



(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)



1-0029930

(51)⁷ H04H 20/89

(13) B

(21) 1-2017-03335

(22) 06/05/2013

(62) 1-2014-03547

(86) PCT/EP2013/059363 06/05/2013

(87) WO2013/171083 21/11/2013

(30) 12305537.8 14/05/2012 EP

(45) 25/11/2021 404

(43) 25/10/2017 355A

(73) DOLBY INTERNATIONAL AB (SE)

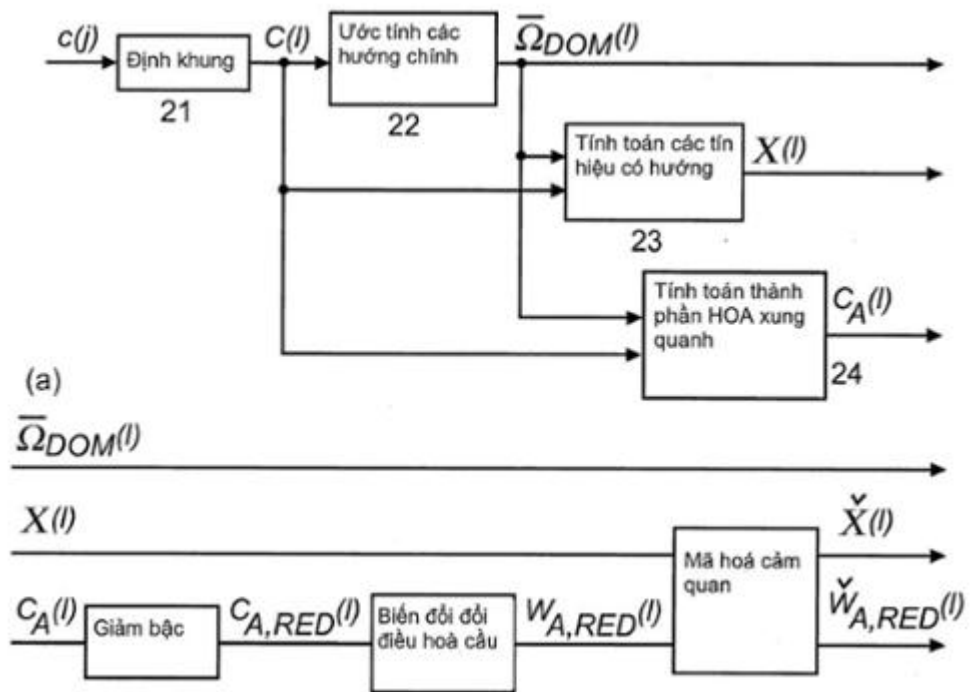
Apollo Building, 3E, Herikerbergweg 1-35, 1101 CN Amsterdam Zuidoost,
Netherlands

(72) KRÜGER, Alexander (DE); KORDON, Sven (DE); BOEHM, Johannes (DE);
BATKE, Johann-Markus (DE).

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

(54) PHƯƠNG PHÁP VÀ THIẾT BỊ GIẢI NÉN TÍN HIỆU AMBISONIC BẬC CAO
ĐÃ ĐƯỢC NÉN VÀ VẬT GHI BẮT BIẾN ĐỌC ĐƯỢC BẰNG MÁY TÍNH

(57) Sáng chế đề cập đến phương pháp và thiết bị giải nén dạng biểu diễn tín hiệu ambisonic bậc cao (HOA - Higher Order Ambisonics) đã được nén. Tín hiệu HOA biểu diễn trường âm thanh hoàn chỉnh ở vùng lân cận của điểm sắc nét, không phụ thuộc vào cài đặt loa phóng thanh. Độ phân giải không gian cao đòi hỏi số hệ số HOA lớn. Theo sáng chế, các hướng âm thanh chính được ước tính và dạng biểu diễn tín hiệu HOA được phân tích thành các tín hiệu có hướng chính trong miền thời gian và thông tin hướng liên quan, và thành phần xung quanh trong miền HOA, tiếp theo là nén thành phần xung quanh bằng cách giảm bậc của nó. Thành phần xung quanh có bậc giảm được biến đổi sang miền không gian, và được mã hóa cảm quan cùng với các tín hiệu có hướng. Ở phía bộ thu, các tín hiệu có hướng được mã hóa và các thành phần xung quanh được mã hóa giảm bậc được giải nén cảm quan, các tín hiệu xung quanh được giải nén cảm quan được biến đổi thành dạng biểu diễn miền tín hiệu HOA có bậc đã giảm, sau đó mở rộng bậc. Dạng biểu diễn tín hiệu HOA tổng được biên soạn lại từ các tín hiệu có hướng, thông tin hướng tương ứng, và thành phần HOA xung quanh bậc ban đầu.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến phương pháp và thiết bị nén và giải nén dạng biểu diễn tín hiệu ambisonic bậc cao, trong đó các thành phần có hướng và thành phần xung quanh được xử lý theo các cách khác nhau.

Tình trạng kỹ thuật của sáng chế

Tín hiệu ambisonic bậc cao (Higher Order Ambisonic - HOA) có ưu điểm bất giữ trường âm thanh hoàn chỉnh trong vùng lân cận của một vị trí cụ thể trong không gian ba chiều, trong đó vị trí này được gọi là 'điểm sắc nét'. Dạng biểu diễn tín hiệu HOA như vậy không phụ thuộc vào cài đặt loa phóng thanh cụ thể, trái ngược với các kỹ thuật dựa trên kênh như âm thanh lập thể hoặc âm thanh vòm. Nhưng tính linh hoạt này có thể gây bất lợi cho quy trình giải mã cần thiết để phát lại dạng biểu diễn tín hiệu HOA trên cài đặt loa phóng thanh cụ thể.

Tín hiệu HOA dựa trên phần mô tả về các biên độ phức của áp suất không khí đối với các số sóng góc riêng k cho các vị trí x ở vùng lân cận của vị trí thính giả mong muốn, mà không mất đi tính tổng quát có thể được giả định là tín hiệu gốc của hệ tọa độ hình cầu, bằng cách sử dụng kỹ thuật mở rộng hàm điều hòa cầu (Spherical Harmonics - SH) cắt cụt. Độ phân giải không gian của dạng biểu diễn này cải thiện với bậc cực đại tăng N của phần mở rộng. Đáng tiếc là, số lượng các hệ số mở rộng O tăng lên bậc hai với bậc N , tức là $O = (N + 1)^2$. Ví dụ, dạng biểu diễn tín hiệu HOA điển hình sử dụng bậc $N = 4$ cần $O = 25$ hệ số tín hiệu HOA. Nếu tốc độ lấy mẫu mong muốn f_s và số bit N_b được dùng trên mỗi mẫu, thì tốc độ bit tổng để truyền dạng biểu diễn tín hiệu HOA được xác định bằng $O \cdot f_s \cdot N_b$, và cuộc truyền dạng biểu diễn tín hiệu HOA bậc $N = 4$ với tốc độ lấy mẫu là $f_s = 48 \text{ kHz}$ sử dụng $N_b = 16$ bit cho mỗi mẫu dẫn đến tốc độ bit là $19,2 \text{ MBit/giây}$. Do đó, việc nén dạng biểu diễn tín hiệu HOA là rất cần thiết.

Tổng quan về các phương pháp nén âm thanh trong không gian hiện có có thể tìm thấy trong đơn sáng chế châu Âu số EP 10306472.1 hoặc trong tài liệu: I. Elfitri,

B. Günel, A.M. Kondo, "Multichannel Audio Coding Based on Analysis by Synthesis", Proceedings of the IEEE, tập 99, số 4, từ trang 657 đến 670, tháng Tư năm 2011.

Các kỹ thuật sau đây là thích hợp hơn đối với sáng chế.

Tín hiệu định dạng B tương đương với dạng biểu diễn tín hiệu ambisonic bậc một có thể được nén bằng cách sử dụng kỹ thuật mã hóa âm thanh có hướng (Directional Audio Coding - DirAC) như được mô tả trong tài liệu: V. Pulkki, "Spatial Sound Reproduction with Directional Audio Coding", Journal of Audio Eng. Society, tập 55(6), trang từ 503 đến 516, 2007. Theo một phiên bản được đề xuất cho các ứng dụng hội nghị từ xa, tín hiệu định dạng B được mã hóa thành tín hiệu vô hướng đơn cũng như thông tin phụ ở dạng hướng đơn và tham số khuếch tán trên mỗi băng tần. Tuy nhiên, kết quả của việc giảm mạnh tốc độ dữ liệu ảnh hưởng đến chất lượng tín hiệu phụ thu được ở dạng bản sao. Ngoài ra, DirAC bị giới hạn ở quy trình nén dạng biểu diễn tín hiệu ambisonic bậc một mà có độ phân giải không gian rất thấp.

Các phương pháp đã biết để nén dạng biểu diễn tín hiệu HOA với $N > 1$ khá hiếm. Một trong số các phương pháp này thực hiện mã hoá trực tiếp các dãy hệ số HOA riêng bằng cách dùng bộ mã hóa-giải mã để mã hóa âm thanh tiên tiến (Advanced Audio Coding - AAC) về mặt cảm quan, xem tài liệu: E. Hellerud, I. Burnett, A. Solvang, U. Peter Svensson, "Encoding Higher Order Ambisonics with AAC", 124th AES Convention, Amsterdam, 2008. Tuy nhiên, vấn đề cố hữu của phương pháp này là mã hóa cảm quan các tín hiệu chưa từng được nghe. Các tín hiệu phát lại được tái tạo thường thu được bằng tổng trọng số của dãy hệ số HOA. Đó là lý do tại sao có xác suất cao để lộ nhiều mã hóa cảm quan khi dạng biểu diễn tín hiệu HOA đã giải nén được kết xuất trên cài đặt loa phóng thanh cụ thể. Nói sâu hơn về mặt kỹ thuật, vấn đề chính của việc để lộ nhiều mã hóa cảm quan là sự tương quan chéo mức cao giữa các dãy hệ số HOA riêng lẻ. Do tín hiệu nhiều được mã hóa trong các dãy hệ số HOA riêng lẻ thường không liên quan với nhau, nên có thể xảy ra sự xếp chồng về cấu trúc của nhiều mã hóa cảm quan trong khi các dãy hệ số HOA không nhiều được hủy đồng thời nếu bị chồng. Một vấn đề khác là sự tương quan chéo đã nêu làm cho hiệu suất của các bộ mã hóa cảm quan bị giảm sút.

Để giảm thiểu những tác động này, đơn sáng chế châu Âu số EP 10306472.1 được đề xuất để biến đổi dạng biểu diễn tín hiệu HOA sang dạng biểu diễn tương đương trong miền không gian trước khi mã hóa cảm quan. Các tín hiệu miền không gian tương ứng với các tín hiệu có hướng thông thường, và sẽ tương ứng với các tín hiệu loa phóng thanh nếu các loa phóng thanh này được bố trí theo đúng các hướng giống như hướng được giả định cho dạng biến đổi miền không gian.

Việc biến đổi sang miền không gian làm giảm mức tương quan chéo giữa các tín hiệu trong miền không gian riêng lẻ. Tuy nhiên, các mức tương quan chéo này không hoàn toàn bị loại bỏ. Ví dụ về mức tương quan chéo tương đối cao là tín hiệu có hướng, mà hướng của tín hiệu này nằm ở giữa các hướng liền kề được bao bởi các tín hiệu miền không gian.

Nhược điểm khác trong đơn sáng chế châu Âu số EP 10306472.1 và tài liệu của Hellerud và các cộng sự nêu trên đó là số tín hiệu được mã hóa cảm quan là $(N + 1)^2$, trong đó N là bậc của dạng biểu diễn tín hiệu HOA. Do đó, tốc độ dữ liệu của dạng biểu diễn tín hiệu HOA đã nén tăng khả tích với bậc ambisonic.

Quy trình nén theo sáng chế thực hiện phân tích dạng biểu diễn trường âm thanh tín hiệu HOA thành thành phần có hướng và thành phần xung quanh. Cụ thể là, đối với việc tính toán thành phần trường âm thanh có hướng, một quy trình mới được mô tả dưới đây được dùng để ước tính một vài hướng âm thanh chính.

Đối với các phương pháp hiện có dùng để ước tính hướng dựa trên tín hiệu ambisonic, tài liệu của Pulkki nêu trên mô tả một phương pháp liên quan đến mã hóa DirAC để ước tính hướng, dựa trên dạng biểu diễn trường âm thanh định dạng B. Hướng thu được từ vector cường độ trung bình để trở đến hướng của dòng năng lượng trường âm thanh. Phương pháp khác dựa trên định dạng B được đề xuất trong tài liệu: D. Levin, S. Gannot, E.A.P. Habets, "Direction-of-Arrival Estimation using Acoustic Vector Sensors in the Presence of Noise", IEEE Proc. of the ICASSP, từ trang 105 đến 108, 2011. Quy trình ước tính hướng được thực hiện lặp đi lặp lại bằng cách tìm kiếm hướng để tạo ra công suất cực đại cho tín hiệu đầu ra chứa chùm âm thanh được lái theo hướng đó.

Tuy nhiên, cả hai phương pháp này bị hạn chế bởi định dạng B trong quá trình ước tính hướng, quá trình này có độ phân giải không gian tương đối thấp. Một nhược

điểm khác là quá trình ước tính chỉ hạn chế ở một hướng chính.

