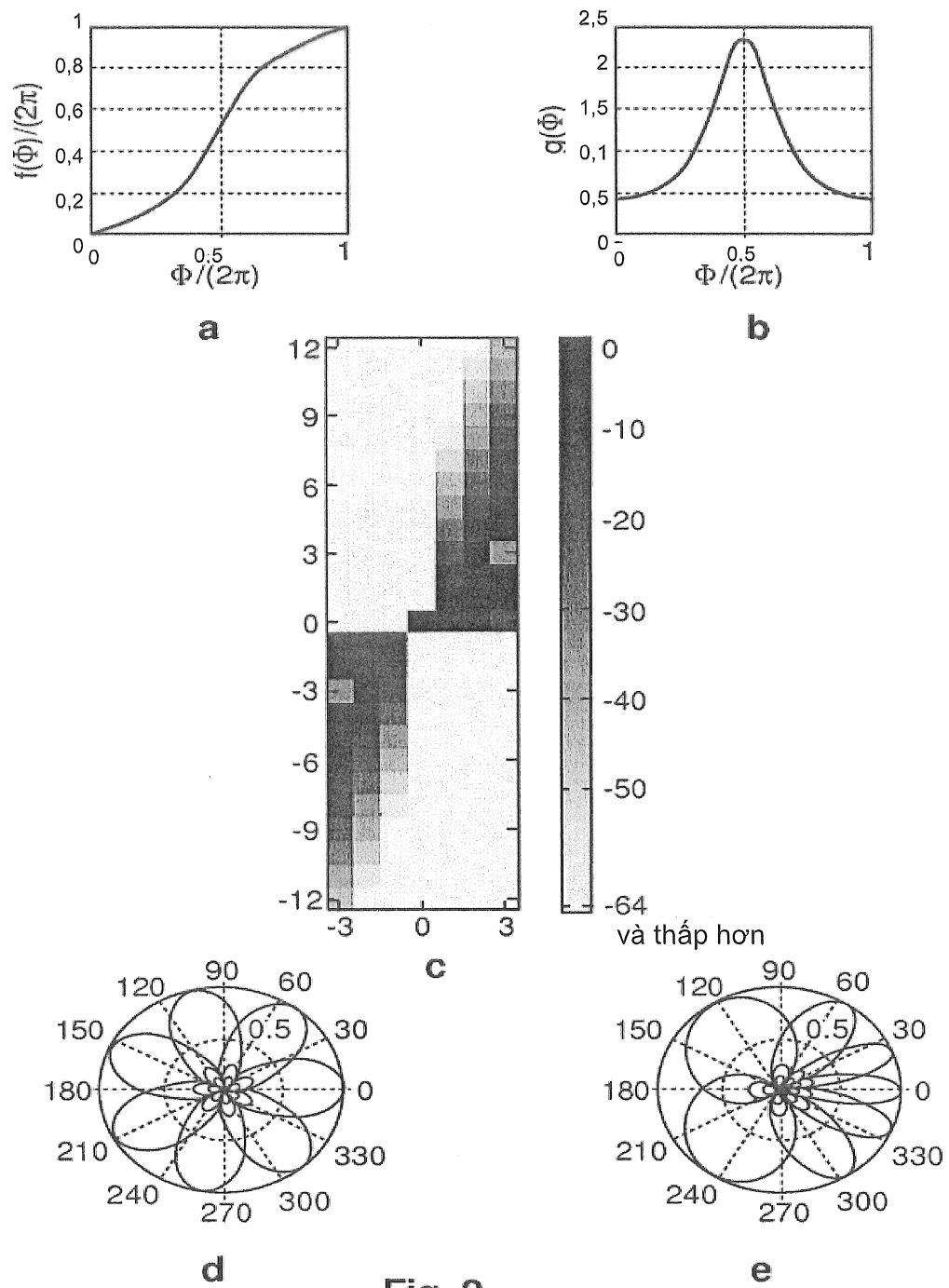


Fig. 2