Các dạng biểu diễn tín hiệu HOA có độ phân giải không gian được cải thiện và do đó cho phép cải thiện việc ước tính một số hướng chính. Các phương pháp hiện có thực hiện ước tính một số hướng dựa trên dạng biểu diễn trường âm thanh tín hiệu HOA là khá hiếm. Phương pháp dựa trên nhận biết nén được đề xuất trong tài liệu: N. Epain, C. Jin, A. van Schaik, "The Application of Compressive Sampling to the Analysis and Synthesis of Spatial Sound Fields", 127th Convention of the Audio Eng. Soc., New York, 2009, và trong tài liệu: A. Wabnitz, N. Epain, A. van Schaik, C Jin, "Time Domain Reconstruction of Spatial Sound Fields Using Compressed Sensing", IEEE Proc. of the ICASSP, từ trang 465 đến 468, 2011. Ý tưởng chính là giả định trường âm thanh nằm rải rác trong không gian, tức là chỉ bao gồm lượng nhỏ tín hiệu có hướng. Sau khi phân bố một lượng lớn các hướng thử nghiệm trên mặt cầu, thuật toán tối ưu hóa được sử dụng để tìm kiếm càng ít hướng thử nghiệm càng tốt cùng với các tín hiệu có hướng tương ứng, sao cho chúng được mô tả bằng dạng biểu diễn tín hiệu HOA cho trước. Phương pháp này tạo ra độ phân giải không gian cải thiện so với độ phân giải thực tế tạo ra bằng dạng biểu diễn tín hiệu HOA cho trước, vì nó tránh được sự phân tán không gian do bậc giới hạn của dạng biểu diễn tín hiệu HOA cho trước. Tuy nhiên, việc thực hiện thuật toán này phụ thuộc rất nhiều vào việc liệu sự giả định phân bố rải rác có thỏa mãn không. Cụ thể là, phương pháp này không được thực hiện nếu trường âm thanh chứa thành phần xung quanh bổ sung phụ bất kỳ, hoặc nếu dạng biểu diễn tín hiệu HOA bị ảnh hưởng bởi nhiễu sẽ xảy ra khi nó được tính từ các bản ghi đa kênh.

Một phương pháp trực giác khác nữa được dùng để biến đổi dạng biểu diễn tín hiệu HOA cho trước sang miền không gian như được mô tả trong tài liệu: B. Rafaely, "Plane-wave decomposition of the sound field on a sphere by spherical convolution", J. Acoust. Soc. Am., tập 4, số 116, từ trang 2149 đến 2157, tháng 10 năm 2004, và sau đó tìm kiếm giá trị lớn nhất theo công suất có hướng. Nhược điểm của phương pháp này là có các thành phần xung quanh dẫn đến làm mờ sự phân bố công suất có hướng và dẫn đến sự dịch chuyển các giá trị lớn nhất trong số các công suất có hướng so với việc không có thành phần xung quanh bất kỳ.

Bản chất kỹ thuật của sáng chế

Vấn đề cần giải quyết theo sáng chế đó là đề xuất kỹ thuật nén các tín hiệu HOA mà nhờ đó vẫn giữ được độ phân giải không gian cao của dạng biểu diễn tín hiệu HOA. Phương pháp giải quyết vấn đề này được mô tả trong điểm 1 phần yêu cầu bảo hộ. Thiết bị sử dụng các phương pháp này được mô tả trong điểm 4 phần yêu cầu bảo hộ.

Sáng chế đề cập đến kỹ thuật nén dạng biểu diễn tín hiệu ambisonic bậc cao (HOA) của trường âm thanh. Theo sáng chế, thuật ngữ 'HOA' biểu thị dạng biểu diễn tín hiệu ambisonic bậc cao cũng như tín hiệu âm thanh được biểu diễn hoặc được mã hóa tương ứng. Các hướng âm thanh chính được ước tính và dạng biểu diễn tín hiệu HOA được phân tích thành một số tín hiệu có hướng chính trong miền thời gian và thông tin hướng liên quan, và thành phần xung quanh trong miền HOA, tiếp theo là nén thành phần xung quanh bằng cách giảm bậc của nó. Sau quá trình phân tích đó, thành phần HOA xung quanh có bậc đã giảm được biến đổi sang miền không gian, và được mã hóa cảm quan cùng với các tín hiệu có hướng.

Ở phía bộ thu hoặc bộ giải mã, các tín hiệu có hướng được mã hóa và thành phần xung quanh được mã hóa giảm bậc được giải nén cảm quan. Các tín hiệu xung quanh đã giải nén cảm quan được biến đổi sang dạng biểu diễn miền tín hiệu HOA có bậc đã giảm, tiếp theo là mở rộng bậc. Dạng biểu diễn tín hiệu HOA tổng được biên soạn lại từ các tín hiệu có hướng và thông tin hướng tương ứng và từ thành phần HOA xung quanh bậc ban đầu.

Có lợi là, thành phần trường âm xung quanh có thể được biểu diễn với độ chính xác đầy đủ bởi dạng biểu diễn tín hiệu HOA có bậc ban đầu thấp hơn, và việc trích các tín hiệu có hướng chính đảm bảo rằng, tiếp theo quy trình nén và giải nén, vẫn đạt được độ phân giải không gian cao.

Theo nguyên lý, sáng chế đề xuất phương pháp nén dạng biểu diễn tín hiệu HOA, phương pháp này bao gồm các bước:

- ước tính các hướng chính, trong đó quy trình ước tính hướng chính này phụ thuộc vào sự phân bố công suất có hướng của các thành phần HOA chính có năng lượng cao;
- phân tích hoặc giải mã dạng biểu diễn tín hiệu HOA thành một số tín hiệu có

hướng chính trong miền thời gian và thông tin hướng liên quan, và thành phần xung quanh dư trong miền tín hiệu HOA, trong đó thành phần xung quanh dư biểu diễn hiệu số giữa dạng biểu diễn tín hiệu HOA và dạng biểu diễn các tín hiệu có hướng chính;

- nén thành phần xung quanh dư này bằng cách giảm bậc của nó khi so sánh với bậc ban đầu của nó;

- biến đổi thành phần HOA xung quanh dư có bậc đã giảm sang miền không gian;

- mã hóa cảm quan các tín hiệu có hướng chính và thành phần HOA xung quanh dư đã biến đổi.

Theo nguyên lý, sáng chế đề xuất phương pháp giải nén dạng biểu diễn tín hiệu HOA mà được nén theo các bước:

- ước tính các hướng chính, trong đó quy trình ước tính hướng chính này phụ thuộc vào sự phân bố công suất có hướng của các thành phần HOA chính có năng lượng cao;

- phân tích hoặc giải mã dạng biểu diễn tín hiệu HOA thành một số tín hiệu có hướng chính trong miền thời gian và thông tin hướng liên quan, và thành phần xung quanh dư trong miền tín hiệu HOA, trong đó thành phần xung quanh dư này biểu diễn hiệu số giữa dạng biểu diễn tín hiệu HOA và dạng biểu diễn của các tín hiệu có hướng chính này;

- nén thành phần xung quanh dư bằng cách giảm bậc của nó khi so sánh với bậc ban đầu của nó;

- biến đổi thành phần HOA xung quanh dư có bậc đã giảm sang miền không gian;

- mã hóa cảm quan các tín hiệu có hướng chính và thành phần HOA xung quanh dư đã biến đổi,

trong đó, phương pháp này bao gồm các bước:

- giải mã cảm quan các tín hiệu có hướng chính được mã hóa cảm quan và thành phần HOA xung quanh dư được biến đổi đã mã hóa cảm quan;

- biến đổi ngược thành phần HOA xung quanh dư được biến đổi đã giải mã cảm quan để thu được dạng biểu diễn miền tín hiệu HOA;

- thực hiện mở rộng bậc của thành phần HOA xung quanh dư được biến đổi ngược để thiết lập thành phần HOA xung quanh bậc ban đầu;

- biên soạn các tín hiệu có hướng chính được giải mã cảm quan, thông tin hướng và thành phần HOA xung quanh được mở rộng bậc ban đầu để thu được dạng biểu diễn tín hiệu HOA.

Theo nguyên lý, sáng chế đề xuất thiết bị nén dạng biểu diễn tín hiệu HOA, thiết bị này bao gồm:

- phương tiện được làm thích ứng để ước tính các hướng chính, trong đó quy trình ước tính định hướng chính này phụ thuộc vào sự phân bố công suất có hướng của các thành phần HOA chính có năng lượng cao;

- phương tiện được làm thích ứng để phân tích hoặc giải mã dạng biểu diễn tín hiệu HOA thành một số tín hiệu có hướng chính trong miền thời gian và thông tin hướng liên quan, và thành phần xung quanh dư trong miền tín hiệu HOA, trong đó thành phần xung quanh dư biểu diễn hiệu số giữa dạng biểu diễn tín hiệu HOA và dạng biểu diễn các tín hiệu có hướng chính này;

- phương tiện được làm thích ứng để nén thành phần xung quanh dư bằng cách giảm bậc của nó khi so sánh với bậc ban đầu của nó;

- phương tiện được làm thích ứng để biến đổi thành phần HOA xung quanh dư có bậc đã giảm sang miền không gian;

- phương tiện được làm thích ứng để mã hóa cảm quan các tín hiệu có hướng chính và thành phần HOA xung quanh dư được biến đổi.

Theo nguyên lý, sáng chế đề xuất thiết bị giải nén dạng biểu diễn tín hiệu HOA mà đã được nén theo các bước:

- ước tính các hướng chính, trong đó quy trình ước tính hướng chính này phụ thuộc vào sự phân bố công suất có hướng của các thành phần HOA chính có năng lượng cao;

- phân tích hoặc giải mã dạng biểu diễn tín hiệu HOA thành một số tín hiệu có hướng chính trong miền thời gian và thông tin hướng liên quan, và thành phần xung quanh dư trong miền tín hiệu HOA, trong đó thành phần xung quanh dư này biểu diễn hiệu số giữa dạng biểu diễn tín hiệu HOA và dạng biểu diễn các tín hiệu có hướng chính này;

- nén thành phần xung quanh dư bằng cách giảm bậc của nó khi so sánh với bậc ban đầu của nó;
 - biến đổi thành phần HOA xung quanh dư có bậc đã giảm sang miền không gian;
 - mã hóa cảm quan các tín hiệu có hướng chính và thành phần HOA xung quanh dư được biến đổi,
- trong đó, thiết bị này bao gồm:
- phương tiện được làm thích ứng để giải mã cảm quan các tín hiệu có hướng chính được mã hóa cảm quan và thành phần HOA xung quanh dư được biến đổi đã mã hóa cảm quan;
 - phương tiện được làm thích ứng để biến đổi ngược thành phần HOA xung quanh dư được biến đổi đã mã hóa cảm quan để thu được dạng biểu diễn miền HOA;
 - phương tiện được làm thích ứng để thực hiện mở rộng bậc của thành phần HOA xung quanh dư được biến đổi ngược để thiết lập thành phần HOA xung quanh bậc ban đầu;
 - phương tiện được làm thích ứng để biên soạn các tín hiệu có hướng chính được mã hóa cảm quan, thông tin hướng và thành phần HOA xung quanh được mở rộng bậc ban đầu để thu được dạng biểu diễn tín hiệu HOA.

Các phương án bổ sung có lợi theo sáng chế được mô tả trong phần yêu cầu bảo hộ phụ thuộc tương ứng.

Mô tả vắn tắt các hình vẽ

Các phương án làm ví dụ của sáng chế được mô tả có dựa vào các hình vẽ kèm theo, trong đó:

Fig.1 là sơ đồ thể hiện hàm phân tán được chuẩn hóa $v_N(\theta)$ cho các bậc ambisonic khác nhau N và cho các góc $\theta \in [0, \pi]$;

Fig.2 là sơ đồ khối của quy trình nén theo sáng chế;

Fig.3 là sơ đồ khối của quy trình giải nén theo sáng chế.

Mô tả chi tiết sáng chế

Các tín hiệu ambisonic mô tả trường âm thanh trong các vùng không có nguồn

bằng cách sử dụng phương pháp mở rộng hàm điều hòa cầu (SH). Tính khả thi của sáng chế này có thể là do tính chất vật lý mà dạng biểu diễn thời gian và không gian của áp suất âm thanh được xác định về cơ bản bằng phương trình sóng.

Phương trình sóng và phương pháp mở rộng hàm điều hòa cầu

Để mô tả chi tiết hơn về tín hiệu ambisonic, trong phần sau đây hệ tọa độ cầu được giả định, trong đó một điểm trong không gian $x = (r, \theta, \phi)^T$ được biểu diễn bằng bán kính $r > 0$ (tức là khoảng cách đến điểm gốc tọa độ), góc nghiêng $\theta \in [0, \pi]$ được đo từ trục cực z, và góc phương vị $\phi \in [0, 2\pi[$ được đo trong mặt phẳng $x=y$ từ trục x. Theo hệ tọa độ cầu này, phương trình sóng cho áp suất âm thanh $p(t, x)$ trong vùng kết nối không có nguồn, trong đó t là thời gian được xác định trong tài liệu: Earl G. Williams, "Fourier Acoustics", tập 93, Applied Mathematical Sciences, Academic Press, 1999:

$$\frac{1}{r^2} \left[\frac{\partial}{\partial r} \left(r^2 \frac{\partial p(t, x)}{\partial r} \right) + \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial p(t, x)}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2 p(t, x)}{\partial \phi^2} \right] - \frac{1}{c_s^2} \frac{\partial^2 p(t, x)}{\partial t^2} = 0 \quad (1)$$

với c_s là vận tốc âm thanh. Do đó, phép biến đổi Fourier của áp suất âm thanh liên quan đến thời gian

$$P(\omega, x) := F_t\{p(t, x)\} \quad (2)$$

$$:= \int_{-\infty}^{\infty} p(t, x) e^{-i\omega t} dt, \quad (3)$$

trong đó, i biểu thị đơn vị ảo, đơn vị này có thể được mở rộng thành dãy SH theo sách của Williams:

$$P(kc_s, (r, \theta, \phi)^T) = \sum_{n=0}^{\infty} \sum_{m=-n}^n p_n^m(kr) Y_n^m(\theta, \phi). \quad (4)$$

Cần lưu ý rằng quy trình mở rộng này có hiệu lực với tất cả các điểm x trong vùng kết nối không có nguồn tương ứng với vùng hội tụ của các dãy.

Theo phương trình (4), k biểu thị số sóng có góc được xác định bằng

$$k := \frac{\omega}{c_s} \quad (5)$$

và $p_n^m(kr)$ là hệ số mở rộng SH chỉ phụ thuộc vào kết quả kr .

Ngoài ra, $Y_n^m(\theta, \phi)$ là hàm SH bậc n và độ m :

$$Y_n^m(\theta, \phi) := \sqrt{\frac{(2n+1)(n-m)!}{4\pi(n+m)!}} P_n^m(\cos \theta) e^{im\phi}, \quad (6)$$

trong đó, $P_n^m(\cos \theta)$ là hàm Legendre kết hợp và $(\cdot)!$ là giai thừa.

Hàm Legendre kết hợp được dùng cho các chỉ số mức không âm m được xác

định qua đa thức Legendre $P_n(x)$ là

$$P_n^m(x) := (-1)^m (1-x^2)^{\frac{m}{2}} \frac{d^m}{dx^m} P_n(x) \quad \text{với } m \geq 0. \quad (7)$$

Đối với các chỉ số mức âm, tức là $m < 0$, hàm Legendre kết hợp được xác định bằng

$$P_n^m(x) := (-1)^m \frac{(n+m)!}{(n-m)!} P_n^{-m}(x) \quad \text{với } m < 0. \quad (8)$$

Đa thức Legendre $P_n(x)$ ($n \geq 0$) có thể được xác định lần lượt bằng cách sử dụng công thức của Rodrigues là

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} (x^2 - 1)^n. \quad (9)$$

Theo kỹ thuật cũ, ví dụ theo tài liệu: M. Poletti, "Unified Description of Ambisonics using Real and Complex Spherical Harmonics", Báo cáo chuyên đề Ambisonics 2009, từ ngày 25 đến 27 tháng 6 năm 2009, Graz, Austria, còn có các định nghĩa của hàm SH mà khác với phương trình (6) bằng hệ số $(-1)^m$ cho các chỉ số mức độ âm.

Cách khác, phép biến đổi Fourier của áp suất âm thanh liên quan đến thời gian có thể được biểu diễn bằng cách sử dụng hàm SH thực $S_n^m(\theta, \phi)$ là

$$P(kc_s, (r, \theta, \phi)^T) = \sum_{n=0}^{\infty} \sum_{m=-n}^n q_n^m(kr) S_n^m(\theta, \phi) \quad (10)$$

Trong tài liệu kỹ thuật, có các định nghĩa khác nhau về hàm SH thực (ví dụ, xem tài liệu của Poletti nêu trên). Có thể có một cách định nghĩa được áp dụng trong suốt bản mô tả này, được xác định bởi

$$S_n^m(\theta, \phi) := \begin{cases} \frac{(-1)^m}{\sqrt{2}} [Y_n^m(\theta, \phi) + Y_n^{m*}(\theta, \phi)] & \text{với } m > 0 \\ Y_n^m(\theta, \phi) & \text{với } m = 0, \\ \frac{(-1)}{i\sqrt{2}} [Y_n^m(\theta, \phi) - Y_n^{m*}(\theta, \phi)] & \text{với } m < 0 \end{cases} \quad (11)$$

trong đó $(\cdot)^*$ biểu thị phép liên hợp phức. Biểu thức khác thu được bằng cách chèn phương trình (6) vào phương trình (11):

$$S_n^m(\theta, \phi) = \sqrt{\frac{(2n+1)(n-m)!}{4\pi(n+m)!}} P_n^m(\cos\theta) \text{trg}_m(\phi), \quad (12)$$

với

$$\text{trg}_m(\phi) := \begin{cases} (-1)^m \sqrt{2} \cos(m\phi) & \text{với } m > 0 \\ 1 & \text{với } m = 0, \\ -\sqrt{2} \sin(m\phi) & \text{với } m < 0 \end{cases} \quad (13)$$

Mặc dù các hàm SH thực có giá trị thực theo định nghĩa, nhưng điều này

không đúng với các hệ số mở rộng tương ứng $q_n^m(kr)$ nói chung.

Hàm SH phức liên quan đến hàm SH thực như sau:

$$Y_n^m(\theta, \phi) = \begin{cases} \frac{q_n^m(kr)}{\sqrt{2}} [S_n^m(\theta, \phi) + iS_n^{-m}(\theta, \phi)] & \text{với } m > 0 \\ S_n^0(\theta, \phi) & \text{với } m = 0 \\ \frac{1}{i\sqrt{2}} [S_n^m(\theta, \phi) + iS_n^{-m}(\theta, \phi)] & \text{với } m < 0 \end{cases} \quad (14)$$

Hàm SH phức $Y_n^m(\theta, \phi)$ cũng như hàm SH thực $S_n^m(\theta, \phi)$ với vector có hướng $\Omega := (\theta, \phi)^T$ tạo ra cơ sở trực giao cho các hàm có giá trị phức bình phương khả tích trên mặt cầu đơn vị S^2 trong không gian ba chiều, và do đó tuân theo các điều kiện

$$\begin{aligned} \int_{S^2} Y_n^m(\Omega) Y_{n'}^{m'}{}^*(\Omega) d\Omega &= \int_0^{2\pi} \int_0^\pi Y_n^m(\theta, \phi) Y_{n'}^{m'}{}^*(\theta, \phi) \sin\theta d\theta d\phi \\ &= \delta_{n-n'} \delta_{m-m'} \end{aligned} \quad (15)$$

$$\int_{S^2} S_n^m(\Omega) S_{n'}^{m'}{}^*(\Omega) d\Omega = \delta_{n-n'} \delta_{m-m'}, \quad (16)$$

trong đó δ biểu thị hàm delta Kronecker. Kết quả thứ hai có thể được suy ra bằng cách sử dụng phương trình (15) và định nghĩa của hàm điều hòa cầu thực trong phương trình (11).

Vấn đề bên trong và các hệ số Ambisonic

Mục đích của ambisonic là dạng biểu diễn của trường âm thanh ở vùng lân cận của điểm gốc tọa độ. Không mất đi tính tổng quát, ở đây vùng quan tâm này được giả định là khối cầu bán kính R có tâm ở điểm gốc tọa độ được xác định theo tập hợp $\{x | 0 \leq r \leq R\}$. Giả thiết quan trọng cho dạng biểu diễn là khối cầu này được giả định không chứa các nguồn âm bất kỳ. Quy trình tìm kiếm dạng biểu diễn của trường âm thanh trong khối cầu này bị giới hạn là 'vấn đề bên trong', ví dụ, trong tài liệu của Williams ở trên.

Có thể thấy rằng đối với vấn đề bên trong, các hệ số mở rộng của hàm SH $p_n^m(kr)$ có thể được biểu diễn là

$$p_n^m(kr) = a_n^m(k) j_n(kr), \quad (17)$$

trong đó $j_n(\cdot)$ biểu thị các hàm Bessel hình cầu bậc một. Từ phương trình (17) tiếp theo là thông tin đầy đủ về trường âm thanh được bao gồm trong các hệ số $a_n^m(k)$, các hệ số này được dùng để chỉ các hệ số ambisonic.

Tương tự, hệ số của phần mở rộng của hàm SH thực $q_n^m(kr)$ có thể được biểu

diễn là

$$q_n^m(kr) = b_n^m(k)j_n(kr), \quad (18)$$

trong đó các hệ số $b_n^m(k)$ được dùng để chỉ các hệ số Ambisonic liên quan đến phần mở rộng bằng cách sử dụng các hàm SH giá trị thực. Các hệ số này liên quan đến $a_n^m(k)$ thông qua

$$b_n^m(k) = \begin{cases} \frac{1}{\sqrt{2}} [(-1)^m a_n^m(k) + a_n^{-m}(k)] & \text{với } m > 0 \\ a_n^0(k) & \text{với } m = 0 \\ \frac{1}{i\sqrt{2}} [a_n^m(k) - (-1)^m a_n^{-m}(k)] & \text{với } m < 0 \end{cases} \quad (19)$$

Phân tích sóng phẳng

Trường âm thanh trong khối cầu không có nguồn âm thanh có tâm ở điểm gốc tọa độ có thể được biểu diễn bởi sự chồng chéo số lượng vô hạn các sóng phẳng có các số sóng góc khác nhau k , ảnh hưởng đến khối cầu từ tất cả các hướng có thể, ví dụ như, như được thể hiện trong tài liệu "Plane-wave decomposition ..." của Rafaely nêu trên. Giả sử rằng biên độ phức của sóng phẳng có số sóng góc k từ hướng Ω_0 được xác định bởi $D(k, \Omega_0)$, có thể được thể hiện theo cách tương tự bằng cách sử dụng phương trình (11) và phương trình (19) mà hệ số ambisonic tương ứng liên quan đến phần mở rộng hàm SH thực được xác định bởi

$$b_{n,\text{sóng phẳng}}^m(k; \Omega_0) = 4\pi i^n D(k, \Omega_0) S_n^m(\Omega_0). \quad (20)$$

Do đó, các hệ số ambisonic cho trường âm thanh do sự chồng chéo của số lượng vô hạn sóng phẳng của số sóng góc k thu được từ phép tích phân trong phương trình (20) trên tất cả các hướng có thể $\Omega_0 \in S^2$:

$$b_n^m(k) = \int_{S^2} b_{n,\text{sóng phẳng}}^m(k; \Omega_0) d\Omega_0 \quad (21)$$

$$= 4\pi i^n \int_{S^2} D(k, \Omega_0) S_n^m(\Omega_0) d\Omega_0. \quad (22)$$

Hàm $D(k, \Omega)$ được gọi là 'mật độ biên' và được giả định là bình phương khả tích trên mặt cầu đơn vị S^2 . Hàm này có thể được mở rộng thành các dãy hàm SH thực là

$$D(k, \Omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^n c_n^m(k) S_n^m(\Omega), \quad (23)$$

trong đó các hệ số mở rộng $c_n^m(k)$ bằng phép tích phân xuất hiện trong phương trình (22), tức là

$$c_n^m(k) = \int_{S^2} D(k, \Omega) S_n^m(\Omega) d\Omega. \quad (24)$$

Bằng cách chèn phương trình (24) vào phương trình (22) có thể thấy rằng hệ

số ambisonic $b_n^m(k)$ là phiên bản được định tỷ lệ của các hệ số mở rộng $c_n^m(k)$, tức là

$$b_n^m(k) = 4\pi i^n c_n^m(k). \quad (25)$$

Khi áp dụng phép biến đổi Fourier ngược liên quan đến thời gian cho hệ số ambisonic được định tỷ lệ $c_n^m(k)$ và cho hàm mật độ biên $D(k, \Omega)$, thu được số lượng miền thời gian tương ứng

$$\tilde{c}_n^m(t) := \mathcal{F}_t^{-1} \left\{ c_n^m \left(\frac{\omega}{c_s} \right) \right\} = \frac{1}{2\pi} \int_{-\infty}^{\infty} c_n^m \left(\frac{\omega}{c_s} \right) e^{i\omega t} d\omega \quad (26)$$

$$d(t, \Omega) := \mathcal{F}_t^{-1} \left\{ D \left(\frac{\omega}{c_s}, \Omega \right) \right\} = \frac{1}{2\pi} \int_{-\infty}^{\infty} D \left(\frac{\omega}{c_s}, \Omega \right) e^{i\omega t} d\omega \quad (27)$$

Sau đó, trong miền thời gian, phương trình (24) có thể được tính bằng

$$\tilde{c}_n^m(t) = \int_{S^2} d(t, \Omega) S_n^m(\Omega) d\Omega. \quad (28)$$

Tín hiệu có hướng trong miền thời gian $d(t, \Omega)$ có thể được biểu diễn bằng phần mở rộng hàm SH thực theo

$$d(t, \Omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^n \tilde{c}_n^m(t) S_n^m(\Omega). \quad (29)$$

Bằng cách sử dụng các hàm SH $S_n^m(\Omega)$ có các giá trị thực, dạng liên hợp phức của nó có thể được biểu diễn bằng

$$d^*(t, \Omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^n \tilde{c}_n^{m*}(t) S_n^m(\Omega). \quad (30)$$

Giả sử tín hiệu miền thời gian $d(t, \Omega)$ là giá trị thực, tức là $d(t, \Omega) = d^*(t, \Omega)$, điều này tiếp theo việc so sánh phương trình (29) với phương trình (30) mà hệ số $\tilde{c}_n^{m*}(t)$ là giá trị thực trong trường hợp này, tức là $\tilde{c}_n^m(t) = \tilde{c}_n^{m*}(t)$.

Các hệ số $\tilde{c}_n^m(t)$ sẽ được dùng để chỉ các hệ số ambisonic miền thời gian được định tỷ lệ trong phần mô tả sau đây.

Sau đây giả sử rằng dạng biểu diễn trường âm thanh được đưa ra theo các hệ số này, việc này sẽ được mô tả chi tiết hơn trong phần dưới đây liên quan đến việc nén.

Lưu ý rằng dạng biểu diễn tín hiệu HOA miền thời gian bằng hệ số $\tilde{c}_n^m(t)$ được dùng cho quy trình xử lý theo sáng chế tương đương với dạng biểu diễn tín hiệu HOA miền tần số tương ứng $c_n^m(k)$. Do đó, kỹ thuật nén và giải nén được mô tả có thể được thực hiện tương đương trong miền tần số với sự biến đổi nhỏ tương ứng của các phương trình.