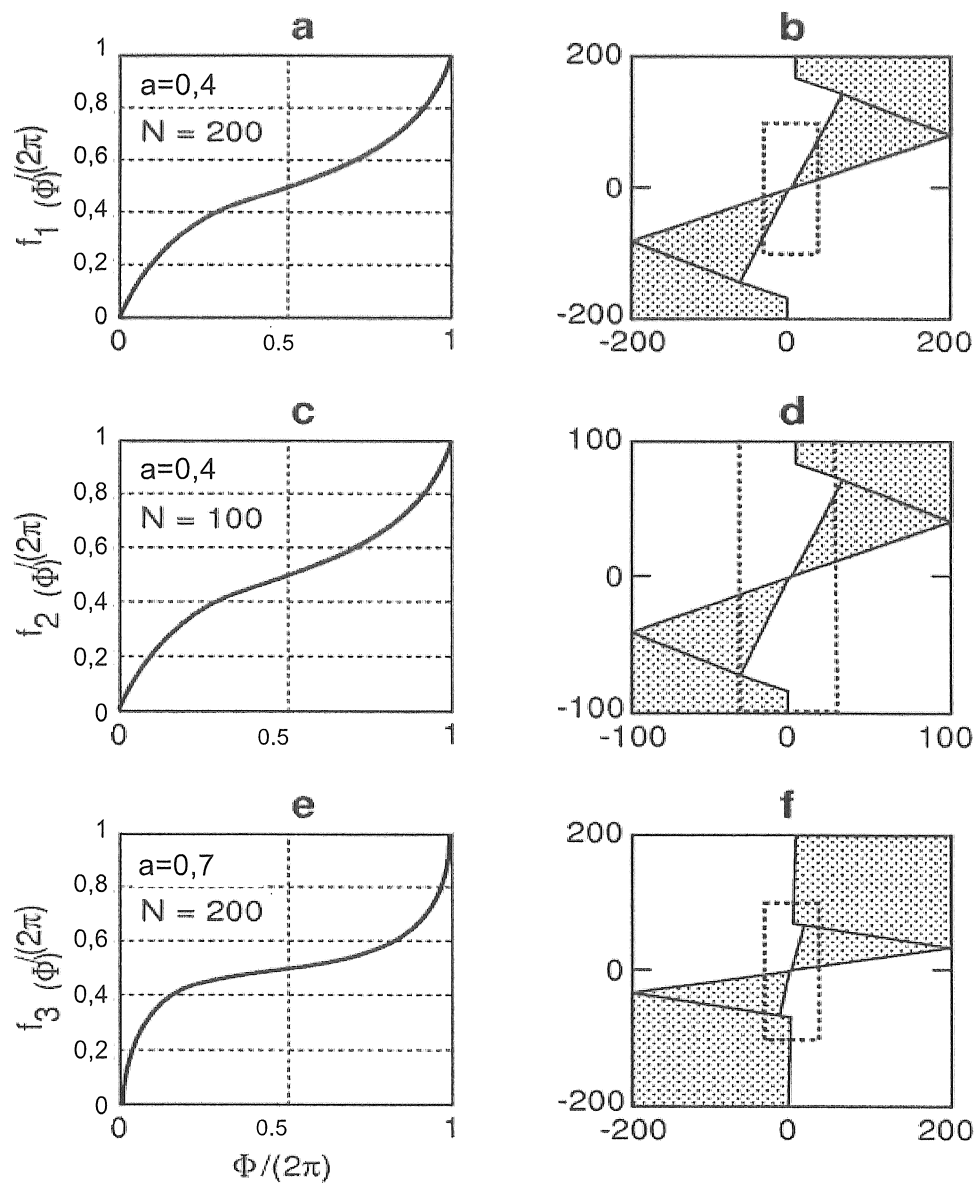


Fig. 3