Độ phân giải không gian có bậc hữu hạn

Trên thực tế, trường âm thanh ở vùng lân cận của điểm gốc tọa độ được mô tả

bằng cách chỉ sử dụng số lượng hữu hạn hệ số ambisonic $c_n^m(k)$ bậc $n \leq N$. Việc tính toán hàm mật độ biên từ các dãy cắt cụt của hàm SH theo

$$D_N(k, \Omega) := \sum_{n=0}^N \sum_{m=-n}^n c_n^m(k) S_n^m(\Omega) \quad (31)$$

đưa ra loại phân tán không gian so với hàm mật độ biên đúng $D(k, \Omega)$, ví dụ tài liệu "Plane-wave decomposition..." nêu trên. Điều này có thể được thực hiện bằng cách tính toán hàm mật độ biên cho sóng phẳng đơn từ hướng Ω_0 bằng cách sử dụng phương trình (31):

$$D_N(k, \Omega) = \sum_{n=0}^N \sum_{m=-n}^n \frac{1}{4\pi i^n} \cdot b_{n, \text{sóng phẳng}}^m(k; \Omega_0) S_n^m(\Omega) \quad (32)$$

$$= D(k, \Omega_0) \sum_{n=0}^N \sum_{m=-n}^n S_n^m(\Omega_0) S_n^m(\Omega) \quad (33)$$

$$= D(k, \Omega_0) \sum_{n=0}^N \sum_{m=-n}^n Y_n^{m*}(\Omega_0) Y_n^m(\Omega) \quad (34)$$

$$= D(k, \Omega_0) \sum_{n=0}^N \frac{2n+1}{4\pi} P_n(\cos \Theta) \quad (35)$$

$$= D(k, \Omega_0) \left[\frac{N+1}{4\pi(\cos \Theta - 1)} (P_{N+1}(\cos \Theta) - P_N(\cos \Theta)) \right] \quad (36)$$

$$= D(k, \Omega_0) v_N(\Theta) \quad (37)$$

với

$$v_N(\Theta) := \frac{N+1}{4\pi(\cos \Theta - 1)} (P_{N+1}(\cos \Theta) - P_N(\cos \Theta)), \quad (38)$$

trong đó Θ biểu thị góc giữa hai vector trở đến các hướng Ω và Ω_0 thỏa mãn điều kiện

$$\cos \Theta = \cos \theta \cos \theta_0 + \cos(\phi - \phi_0) \sin \theta \sin \theta_0. \quad (39)$$

Theo phương trình (34) các hệ số ambisonic cho sóng phẳng xác định trong phương trình (20) được sử dụng, trong khi ở phương trình (35) và (36), một số định lý toán học được vận dụng, ví dụ, tài liệu "Plane-wave decomposition..." ở trên. Tính chất trong phương trình (33) có thể được thể hiện bằng cách sử dụng phương trình (14).

So sánh phương trình (37) với hàm mật độ biên đúng

$$(k, \Omega) = D(k, \Omega_0) \frac{\delta(\Theta)}{2\pi}, \quad (40)$$

trong đó $\delta(\cdot)$ biểu thị hàm delta Dirac, sự phân bố không gian trở nên rõ ràng từ việc thay thế hàm delta Dirac được định tỷ lệ bằng hàm phân tán $v_N(\Theta)$ mà, sau khi được chuẩn hóa bằng giá trị cực đại của nó, được minh họa trên Fig.1 đối với các bậc ambisonic khác nhau N và các góc $\Theta \in [0, \pi]$.

Do số không đầu tiên $v_N(\Theta)$ được đặt gần ở $\frac{\pi}{N}$ với $N \geq 4$ (xem "Plane-wave decomposition..." nêu trên), hiệu ứng phân tán bị giảm (và do đó độ phân giải không gian được cải thiện) khi tăng bậc ambisonic N .

Với $N \rightarrow \infty$ hàm phân tán $v_N(\theta)$ hội tụ với hàm delta Dirac được định tỷ lệ. Có thể thấy được điều này nếu hệ thức đầy đủ cho đa thức Legendre

$$\sum_{n=0}^{\infty} \frac{2n+1}{2} P_n(x) P_n(x') = \delta(x - x') \quad (41)$$

được sử dụng cùng với phương trình (35) để biểu diễn giới hạn $v_N(\theta)$ với $N \rightarrow \infty$ là

$$\lim_{N \rightarrow \infty} v_N(\theta) = \frac{1}{2\pi} \sum_{n=0}^{\infty} \frac{2n+1}{2} P_n(\cos\theta) \quad (42)$$

$$= \frac{1}{2\pi} \sum_{n=0}^{\infty} \frac{2n+1}{2} P_n(\cos\theta) P_n(1) \quad (43)$$

$$= \frac{1}{2\pi} \delta(\cos\theta - 1) \quad (44)$$

$$= \frac{1}{2\pi} \delta(\theta). \quad (45)$$

Khi xác định vector của hàm SH thực bậc $n \leq N$ bởi

$$(\mathbf{\Omega}) := (S_0^0(\mathbf{\Omega}), S_1^{-1}(\mathbf{\Omega}), S_1^0(\mathbf{\Omega}), S_1^1(\mathbf{\Omega}), S_2^{-2}(\mathbf{\Omega}), \dots, S_N^N(\mathbf{\Omega}))^T \in \mathbb{R}^O, \quad (46)$$

trong đó $O = (N + 1)^2$ và trong đó $(\cdot)^T$ biểu thị sự chuyển vị, sự so sánh phương trình (37) với phương trình (33) thể hiện rằng hàm phân tán có thể được biểu diễn qua tích vô hướng của hai vector SH thực là $v_N(\theta) = \mathbf{S}^T(\mathbf{\Omega}) \mathbf{S}(\mathbf{\Omega}_0)$. (47)

Sự phân tán có thể được biểu diễn một cách tương đương trong miền thời gian là

$$d_N(t, \mathbf{\Omega}) := \sum_{n=0}^N \sum_{m=-n}^n \tilde{c}_n^m(t) S_n^m(\mathbf{\Omega}) \quad (48)$$

$$= d(t, \mathbf{\Omega}_0) v_N(\theta). \quad (49)$$

Quy trình lấy mẫu

Đối với một số ứng dụng, mong muốn là xác định các hệ số ambisonic miền thời gian được định tỷ lệ $\tilde{c}_n^m(t)$ từ các mẫu của hàm mật độ biên miền thời gian $d(t, \mathbf{\Omega})$ với số lượng hữu hạn J của các hướng rời rạc $\mathbf{\Omega}_j$. Phép tích phân ở phương trình (28) sau đó được tính xấp xỉ bằng tổng hữu hạn theo B. Rafaely, "Analysis and Design of Spherical Microphone Arrays", IEEE Transactions on Speech and Audio Processing, Tập 13, số 1, từ trang 135 đến 143, tháng một năm 2005:

$$\tilde{c}_n^m(t) \approx \sum_{j=1}^J g_j \cdot d(t, \mathbf{\Omega}_j) S_n^m(\mathbf{\Omega}_j), \quad (50)$$

trong đó g_j biểu thị một số trọng số lấy mẫu được chọn thích hợp. Tuy nhiên, trái với tài liệu "Analysis and Design ...", phép tính xấp xỉ (50) được dùng để chỉ dạng biểu diễn miền thời gian bằng cách sử dụng hàm SH thực chứ không phải là dạng biểu diễn miền tần số sử dụng các hàm SH phức. Điều kiện cần thiết để phép tính xấp xỉ (50) trở nên chính xác là mật độ biên có bậc điều hòa giới hạn N , có nghĩa là

$$\tilde{c}_n^m(t) = 0 \quad \text{với } n > N. \quad (51)$$

Nếu điều kiện này không được đáp ứng, thì phép tính xấp xỉ (50) bị sai số lấy mẫu không gian, chẳng hạn, trong tài liệu B. Rafaely, "Spatial Aliasing in Spherical Microphone Arrays", IEEE Transactions on Signal Processing, tập 55, số 3, trang từ 1003 đến 1010, tháng ba năm 2007.

Điều kiện cần thiết thứ hai đòi hỏi các điểm lấy mẫu Ω_j và các trọng số tương ứng thỏa mãn các điều kiện tương ứng xác định trong tài liệu "Analysis and Design ...":

$$\sum_{j=1}^J g_j S_{n'}^{m'}(\Omega_j) S_n^m(\Omega_j) = \delta_{n-n'} \delta_{m-m'} \quad \text{với } m, m' \leq N. \quad (52)$$

Các điều kiện (51) và (52) đáp ứng đồng thời đầy đủ để lấy mẫu chính xác.

Điều kiện lấy mẫu (52) bao gồm tập hợp các phương trình tuyến tính, các phương trình này có thể được tính toán ngắn gọn bằng cách sử dụng phương trình ma trận đơn là $\Psi^H = I$, (53)

trong đó Ψ là ma trận chế độ được xác định bằng

$$\Psi := [S(\Omega_1) \dots S(\Omega_J)] \in R^{O \times J} \quad (54)$$

và G biểu thị ma trận có các trọng số trên đường chéo của nó, tức là

$$G := \text{diag}(g_1, \dots, g_J). \quad (55)$$

Từ phương trình (53) có thể thấy rằng điều kiện cần thiết để đáp ứng cho phương trình (52) là số lượng điểm lấy mẫu J thỏa mãn $J \geq O$. Ở các điểm lấy mẫu J tập hợp các giá trị của mật độ biên miền thời gian thành vector

$$\mathbf{w}(t) := (D(t, \Omega_1), \dots, D(t, \Omega_J))^T, \quad (56)$$

và xác định vector của các hệ số ambisonic miền thời gian được định tỷ lệ theo

$$\mathbf{c}(t) := (\tilde{c}_0^0(t), \tilde{c}_1^{-1}(t), \tilde{c}_1^0(t), \tilde{c}_1^1(t), \tilde{c}_2^{-2}(t), \dots, \tilde{c}_O^O(t))^T, \quad (57)$$

cả hai vector có liên quan đến việc mở rộng của các hàm SH (29). Hệ thức này tạo ra theo hệ phương trình tuyến tính: $\mathbf{w}(t) = \Psi^H \mathbf{c}(t)$. (58)

Bằng cách sử dụng ký hiệu vector được đưa vào, việc tính toán các hệ số ambisonic miền thời gian được định tỷ lệ từ các giá trị của các mẫu hàm mật độ biên miền thời gian có thể được viết là $\mathbf{w}(t) \approx \Psi G \mathbf{w}(t)$. (59)

Nếu bậc ambisonic N được cố định, thì bậc này thường không thể tính số lượng $J \geq O$ của các điểm lấy mẫu Ω_j và các trọng số tương ứng sao cho phương

trình (52) đáp ứng điều kiện lấy mẫu. Tuy nhiên, nếu các điểm lấy mẫu được chọn sao cho điều kiện lấy mẫu được làm xấp xỉ, thì hạng của ma trận chế độ Ψ là O và số điều kiện của nó thấp. Trong trường hợp này, có phương trình giả ngược $\Psi^+ := (\Psi\Psi^H)^{-1}\Psi\Psi^+$ (60) của ma trận chế độ Ψ và phép tính xấp xỉ hợp lý của vector hệ số ambisonic miền thời gian được định tỷ lệ $c(t)$ từ vector của các mẫu hàm mật độ biên miền thời gian được xác định bởi $c(t) \approx \Psi^+w(t)$. (61)

Nếu $J = O$ và hạng của ma trận chế độ là O thì phương trình giả ngược của nó trùng với phương trình nghịch đảo của nó vì

$$\Psi^+ = (\Psi\Psi^H)^{-1}\Psi = \Psi^{-H}\Psi^{-1}\Psi = \Psi^{-H}. \quad (62)$$

Ngoài ra nếu điều kiện lấy mẫu của phương trình (52) được thỏa mãn, thì $\Psi^{-H} = \Psi G$ (63) và cả hai phép xấp xỉ trong phương trình (59) và (61) là tương đương và được tính chính xác.

Vector $w(t)$ có thể được hiểu là vector của các tín hiệu không gian miền thời gian. Việc biến đổi từ miền tín hiệu HOA sang miền không gian có thể được thực hiện, ví dụ bằng cách sử dụng phương trình (58). Loại biến đổi này được gọi là 'biến đổi hàm điều hòa cầu (SHT)' trong ứng dụng này và được sử dụng khi thành phần tín hiệu HOA xung quanh có bậc giảm được biến đổi sang miền không gian. Giả định ngầm rằng các điểm lấy mẫu không gian Ω_j cho SHT gần như thỏa mãn điều kiện lấy mẫu trong phương trình (52) với $g_j \approx \frac{4\pi}{O}$ với $j = 1, \dots, J$ và $J = O$. Theo sự giả định này SHT ma trận thỏa mãn $\Psi^H \approx \frac{4\pi}{O}\Psi^{-1}$. Trong trường hợp việc định tỷ lệ tuyệt đối cho SHT là không quan trọng, hằng số $\frac{4\pi}{O}$ có thể được bỏ qua.

Nén

Sáng chế đề cập đến phương pháp nén dạng biểu diễn tín hiệu HOA đã cho. Như được nêu trên đây, dạng biểu diễn tín hiệu HOA được phân tích thành số tín hiệu có hướng chính định trước trong miền thời gian và thành phần xung quanh trong miền tín hiệu HOA, tiếp theo là nén dạng biểu diễn tín hiệu HOA của thành phần xung quanh bằng cách giảm bậc của nó. Thao tác này sử dụng giả định mà được hỗ trợ bằng các quy trình thử nghiệm nghe, trong đó thành phần trường âm xung quanh có thể được biểu diễn với độ chính xác đầy đủ bằng dạng biểu diễn tín hiệu HOA có bậc thấp. Quy trình trích các tín hiệu có hướng chính đảm bảo rằng, tiếp theo quy

trình nén và giải nén tương ứng, độ phân giải không gian cao được giữ lại.

Sau khi phân tích, thành phần HOA xung quanh có bậc đã giảm được biến đổi sang miền không gian, và được mã hóa cảm quan cùng với các tín hiệu có hướng như được mô tả trong phần *các phương án làm ví dụ* của đơn sáng chế châu Âu số EP 10306472.1.

Quy trình nén bao gồm hai bước liên tiếp được mô tả trên Fig. 2. Định nghĩa chính xác về các tín hiệu riêng lẻ được mô tả trong phần *Chi tiết về phương pháp nén* sau đây.

Ở bước hoặc giai đoạn thứ nhất thể hiện trên Fig.2a, trong bộ ước tính hướng chính 22, các hướng chính được ước tính và quy trình phân tích tín hiệu Ambisonic $C(l)$ thành thành phần có hướng và thành phần dư hoặc thành phần xung quanh được thực hiện, trong đó l biểu thị chỉ số khung hình. Thành phần có hướng được tính ở bước tính toán tín hiệu có hướng hoặc bước 23, do đó dạng biểu diễn ambisonic được chuyển đổi sang tín hiệu miền thời gian được biểu diễn bằng tập hợp D tín hiệu có hướng thông thường $X(l)$ có các hướng tương ứng $\bar{\Omega}_{\text{DOM}}(l)$. Thành phần xung quanh dư được tính toán ở bước tính toán thành phần HOA xung quanh hoặc bước 24, và được biểu diễn bởi các hệ số miền tín hiệu $\text{HOAC}_A(l)$.

Ở bước thứ hai thể hiện trên Fig.2b, quy trình mã hóa cảm quan các tín hiệu có hướng $X(l)$ và thành phần tín hiệu HOA xung quanh $C_A(l)$ được thực hiện như sau:

- Các tín hiệu có hướng miền thời gian thông thường $X(l)$ có thể là được nén riêng trong bộ mã hóa cảm quan 27 bằng cách sử dụng kỹ thuật nén cảm quan đã biết bất kỳ.
- Quy trình nén thành phần miền tín hiệu HOA xung quanh $C_A(l)$ được thực hiện ở hai giai đoạn hoặc các bước phụ.

Giai đoạn hoặc bước phụ thứ nhất 25 thực hiện giảm bậc ambisonic gốc từ N xuống N_{RED} , ví dụ $N_{\text{RED}} = 2$, thu được thành phần HOA xung quanh $C_{A,\text{RED}}(l)$. Ở đây, giả sử rằng thành phần trường âm xung quanh có thể được biểu diễn với độ chính xác đầy đủ bởi tín hiệu HOA bậc thấp. Giai đoạn hoặc bước phụ thứ hai 26 dựa trên quy trình nén được mô tả trong đơn sáng chế châu Âu số EP 10306472.1. Các tín hiệu HOA $O_{\text{RED}} := (N_{\text{RED}} + 1)^2 C_{A,\text{RED}}(l)$ của thành phần trường âm xung quanh đã được tính ở giai đoạn/bước phụ 25, được biến đổi thành tín hiệu tương đương O_{RED}

$W_{A,RED}(l)$ trong miền không gian bằng cách áp dụng phép biến đổi hàm điều hòa cầu, tạo ra các tín hiệu miền thời gian thông thường có thể được nhập vào dãy các bộ mã hóa-giải mã cảm quan song song 27. Kỹ thuật mã hóa hoặc nén cảm quan đã biết bất kỳ có thể được áp dụng. Các tín hiệu có hướng được mã hóa $\tilde{X}(l)$ và tín hiệu miền không gian được mã hóa giảm bậc $\tilde{W}_{A,RED}(l)$ được xuất ra và có thể được truyền hoặc lưu trữ.

Có lợi là, quy trình nén cảm quan tất cả các tín hiệu miền thời gian $X(l)$ và $W_{A,RED}(l)$ có thể được thực hiện đồng thời trong bộ mã hóa cảm quan 27 để cải thiện hiệu quả mã hóa toàn phần bằng cách sử dụng quan hệ tương quan giữa các kênh có thể có.

Giải nén

Quy trình giải nén cho tín hiệu nhận được hoặc được phát lại được mô tả trên Fig.3. Tương tự quy trình nén, quy trình giải nén bao gồm hai bước liên tiếp.

Ở bước hoặc giai đoạn thứ nhất thể hiện trên Fig.3a, ở bước giải mã cảm quan 31, quy trình giải mã cảm quan hoặc giải nén các tín hiệu có hướng được mã hóa $\tilde{X}(l)$ và tín hiệu miền không gian được mã hóa giảm bậc $\tilde{W}_{A,RED}(l)$ được thực hiện, trong đó $\tilde{X}(l)$ là thành phần biểu diễn và $\tilde{W}_{A,RED}(l)$ là thành phần tín hiệu HOA xung quanh. Tín hiệu miền không gian được giải mã cảm quan hoặc được giải nén $\tilde{W}_{A,RED}(l)$ được biến đổi ở bộ biến đổi hàm điều hòa cầu ngược 32 sang dạng biểu diễn miền tín hiệu HOA $\hat{C}_{A,RED}(l)$ bậc N_{RED} qua phép biến đổi hàm điều hòa cầu ngược. Sau đó, ở bước hoặc giai đoạn mở rộng bậc 33, dạng biểu diễn tín hiệu HOA phù hợp $\hat{C}_A(l)$ bậc N được ước tính từ $\hat{C}_{A,RED}(l)$ bằng cách mở rộng bậc.

Ở bước hoặc giai đoạn thứ hai thể hiện trên Fig. 3b, dạng biểu diễn tín hiệu HOA tổng $\hat{C}(l)$ được biên soạn lại trong bộ hợp dịch tín hiệu HOA 34 từ các tín hiệu có hướng $\hat{X}(l)$ và thông tin hướng tương ứng $\bar{\Omega}_{DOM}(l)$ cũng như từ thành phần HOA xung quanh bậc ban đầu $\hat{C}_A(l)$.

Giảm tốc độ dữ liệu có thể thu được

Một vấn đề được giải quyết theo sáng chế là giảm đáng kể tốc độ dữ liệu so với các phương pháp nén hiện có đối với các dạng biểu diễn tín hiệu HOA. Trong phần mô tả sau đây, tỷ lệ nén thu được so với dạng biểu diễn tín hiệu HOA không nén

được mô tả. Tỷ lệ nén thu được từ quá trình so sánh tốc độ dữ liệu cần thiết cho cuộc truyền tín hiệu HOA không nén $C(l)$ bậc N với tốc độ dữ liệu cần thiết cho cuộc truyền dạng biểu diễn được nén bao gồm D tín hiệu có hướng được mã hóa cảm quan $X(l)$ với các hướng tương ứng $\bar{\Omega}_{\text{DOM}}(l)$ và N_{RED} tín hiệu miền không gian được mã hóa cảm quan $W_{A,\text{RED}}(l)$ biểu diễn thành phần tín hiệu HOA xung quanh.

Để truyền tín hiệu HOA không nén $C(l)$ cần thiết có tốc độ dữ liệu $O \cdot f_s \cdot N_b$. Ngược lại, việc truyền D tín hiệu có hướng được mã hóa cảm quan $X(l)$ cần có tốc độ dữ liệu là $D \cdot f_{b,\text{COD}}$, trong đó $f_{b,\text{COD}}$ biểu thị tốc độ bit của các tín hiệu được mã hóa cảm quan. Tương tự, việc truyền tín hiệu miền không gian được mã hóa cảm quan N_{RED} , các tín hiệu $W_{A,\text{RED}}(l)$ cần tốc độ bit là $O_{\text{RED}} \cdot f_{b,\text{COD}}$. Các hướng $\bar{\Omega}_{\text{DOM}}(l)$ được giả định sẽ được tính toán dựa trên tốc độ thấp hơn nhiều so với tốc độ lấy mẫu f_s , tức là các hướng được giả định sẽ được cố định trong khoảng thời gian của khung tín hiệu bao gồm B mẫu, ví dụ $B = 1200$ cho tốc độ lấy mẫu là $f_s = 48\text{kHz}$, và quy trình dùng chung tốc độ dữ liệu tương ứng có thể được bỏ qua để tính toán tốc độ dữ liệu tổng của tín hiệu HOA được nén.

Do đó, cuộc truyền dạng biểu diễn được nén cần tốc độ dữ liệu khoảng $(D + O_{\text{RED}}) \cdot f_{b,\text{COD}}$. Do đó, tỷ lệ nén r_{COMPR} là

$$r_{\text{COMPR}} \approx \frac{O \cdot f_s \cdot N_b}{(D + O_{\text{RED}}) \cdot f_{b,\text{COD}}} \quad (64)$$

Ví dụ, việc nén dạng biểu diễn tín hiệu HOA bậc $N = 4$ sử dụng tốc độ lấy mẫu $f_s = 48\text{kHz}$ và $N_b = 16$ bit trên mỗi mẫu thành dạng biểu diễn có $D = 3$ hướng chính sử dụng bậc tín hiệu HOA đã giảm $N_{\text{RED}} = 2$ và tốc độ bit $64 \frac{\text{kbit}}{\text{s}}$ sẽ dẫn đến tỷ lệ nén là $r_{\text{COMPR}} \approx 25$. Cuộc truyền dạng biểu diễn được nén cần tốc độ dữ liệu khoảng là $768 \frac{\text{kbit}}{\text{s}}$.

Giảm khả năng xảy ra việc để lộ nhiễu mã hóa

Như được mô tả trong phần tình trạng kỹ thuật, quy trình nén cảm quan các tín hiệu miền không gian được mô tả trong đơn sáng chế châu Âu số EP 10306472.1 bị ảnh hưởng bởi các mối tương quan chéo còn lại giữa các tín hiệu mà có thể dẫn đến việc để lộ nhiễu mã hóa cảm quan. Theo sáng chế, các tín hiệu có hướng chính trước tiên được tách từ dạng biểu diễn trường âm thanh tín hiệu HOA trước khi được mã

hóa cảm quan. Tức là, khi biên soạn dạng biểu diễn tín hiệu HOA, sau khi giải mã cảm quan, nhiều mã hóa có độ định hướng không gian chính xác giống như các tín hiệu có hướng. Cụ thể là, mức đóng góp của nhiều mã hóa cũng như tín hiệu có hướng cho hướng tùy ý bất kỳ được mô tả tất định bằng hàm phân tán không gian được giải thích trong phần *Độ phân giải không gian có bậc hữu hạn*. Nói cách khác, vào bất kỳ thời điểm nào thì vectơ hệ số HOA biểu diễn nhiều mã hóa là bội số của vectơ hệ số HOA biểu diễn tín hiệu có hướng. Do đó, tổng trọng số tùy ý của các hệ số HOA có nhiều sẽ không dẫn đến việc để lộ bất kỳ nhiều mã hóa cảm quan.

Ngoài ra, thành phần xung quanh có bậc đã giảm được xử lý chính xác như được đề xuất trong đơn sáng chế châu Âu số EP 10306472.1, nhưng do theo định nghĩa, các tín hiệu miền không gian của thành phần xung quanh có mối tương quan khá thấp, nên xác suất để lộ nhiều cảm quan là thấp.

Ước tính hướng được cải thiện

Việc ước tính hướng theo sáng chế phụ thuộc vào sự phân bố công suất có hướng của thành phần tín hiệu HOA chính có năng lượng cao. Sự phân bố công suất có hướng được tính từ ma trận tương quan bị giảm hạng của dạng biểu diễn tín hiệu HOA thu được bằng cách phân tích giá trị riêng của ma trận tương quan của dạng biểu diễn tín hiệu HOA.

So với quy trình ước tính hướng được dùng trong tài liệu "Plane-wave decomposition ..." nêu trên, quy trình này có ưu điểm là chính xác hơn do việc tập trung vào thành phần HOA chính có năng lượng cao thay vì sử dụng dạng biểu diễn tín hiệu HOA hoàn chỉnh để ước tính hướng giúp làm giảm độ mờ trong không gian về sự phân bố công suất có hướng.

So với quy trình ước tính hướng được đề xuất trong tài liệu: " The Application of Compressive Sampling to the Analysis and Synthesis of Spatial Sound Fields" và "Time Domain Reconstruction of Spatial Sound Fields Using Compressed Sensing " nêu trên, quy trình này có ưu điểm ổn định hơn. Lý do là ở chỗ quy trình phân tích dạng biểu diễn tín hiệu HOA thành thành phần có hướng và thành phần xung quanh hầu như không thể được thực hiện tốt, do đó vẫn có nhiều thành phần xung quanh nhỏ trong thành phần có hướng. Sau đó, phương pháp lấy mẫu nén giống như trong hai tài liệu này không thực hiện ước tính hướng hợp lý do độ nhạy cao của

chúng so với sự có mặt của các tín hiệu xung quanh.

Có lợi là, việc ước tính hướng theo sáng chế không bị ảnh hưởng bởi vấn đề này.

Các ứng dụng khác của quy trình phân tích dạng biểu diễn tín hiệu HOA

Quy trình phân tích được mô tả đối với dạng biểu diễn tín hiệu HOA thành một số tín hiệu có hướng với thông tin hướng liên quan và thành phần xung quanh trong miền tín hiệu HOA có thể được dùng cho quy trình kết xuất tín hiệu thích ứng dạng DirAC của dạng biểu diễn tín hiệu HOA theo quy trình được đề xuất trong tài liệu Pulkki "Spatial Sound Reproduction with Directional Audio Coding".

Mỗi thành phần tín hiệu HOA có thể được kết xuất khác nhau do các đặc tính vật lý của hai thành phần là khác nhau. Ví dụ, các tín hiệu có hướng có thể được kết xuất đến loa phóng thanh bằng cách sử dụng kỹ thuật quét tín hiệu tương tự như kỹ thuật quét biên độ dựa trên vector (Vector Based Amplitude Panning - VBAP), ví dụ, trong tài liệu của V. Pulkki, "Virtual Sound Source Positioning Using Vector Base Amplitude Panning", Journal of Audio Eng. Society, tập 45, số 6, trang 456-466, 1997. Thành phần tín hiệu HOA xung quanh có thể được kết xuất bằng cách sử dụng kỹ thuật kết xuất HOA chuẩn đã biết.

Quy trình kết xuất như vậy không bị hạn chế bởi dạng biểu diễn tín hiệu Ambisonic bậc '1' và do đó có thể được xem như phần mở rộng của quy trình kết xuất dạng DirAC cho dạng biểu diễn tín hiệu HOA bậc $N > 1$.

Quy trình ước tính một vài hướng từ dạng biểu diễn tín hiệu HOA có thể được dùng cho mọi loại phân tích trường âm thanh liên quan.

Các đoạn sau mô tả chi tiết hơn các bước xử lý tín hiệu.

Phương pháp nén

Định nghĩa định dạng đầu vào

Đối với đầu vào, hệ số HOA miền thời gian được định tỷ lệ $\tilde{c}_n^m(t)$ xác định trong phương trình (26) được giả định sẽ được lấy mẫu với tốc độ $f_s = \frac{1}{T_s}$. Vector $c(j)$ được xác định bao gồm tất cả hệ số thuộc thời gian lấy mẫu $t = jT_s$, $j \in \mathbb{Z}$, theo $c(j) := [\tilde{c}_0^0(jT_s), \tilde{c}_1^{-1}(jT_s), \tilde{c}_1^0(jT_s), \tilde{c}_1^1(jT_s), \tilde{c}_2^{-2}(jT_s), \dots, \tilde{c}_N^N(jT_s)]^T \in \mathbb{R}^O$. (65)

Tạo khung

Vector đến $c(j)$ có các hệ số HOA định tỷ lệ được tạo khung trong bước hoặc

giai đoạn tạo khung 21 thành các khung không chồng nhau có chiều dài B theo

$$C(l) := [c(lB + 1) \ c(lB + 2) \ \dots \ c(lB + B)] \in R^{O \times B}. \quad (66)$$

Giả sử tốc độ lấy mẫu $f_s = 48kHz$, chiều dài khung phù hợp là $B = 1200$ mẫu tương ứng với khoảng thời gian khung $25ms$.

Ước lượng các hướng chính

Đối với quy trình ước tính hướng chính, ma trận tương quan sau

$$B(l) := \frac{1}{LB} \sum_{l'=0}^{L-1} C(l-l') C^T(l-l') \in \mathbb{R}^{O \times O}. \quad (67)$$

được tính toán. Phép tổng trên khung hiện thời l và $L-l$ khung trước đó chỉ ra rằng sự phân tích hướng dựa vào nhóm các khung chồng dài có $L \cdot B$ mẫu, tức là, đối với mỗi khung hiện thời, nội dung của các khung liền kề được xem xét. Việc này đóng góp vào sự ổn định của quy trình phân tích hướng vì hai lý do: các khung dài hơn dẫn đến số lượng quan sát lớn hơn, và quy trình ước tính hướng được làm mịn do các khung chồng nhau.

Giả sử $f_s = 48kHz$ và $B = 1200$, giá trị hợp lý cho L là 4 tương ứng với toàn bộ khoảng thời gian khung $100ms$.

Tiếp theo, quy trình phân tích giá trị riêng của ma trận tương quan $B(l)$ được xác định theo $(l) = V(l) \Lambda(l) V^T(l)$, (68)

trong đó, ma trận $V(l)$ bao gồm các vector riêng $v_i(l)$, $1 \leq i \leq O$, là $V(l) := [v_1(l) \ v_2(l) \ \dots \ v_O(l)] \in R^{O \times O}$ (69)

và ma trận $\Lambda(l)$ là ma trận đường chéo có các giá trị riêng tương ứng $\lambda_i(l)$, $1 \leq i \leq O$, trên đường chéo của nó:

$$\Lambda(l) := \text{diag}(\lambda_1(l), \lambda_2(l), \dots, \lambda_O(l)) \in R^{O \times O}. \quad (70)$$

Giả sử rằng các giá trị riêng được lập chỉ mục theo bậc giảm dần, tức là $\lambda_1(l) \geq \lambda_2(l) \geq \dots \geq \lambda_O(l)$. (71)

Sau đó, tập chỉ mục $\{1, \dots, \tilde{I}(l)\}$ của các giá trị riêng chính được tính. Một khả năng để quản lý tập chỉ mục này là xác định tỷ lệ công suất có hướng trên công suất xung quanh tối thiểu mong muốn của dải rộng DAR_{MIN} và sau đó xác định $\tilde{I}(l)$ sao cho

$$10 \log_{10} \left(\frac{\lambda_i(l)}{\lambda_1(l)} \right) \geq -DAR_{MIN} \ \forall i \leq \tilde{I}(l) \text{ và } 10 \log_{10} \left(\frac{\lambda_i(l)}{\lambda_1(l)} \right) > -DAR_{MIN} \\ \text{với } i = \tilde{I}(l) + 1. \quad (72)$$

Sự lựa chọn hợp lý cho DAR_{MIN} là 15dB. Số lượng giá trị riêng chính còn bị hạn chế không lớn hơn D để tập trung vào không nhiều hơn D hướng chính. Quy trình này được thực hiện bằng cách thay thế tập chỉ mục $\{1, \dots, \tilde{I}(l)\}$ bằng $\{1, \dots, I(l)\}$, trong đó

$$J(l) := \max(\tilde{J}(l), D). \quad (73)$$

Tiếp theo, dạng xấp xỉ bậc $I(l)$ của $B(l)$ thu được bằng

$$\mathbf{B}_J(l) := \mathbf{V}_J(l) \mathbf{\Lambda}_J(l) \mathbf{V}_J^T(l), \text{ trong đó} \quad (74)$$

$$\mathbf{V}_J(l) := [\mathbf{v}_1(l) \quad \mathbf{v}_2(l) \quad \dots \quad \mathbf{v}_{J(l)}(l)] \in \mathbb{R}^{O \times J(l)}, \quad (75)$$

$$\mathbf{\Lambda}_J(l) := \text{diag}(\lambda_1(l), \lambda_2(l), \dots, \lambda_{J(l)}(l)) \in \mathbb{R}^{J(l) \times J(l)}. \quad (76)$$

Ma trận này phải bao gồm các đóng góp của các thành phần có hướng chính so với $B(l)$.

Sau đó, vector

$$\boldsymbol{\sigma}^2(l) := \text{diag}(\mathbf{\Xi}^T \mathbf{B}_J(l) \mathbf{\Xi}) \in \mathbb{R}^Q \quad (77)$$

$$= (\mathbf{S}_1^T \mathbf{B}_J(l) \mathbf{S}_1, \dots, \mathbf{S}_Q^T \mathbf{B}_J(l) \mathbf{S}_Q)^T \quad (78)$$

được tính toán, trong đó $\mathbf{\Xi}$ là ma trận chế độ liên quan đến một lượng lớn các hướng thử nghiệm được phân tán gần bằng nhau $\Omega_q := (\theta_q, \phi_q)$, $1 \leq q \leq Q$, trong đó $\theta_q \in [0, \pi]$ biểu thị góc nghiêng $\theta \in [0, \pi]$ được đo từ trục cực z và $\phi_q \in [-\pi, \pi[$ biểu thị góc phương vị được đo ở mặt phẳng $x=y$ từ trục x .

$$\text{Ma trận chế độ } \mathbf{E} \text{ được xác định bằng } \mathbf{E} := [S_1 \ S_2 \ \dots \ S_Q] \in \mathbb{R}^{O \times Q} \quad (79)$$

$$\text{với } S_q := [S_0^0(\Omega_q), S_1^{-1}(\Omega_q), S_1^0(\Omega_q), S_1^{-1}(\Omega_q), S_2^{-2}(\Omega_q), \dots, S_N^N(\Omega_q)]^T \quad (80)$$

với $1 \leq q \leq Q$.

Các phần tử $\sigma_q^2(l)$ của $\boldsymbol{\sigma}^2(l)$ là các giá trị xấp xỉ của công suất của các sóng phẳng, tương ứng với các tín hiệu có hướng chính, tác động tới các hướng Ω_q . Sự giải thích theo lý thuyết cho điều đó được đưa ra trong phần dưới *Giải thích thuật toán tìm kiếm hướng*.

Từ $\boldsymbol{\sigma}^2(l)$, số lượng $\tilde{D}(l)$ của các hướng chính $\boldsymbol{\Omega}_{\text{CURRDOM}, \tilde{d}}(l)$, $1 \leq \tilde{d} \leq \tilde{D}(l)$, cho quy trình xác định các thành phần tín hiệu có hướng được tính toán. Do đó số hướng chính được thu hẹp để đáp ứng $\tilde{D}(l) \leq D$ để đảm bảo tốc độ dữ liệu không đổi. Tuy nhiên, nếu tốc độ dữ liệu biến đổi được chấp nhận, thì số hướng chính có thể được thích ứng với cảnh âm thanh hiện tại.

Có một khả năng để tính toán hướng chính $\tilde{D}(l)$ là thiết lập hướng chính thứ nhất với công suất cực đại, tức là $\Omega_{CURRDOM,1}(l) = \Omega_{q_1}$ với $q_1 := \operatorname{argmax}_{q \in M_1} \sigma_q^2(l)$ và $M_1 := \{1, 2, \dots, Q\}$. Giả sử rằng giá trị công suất lớn nhất được tạo ra bởi tín hiệu có hướng chính, và xem xét thực tế là việc sử dụng dạng biểu diễn tín hiệu HOA bậc hữu hạn N dẫn đến sự phân tán các tín hiệu có hướng trong không gian (ví dụ, tài liệu "Plane-wave decomposition ..." nêu trên), nên có thể kết luận rằng trong miền lân cận có hướng $\Omega_{CURRDOM,1}(l)$ xuất hiện các thành phần công suất thuộc về cùng tín hiệu có hướng. Do sự phân tán tín hiệu không gian có thể được biểu diễn bằng hàm $v_N(\theta_{q,q_1})$ (xem phương trình (38)), trong đó $\theta_{q,q_1} := \angle(\Omega_q, \Omega_{q_1})$ biểu thị góc giữa Ω_q và $\Omega_{CURRDOM,1}(l)$, công suất thuộc về tín hiệu có hướng giảm theo $v_N^2(\theta_{q,q_1})$. Do đó, thích hợp để loại trừ tất cả các hướng Ω_q trong miền lân cận có hướng Ω_{q_1} với $\theta_{q,1} \leq \theta_{MIN}$ dùng để tìm kiếm hướng chính khác. Khoảng cách θ_{MIN} có thể được chọn là điểm gốc đầu tiên của $v_N(x)$, điểm gốc này được xác định là $\frac{\pi}{N}$ với $N \geq 4$. Sau đó hướng chính thứ hai được thiết lập để có công suất cực đại theo các hướng còn lại $\Omega_q \in M_2$ với $M_2 := \{q \in M_1 | \theta_{q,1} > \theta_{MIN}\}$. Các hướng chính còn lại được xác định theo cách tương tự.

Số lượng hướng chính $\tilde{D}(l)$ có thể được xác định bằng cách coi như giá trị lũy thừa $\sigma_{q_d}^2(l)$ được gán cho các hướng chính riêng Ω_{q_d} và tìm kiếm trường hợp trong đó tỷ lệ $\frac{\sigma_{q_1}^2(l)}{\sigma_{q_d}^2(l)}$ vượt quá giá trị của tỷ lệ công suất có hướng trên công suất xung quanh mong muốn DAR_{MIN} . Nghĩa là $\tilde{D}(l)$ thỏa mãn

$$10 \log_{10} \left(\frac{\sigma_{q_1}^2(l)}{\sigma_{q_{\tilde{D}(l)}}^2(l)} \right) \leq DAR_{MIN} \wedge \left[10 \log_{10} \left(\frac{\sigma_{q_1}^2(l)}{\sigma_{q_{\tilde{D}(l)+1}}^2(l)} \right) > DAR_{MIN} \vee \tilde{D}(l) = D \right]. \quad (81)$$

Toàn bộ quy trình để tính toán tất cả hướng chính có thể được thực hiện như sau:

Thuật toán 1: Tìm kiếm các hướng chính xác định sự phân phối công suất trên hình cầu

```

PowerFlag = true
 $\tilde{d} = 1$ 
 $\mathcal{M}_1 = \{1, 2, \dots, Q\}$ 
repeat
   $q_{\tilde{d}} = \operatorname{argmax}_{q \in \mathcal{M}_{\tilde{d}}} \sigma_q^2(l)$ 
  if  $\left[ \tilde{d} > 1 \wedge 10 \log_{10} \left( \frac{\sigma_{q_1}^2(l)}{\sigma_{q_{\tilde{d}}}^2(l)} \right) > \text{DAR}_{\text{MIN}} \right]$  then
    PowerFlag = false
  else
     $\Omega_{\text{CURRDOM}, \tilde{d}}(l) = \Omega_{q_{\tilde{d}}}$ 
     $\mathcal{M}_{\tilde{d}+1} = \{q \in \mathcal{M}_{\tilde{d}} \mid \angle(\Omega_q, \Omega_{q_{\tilde{d}}}) > \Theta_{\text{MIN}}\}$ 
     $\tilde{d} = \tilde{d} + 1$ 
  end if
until  $\left[ \tilde{d} > D \vee \text{PowerFlag} = \text{false} \right]$ 
 $\tilde{D}(l) = \tilde{d} - 1$ 

```

Tiếp theo, các hướng $\Omega_{\text{CURRDOM}, \tilde{d}}(l)$, $1 \leq \tilde{d} \leq \tilde{D}(l)$ thu được ở khung hiện thời được làm mịn với các hướng của các khung trước đó, tạo ra các hướng được làm mịn $\bar{\Omega}_{\text{DOM}, d}(l)$, $1 \leq d \leq D$. Thao tác này có thể được chia thành hai phần liên tục:

- (a) Hướng chính hiện thời $\Omega_{\text{CURRDOM}, \tilde{d}}(l)$, $1 \leq \tilde{d} \leq \tilde{D}(l)$, được gán cho các hướng được làm mịn $\bar{\Omega}_{\text{DOM}, d}(l-1)$, $1 \leq d \leq D$, từ khung trước đó. Hàm gán $f_{A, l}: \{1, \dots, \tilde{D}(l)\} \rightarrow \{1, \dots, D\}$ được xác định sao cho tổng các góc giữa các hướng đã gán

$$\sum_{\tilde{d}=1}^{\tilde{D}(l)} \angle \left(\Omega_{\text{CURRDOM}, \tilde{d}}(l), \bar{\Omega}_{\text{DOM}, f_{A, l}(\tilde{d})}(l-1) \right) \quad (82)$$

được tối thiểu hóa. Vấn đề gán như vậy có thể được giải quyết bằng cách sử dụng thuật toán Hungary đã biết, chẳng hạn, H.W. Kuhn, "The Hungarian method for the assignment problem", nghiên cứu logistic Naval hàng quý 2, số 1-2, từ trang 83 đến 97, 1955. Các góc giữa các hướng hiện thời $\Omega_{\text{CURRDOM}, \tilde{d}}(l)$ và các hướng không hoạt động (xem dưới đây để giải thích thuật ngữ 'hướng không hoạt động') từ khung trước đó $\bar{\Omega}_{\text{DOM}, d}(l-1)$ được thiết lập là $2\Theta_{\text{MIN}}$. Thao tác này có hiệu quả là các hướng hiện thời $\Omega_{\text{CURRDOM}, \tilde{d}}(l)$, các hướng này có khoảng cách gần hơn $2\Theta_{\text{MIN}}$ so với các hướng hoạt động nêu trên $\bar{\Omega}_{\text{DOM}, d}(l-1)$, được cố gắng để gán cho chúng. Nếu khoảng cách vượt quá $2\Theta_{\text{MIN}}$, thì hướng hiện thời tương ứng được gán định thuộc tín hiệu mới, có nghĩa là nó

được ưu tiên gán cho một hướng không hoạt động trước đây $\bar{\Omega}_{\text{DOM},d}(l-1)$.

Chú ý: khi cho phép thuật toán nén toán bộ có độ trễ lớn hơn, việc gán các giá trị ước tính hướng liên tục có thể được thực hiện chính xác hơn. Ví dụ, thay đổi hướng đột ngột có thể được xác định tốt hơn mà không trộn chúng với các giá trị ngoại lai do sai số dự đoán.

- (b) Các hướng được làm mịn $\bar{\Omega}_{\text{DOM},d}(l-1)$, $1 \leq d \leq D$ được tính bằng cách sử dụng giá trị gán từ bước (a). Việc làm mịn dựa trên hình cầu học chứ không phải là hình học 3clit. Đối với mỗi hướng chính hiện thời $\Omega_{\text{CURRDOM},\tilde{d}}(l)$, $1 \leq \tilde{d} \leq \tilde{D}(l)$, quy trình làm mịn được thực hiện dọc theo hình cung nhỏ của vòng tròn lớn đi qua hai điểm trên mặt cầu, hai điểm này được xác định theo các hướng $\Omega_{\text{CURRDOM},\tilde{d}}(l)$ và $\bar{\Omega}_{\text{DOM},d}(l-1)$. Rõ ràng là, các góc phương vị và góc nghiêng được làm mịn độc lập bằng cách tính trung bình động được lấy trọng số theo hàm mũ nhân với hệ số làm mịn α_Ω . Đối với góc nghiêng việc này dẫn đến phép toán làm mịn sau đây:

$$\begin{aligned} \bar{\theta}_{\text{DOM},f_{\mathcal{A},l}(\tilde{d})}(l) &= (1 - \alpha_\Omega) \cdot \bar{\theta}_{\text{DOM},f_{\mathcal{A},l}(\tilde{d})}(l-1) + \alpha_\Omega \cdot \theta_{\text{DOM},\tilde{d}}(l) , \\ 1 \leq \tilde{d} \leq \tilde{D}(l) . \end{aligned} \quad (83)$$

Đối với góc phương vị, quy trình làm mịn phải được sửa đổi để thu được độ mịn chính xác ở quy trình chuyển tiếp từ $\pi - \varepsilon$ sang $-\pi$, $\varepsilon > 0$, và quy trình chuyển tiếp theo hướng ngược. Quy trình này có thể được xem xét bằng cách tính toán trước tiên môđun góc sai phân 2π là

$$\Delta_{\phi,[0,2\pi],\tilde{d}}(l) := \left[\phi_{\text{DOM},\tilde{d}}(l) - \bar{\phi}_{\text{DOM},f_{\mathcal{A},l}(\tilde{d})}(l-1) \right] \bmod 2\pi , \quad (84)$$

mà được biến đổi sang khoảng $[-\pi, \pi[$ bằng

$$\Delta_{\phi,[-\pi,\pi],\tilde{d}}(l) := \begin{cases} \Delta_{\phi,[0,2\pi],\tilde{d}}(l) & \text{với } \Delta_{\phi,[0,2\pi],\tilde{d}}(l) < \pi \\ \Delta_{\phi,[0,2\pi],\tilde{d}}(l) - 2\pi & \text{với } \Delta_{\phi,[0,2\pi],\tilde{d}}(l) \geq \pi \end{cases} . \quad (85)$$

Môđun góc phương vị chính được làm mịn 2π được xác định là

$$\bar{\phi}_{\text{DOM},[0,2\pi],\tilde{d}}(l) := \left[\bar{\phi}_{\text{DOM},\tilde{d}}(l-1) + \alpha_\Omega \cdot \Delta_{\phi,[-\pi,\pi],\tilde{d}}(l) \right] \bmod 2\pi \quad (86)$$

và cuối cùng được biến đổi nằm trong khoảng $[-\pi, \pi[$ bằng

$$\bar{\phi}_{\text{DOM},\tilde{d}}(l) = \begin{cases} \bar{\phi}_{\text{DOM},[0,2\pi],\tilde{d}}(l) & \text{với } \bar{\phi}_{\text{DOM},[0,2\pi],\tilde{d}}(l) < \pi \\ \bar{\phi}_{\text{DOM},[0,2\pi],\tilde{d}}(l) - 2\pi & \text{với } \bar{\phi}_{\text{DOM},[0,2\pi],\tilde{d}}(l) \geq \pi \end{cases} . \quad (87)$$

Trong trường hợp $\tilde{D}(l) < D$, có các hướng $\bar{\Omega}_{\text{DOM},d}(l-1)$ từ khung trước đó mà

không nhận được hướng chính hiện thời được gán. Tập hợp chỉ số tương ứng được biểu thị bởi

$$\mathcal{M}_{NA}(l) := \{1, \dots, D\} \setminus \{f_{\mathcal{A},l}(\tilde{d}) | 1 \leq \tilde{d} \leq D\}. \quad (88)$$

Các hướng tương ứng được sao chép từ khung cuối cùng, tức là $\bar{\boldsymbol{\Omega}}_{DOM,d}(l) = \bar{\boldsymbol{\Omega}}_{DOM,d}(l-1)$ với $l \in M_{NA}(l)$. (89)

Các hướng mà không được gán số khung định trước L_{IA} được gọi là không hoạt động.

Sau đó tập chỉ mục gồm các hướng hoạt động biểu thị là $M_{ACT}(l)$ được tính toán. Bản số của nó được biểu thị bằng $D_{ACT}(l) := |M_{ACT}(l)|$.

Sau đó, tất cả các hướng được làm mịn được ghép vào ma trận đơn hướng là

$$\bar{\boldsymbol{\Omega}}_{DOM}(l) := [\bar{\boldsymbol{\Omega}}_{DOM,1}(l) \quad \bar{\boldsymbol{\Omega}}_{DOM,2}(l) \quad \dots \quad \bar{\boldsymbol{\Omega}}_{DOM,D}(l)]. \quad (90)$$

Tính toán tín hiệu có hướng

Việc tính toán tín hiệu có hướng dựa trên sự so khớp chế độ. Cụ thể, việc tìm kiếm được thực hiện đối với các tín hiệu có hướng mà dạng biểu diễn tín hiệu HOA của nó tạo ra giá trị xấp xỉ tốt nhất cho tín hiệu HOA đã cho. Do sự thay đổi các hướng giữa các khung liên tiếp có thể dẫn đến sự gián đoạn của các tín hiệu có hướng, nên việc ước tính các tín hiệu có hướng cho các khung chồng nhau có thể được tính toán, tiếp theo là làm mịn kết quả của các khung chồng nhau liên tục bằng cách sử dụng hàm cửa sổ phù hợp. Tuy nhiên, việc làm mịn tạo ra độ trễ của khung đơn.

Việc ước tính chi tiết của các tín hiệu có hướng được giải thích như sau:

Thứ nhất, ma trận chế độ dựa trên các hướng hoạt động được làm mịn được tính theo

$$\mathbf{E}_{ACT}(l) := [\mathbf{S}_{DOM,d_{ACT,1}}(l) \quad \mathbf{S}_{DOM,d_{ACT,2}}(l) \quad \dots \quad \mathbf{S}_{DOM,d_{ACT,D_{ACT}(l)}}(l)] \in \mathbb{R}^{O \times D_{ACT}(l)} \quad (91)$$

với $S_{DOM,d}(l) :=$

$$[S_0^0(\bar{\boldsymbol{\Omega}}_{DOM,d}(l)), S_1^{-1}(\bar{\boldsymbol{\Omega}}_{DOM,d}(l)), S_1^0(\bar{\boldsymbol{\Omega}}_{DOM,d}(l)), \dots, S_N^N(\bar{\boldsymbol{\Omega}}_{DOM,d}(l))]^T \in \mathbb{R}^O, \quad (92)$$

trong đó $d_{ACT,j}$, $1 \leq j \leq D_{ACT}(l)$ biểu thị các chỉ số của các hướng hoạt động.

Tiếp theo, ma trận $X_{INST}(l)$ được tính toán để chứa các ước tính không mịn của tất cả các tín hiệu có hướng cho khung thứ $(l-1)$ và thứ l :

$$\mathbf{X}_{\text{INST}}(l) := [\mathbf{x}_{\text{INST}}(l, 1) \quad \mathbf{x}_{\text{INST}}(l, 2) \quad \dots \quad \mathbf{x}_{\text{INST}}(l, 2B)] \in \mathbb{R}^{D \times 2B} \quad (93)$$

với

$$\mathbf{x}_{\text{INST}}(l, j) = [x_{\text{INST},1}(l, j), x_{\text{INST},2}(l, j), \dots, x_{\text{INST},D}(l, j)]^T \in \mathbb{R}^D, 1 \leq j \leq 2B. \quad (94)$$

Quy trình này được hoàn thành trong hai bước. Ở bước thứ nhất, các mẫu tín hiệu có hướng ở các hàng tương ứng với các hướng không hoạt động được thiết lập bằng 0, tức là

$$x_{\text{INST},d}(l, j) = 0 \quad \forall 1 \leq j \leq 2B, \quad \text{nếu } d \notin \mathcal{M}_{\text{ACT}}(l). \quad (95)$$

Ở bước thứ hai, các mẫu tín hiệu có hướng tương ứng với các hướng hoạt động thu được bằng cách đầu tiên sắp xếp chúng vào ma trận theo

$$\mathbf{X}_{\text{INST,ACT}}(l) := \begin{bmatrix} x_{\text{INST},d_{\text{ACT},1}}(l, 1) & \dots & x_{\text{INST},d_{\text{ACT},1}}(l, 2B) \\ \vdots & \ddots & \vdots \\ x_{\text{INST},d_{\text{ACT},D_{\text{ACT}}(l)}}(l, 1) & \dots & x_{\text{INST},d_{\text{ACT},D_{\text{ACT}}(l)}}(l, 2B) \end{bmatrix}. \quad (96)$$

Ma trận này sau đó được tính để tối thiểu hóa chuẩn Oclit có sai số $\mathbf{E}_{\text{ACT}}(l) \mathbf{X}_{\text{INST,ACT}}(l) - [\mathbf{C}(l-1) \quad \mathbf{C}(l)]$. (97)

Giải pháp được xác định bởi

$$\mathbf{X}_{\text{INST,ACT}}(l) = [\mathbf{E}_{\text{ACT}}^T(l) \mathbf{E}_{\text{ACT}}(l)]^{-1} \mathbf{E}_{\text{ACT}}^T(l) [\mathbf{C}(l-1) \quad \mathbf{C}(l)]. \quad (98)$$

Các giá trị ước tính của tín hiệu có hướng $x_{\text{INST},d}(l, j)$, $1 \leq d \leq D$, được tạo cửa sổ bằng hàm cửa sổ phù hợp $w(j)$:

$$x_{\text{INST,WIN},d}(l, j) := x_{\text{INST},d}(l, j) \cdot w(j), 1 \leq j \leq 2B. \quad (99)$$

Ví dụ về hàm cửa sổ được tạo ra bởi cửa sổ Hamming có chu kỳ được xác định bằng

$$w(j) := \begin{cases} K_w \left[0.54 - 0.46 \cos\left(\frac{2\pi j}{2B+1}\right) \right] & \text{với } 1 \leq j \leq 2B \\ 0 & \text{hoặc} \end{cases}, \quad (100)$$

trong đó K_w là hệ số tỷ lệ mà được xác định sao cho tổng của các cửa sổ dịch chuyển bằng '1'. Các tín hiệu có hướng được làm mịn cho khung thứ $(l-1)$ được tính bằng cách xếp chồng thích hợp các giá trị ước tính không được làm mịn được tạo cửa sổ theo

$$x_d((l-1)B + j) = x_{\text{INST,WIN},d}(l-1, B + j) + x_{\text{INST,WIN},d}(l, j). \quad (101)$$

Các mẫu của tất cả tín hiệu có hướng được làm mịn cho khung thứ $(l-1)$ được sắp xếp trong ma trận $\mathbf{X}(l-1)$ là

$$\mathbf{X}(l-1) := [\mathbf{x}((l-1)B + 1) \quad \mathbf{x}((l-1)B + 2) \quad \dots \quad \mathbf{x}((l-1)B + B)] \in \mathbb{R}^{D \times B}$$

(102)

với $(j) = [x_1(j), x_2(j), \dots, x_D(j)]^T \in \mathbb{R}^D$. (103)

Tính toán thành phần HOA xung quanh

Thành phần HOA xung quanh $C_A(l-1)$ thu được bằng cách lấy dạng biểu diễn tín hiệu HOA tổng $C(l-1)$ trừ đi thành phần HOA có hướng tổng $C_{DIR}(l-1)$ theo

$$\mathbf{C}_A(l-1) := \mathbf{C}(l-1) - \mathbf{C}_{DIR}(l-1) \in \mathbb{R}^{O \times B}, \quad (104)$$

trong đó $C_{DIR}(l-1)$ được xác định bởi

$$\begin{aligned} \mathbf{C}_{DIR}(l-1) &:= \mathbf{\Xi}_{DOM}(l-1) \begin{bmatrix} x_{INST,WIN,1}(l-1, B+1) & & x_{INST,WIN,1}(l-1, 2B) \\ \vdots & \ddots & \vdots \\ x_{INST,WIN,D}(l-1, B+1) & & x_{INST,WIN,D}(l-1, 2B) \end{bmatrix} \\ &+ \mathbf{\Xi}_{DOM}(l) \begin{bmatrix} x_{INST,WIN,1}(l, 1) & & x_{INST,WIN,1}(l, B) \\ \vdots & \ddots & \vdots \\ x_{INST,WIN,D}(l, 1) & & x_{INST,WIN,D}(l, B) \end{bmatrix}, \quad (105) \end{aligned}$$

và trong đó $\mathbf{\Xi}_{DOM}(l)$ biểu thị ma trận chế độ dựa trên tất cả các hướng làm mịn được xác định bằng

$$\mathbf{\Xi}_{DOM}(l) := [\mathbf{S}_{DOM,1}(l) \quad \mathbf{S}_{DOM,2}(l) \quad \dots \quad \mathbf{S}_{DOM,D}(l)] \in \mathbb{R}^{O \times D}. \quad (106)$$

Do việc tính toán thành phần HOA có hướng tổng cũng dựa trên việc làm mịn không gian các thành phần HOA có hướng tổng tức thời chồng liên tục, nên thành phần HOA xung quanh cũng thu được có độ trễ của một khung đơn.

Giảm bậc cho thành phần HOA xung quanh

Thể hiện $C_A(l-1)$ qua các thành phần của nó là

$$\mathbf{C}_A(l-1) = \begin{bmatrix} c_{0,A}^0((l-1)B+1) & & c_{0,A}^0((l-1)B+B) \\ \vdots & \ddots & \vdots \\ c_{N,A}^N((l-1)B+1) & & c_{N,A}^N((l-1)B+B) \end{bmatrix}, \quad (107)$$

việc giảm bậc được thực hiện bằng cách giảm tất cả hệ số HOA $c_{n,A}^m(j)$ có $n > N_{RED}$:

$$\mathbf{C}_{A,RED}(l-1) := \begin{bmatrix} c_{0,A}^0((l-1)B+1) & & c_{0,A}^0((l-1)B+B) \\ \vdots & \ddots & \vdots \\ c_{N_{RED},A}^{N_{RED}}((l-1)B+1) & & c_{N_{RED},A}^{N_{RED}}((l-1)B+B) \end{bmatrix} \in \mathbb{R}^{O_{RED} \times B}. \quad (108)$$

Biến đổi hàm điều hòa cầu cho thành phần HOA xung quanh

Phép biến đổi hàm điều hòa cầu được thực hiện bằng cách nhân thành phần HOA xung quanh có bậc đã giảm $C_{A,RED}(l)$ với nghịch đảo của ma trận chế độ

$$\mathbf{\Xi}_A := [\mathbf{S}_{A,1} \quad \mathbf{S}_{A,2} \quad \dots \quad \mathbf{S}_{A,O_{RED}}] \in \mathbb{R}^{O_{RED} \times O_{RED}} \quad (109)$$

với $\mathbf{S}_{A,d} := [S_0^0(\boldsymbol{\Omega}_{A,d}), S_1^{-1}(\boldsymbol{\Omega}_{A,d}), S_1^0(\boldsymbol{\Omega}_{A,d}), \dots, S_{N_{RED}}^{N_{RED}}(\boldsymbol{\Omega}_{A,d})]^T \in \mathbb{R}^{O_{RED}}$, (110)

dựa trên O_{RED} là các hướng phân tán đều $\boldsymbol{\Omega}_{A,d}$, $1 \leq d \leq O_{RED}$: $\mathbf{W}_{A,RED}(l) = (\mathbf{E}_A)^{-1} \mathbf{C}_{A,RED}(l)$. (111)

Giải nén

Biến đổi hàm điều hòa cầu ngược

Các tín hiệu miền không gian được giải nén cảm quan $\widehat{\mathbf{W}}_{A,RED}(l)$ được biến đổi sang dạng biểu diễn miền HOA $\widehat{\mathbf{C}}_{A,RED}(l)$ bậc N_{RED} qua phép biến đổi hàm điều hòa cầu ngược bằng

$$\widehat{\mathbf{C}}_{A,RED}(l) = \mathbf{E}_A \widehat{\mathbf{W}}_{A,RED}(l). \quad (112)$$

Mở rộng bậc

Bậc ambisonic của dạng biểu diễn tín hiệu HOA $\widehat{\mathbf{C}}_{A,RED}(l)$ được mở rộng đến N bằng cách thêm các số không theo

$$\widehat{\mathbf{C}}_A(l) := \begin{bmatrix} \widehat{\mathbf{C}}_{A,RED}(l) \\ \mathbf{0}_{(O-O_{RED}) \times B} \end{bmatrix} \in \mathbb{R}^{O \times B}, \quad (113)$$

trong đó $\mathbf{0}_{m \times n}$ là ma trận số không có m hàng và n cột.

Thành phần hệ số HOA

Các hệ số HOA được giải nén cuối cùng bao gồm thêm thành phần HOA có hướng và xung quanh theo $\widehat{\mathbf{C}}(l-1) := \widehat{\mathbf{C}}_A(l-1) + \widehat{\mathbf{C}}_{DIR}(l-1)$. (114)

Ở giai đoạn này, độ trễ của khung đơn lại được đưa vào để cho phép thành phần HOA có hướng được tính toán dựa trên sự làm mịn không gian. Bằng cách làm như vậy, tránh được sự gián đoạn không mong muốn có thể có trong thành phần có hướng của trường âm thanh do sự thay đổi của các hướng giữa các khung liên tiếp.

Để tính toán thành phần HOA có hướng được làm mịn, hai khung liên tục có chứa các giá trị ước tính của tất cả các tín hiệu có hướng riêng được ghép thành khung dài đơn là $\widehat{\mathbf{X}}_{INST}(l) := [\widehat{\mathbf{X}}(l-1) \quad \widehat{\mathbf{X}}(l)] \in \mathbb{R}^{D \times 2B}$. (115)

Mỗi phần trích tín hiệu riêng chứa trong khung dài này được nhân với hàm cửa sổ, chẳng hạn, giống như phương trình (100). Khi biểu diễn khung dài $\widehat{\mathbf{X}}_{INST}(l)$ qua các thành phần của nó bằng

$$\widehat{\mathbf{X}}_{INST}(l) = \begin{bmatrix} \widehat{x}_{INST,1}(l, 1) & & \widehat{x}_{INST,1}(l, 2B) \\ \vdots & \ddots & \vdots \\ \widehat{x}_{INST,D}(l, 1) & & \widehat{x}_{INST,D}(l, 2B) \end{bmatrix}, \quad (116)$$

thao tác tạo cửa sổ có thể được thực hiện như tính toán các đoạn trích tín hiệu được tạo cửa sổ $\hat{x}_{INST,WIN,d}(l,j)$, $1 \leq d \leq D$, bằng

$$\hat{x}_{INST,WIN,d}(l,j) = \hat{x}_{INST,d}(l,j) \cdot w(j), 1 \leq j \leq 2B, 1 \leq d \leq D. \quad (117)$$

Cuối cùng, thành phần HOA có hướng tổng $C_{DIR}(l-1)$ thu được bằng cách mã hóa tất cả đoạn trích tín hiệu có hướng được tạo cửa sổ thành các hướng phù hợp và chồng các đoạn trích này theo kiểu gối lên nhau:

$$\begin{aligned} \hat{C}_{DIR}(l-1) = \Xi_{DOM}(l-1) & \begin{bmatrix} \hat{x}_{INST,WIN,1}(l-1, B+1) & & \hat{x}_{INST,WIN,1}(l-1, 2B) \\ \vdots & \ddots & \vdots \\ \hat{x}_{INST,WIN,D}(l-1, B+1) & & \hat{x}_{INST,WIN,D}(l-1, 2B) \end{bmatrix} \\ + \Xi_{DOM}(l) & \begin{bmatrix} \hat{x}_{INST,WIN,1}(l, 1) & & \hat{x}_{INST,WIN,1}(l, B) \\ \vdots & \ddots & \vdots \\ \hat{x}_{INST,WIN,D}(l, 1) & & \hat{x}_{INST,WIN,D}(l, B) \end{bmatrix}. \end{aligned} \quad (118)$$

Giải thích thuật toán tìm kiếm hướng

Trong phần sau đây, quá trình thúc đẩy được giải thích sau quy trình tìm kiếm hướng được mô tả trong phần *Ước tính các hướng chính*. Quá trình này dựa trên một số giả thiết mà ban đầu được xác định.

Các giả thiết

Nói chung vector hệ số HOA $c(j)$ liên quan đến hàm mật độ biên miền thời gian $d(j, \Omega)$ qua phương trình

$$c(j) = \int_{S^2} d(j, \Omega) \mathbf{S}(\Omega) d\Omega, \quad (119)$$

được giả định tuân theo mẫu sau:

$$c(j) = \sum_{i=1}^I x_i(j) \mathbf{S}(\Omega_{x_i}(l)) + c_A(j) \quad \text{với } lB+1 \leq j \leq (l+1)B. \quad (120)$$

Mô hình này quy định rằng vector hệ số HOA $c(j)$ một mặt được tạo ra bằng I tín hiệu nguồn có hướng chính $x_i(j)$, $1 \leq i \leq I$, đến từ các hướng $\Omega_{x_i}(l)$ trong khung thứ l . Cụ thể là, các hướng được giả định là cố định trong khoảng thời gian của khung đơn. Số tín hiệu nguồn chính I được giả định rõ ràng nhỏ hơn tổng số hệ số HOA O . Ngoài ra, chiều dài khung B được giả định rõ ràng lớn hơn O . Mặt khác, vector $c(j)$ bao gồm thành phần dư $c_A(j)$, thành phần này có thể được coi như biểu diễn trường âm thanh xung quanh đẳng hướng lý tưởng.

Các thành phần vector hệ số HOA riêng được giả định có các tính chất sau:

- Các tín hiệu nguồn chính được giả định là trung bình bằng không, tức là

$$\sum_{j=lB+1}^{(l+1)B} x_i(j) \approx 0 \quad \forall 1 \leq i \leq I, \quad (121)$$

- Và giả sử là không liên quan với nhau, tức là

$$\frac{1}{B} \sum_{j=lB+1}^{(l+1)B} x_i(j) x_{i'}(j) \approx \delta_{i-i'} \bar{\sigma}_{x_i}^2(l) \quad \forall 1 \leq i, i' \leq I \quad (122)$$

với $\bar{\sigma}_{x_i}^2(l)$ là công suất trung bình của tín hiệu thứ i cho khung thứ l .

- Các tín hiệu nguồn chính được giả định là không liên quan với thành phần xung quanh của vector hệ số HOA, tức là

$$\frac{1}{B} \sum_{j=lB+1}^{(l+1)B} x_i(j) \mathbf{c}_A(j) \approx \mathbf{0} \quad \forall 1 \leq i \leq I. \quad (123)$$

- Vector thành phần HOA xung quanh được giả định là trung bình bằng không và được giả định có ma trận hiệp biến

$$\Sigma_A(l) := \frac{1}{B} \sum_{j=lB+1}^{(l+1)B} \mathbf{c}_A(j) \mathbf{c}_A^T(j). \quad (124)$$

- Tỷ lệ công suất có hướng trên công suất xung quanh $DAR(l)$ của mỗi khung l được

$$\text{xác định ở đây bằng } DAR(l) := 10 \log_{10} \left[\frac{\max_{1 \leq i \leq I} \bar{\sigma}_{x_i}^2(l)}{\|\Sigma_A(l)\|^2} \right], \quad (125)$$

được giả định lớn hơn giá trị được mong muốn định trước DAR_{MIN} , tức là

$$DAR(l) \geq DAR_{MIN}. \quad (126)$$

Giải thích sự tìm kiếm hướng

Để giải thích, trường hợp dưới đây được xem xét trong đó ma trận tương quan $B(l)$ (xem phương trình (67)) được tính toán chỉ dựa trên các mẫu của khung thứ l mà không xem xét các mẫu của $L - 1$ khung trước đó. Thao tác này tương ứng với hoạt động thiết lập $L = 1$. Do đó, ma trận tương quan có thể được biểu diễn bằng

$$\mathbf{B}(l) = \frac{1}{B} \mathbf{C}(l) \mathbf{C}^T(l) \quad (127)$$

$$= \frac{1}{B} \sum_{j=lB+1}^{(l+1)B} \mathbf{c}(j) \mathbf{c}^T(j). \quad (128)$$

Bằng việc thế giả thiết mô hình trong phương trình (120) vào phương trình (128) và sử dụng các phương trình (122) và (123) và định nghĩa trong phương trình (124), ma trận tương quan $B(l)$ có thể được tính xấp xỉ bằng

$$\begin{aligned} \mathbf{B}(l) &= \frac{1}{B} \sum_{j=lB+1}^{(l+1)B} [\sum_{i=1}^I x_i(j) \mathbf{S}(\boldsymbol{\Omega}_{x_i}(l)) + \mathbf{c}_A(j)] [\sum_{i'=1}^I x_{i'}(j) \mathbf{S}(\boldsymbol{\Omega}_{x_{i'}}(l)) + \mathbf{c}_A(j)]^T \\ &= \sum_{i=1}^I \sum_{i'=1}^I \mathbf{S}(\boldsymbol{\Omega}_{x_i}(l)) \mathbf{S}^T(\boldsymbol{\Omega}_{x_{i'}}(l)) \frac{1}{B} \sum_{j=lB+1}^{(l+1)B} x_i(j) x_{i'}(j) \\ &+ \sum_{i=1}^I \mathbf{S}(\boldsymbol{\Omega}_{x_i}(l)) \frac{1}{B} \sum_{j=lB+1}^{(l+1)B} x_i(j) \mathbf{c}_A^T(j) + \sum_{i'=1}^I \frac{1}{B} \sum_{j=lB+1}^{(l+1)B} x_{i'}(j) \mathbf{c}_A(j) \mathbf{S}^T(\boldsymbol{\Omega}_{x_{i'}}(l)) \\ &+ \frac{1}{B} \sum_{j=lB+1}^{(l+1)B} \mathbf{c}_A(j) \mathbf{c}_A^T(j) \end{aligned} \quad (130)$$

$$\approx \sum_{i=1}^I \bar{\sigma}_{x_i}^2(l) \mathbf{S}(\boldsymbol{\Omega}_{x_i}(l)) \mathbf{S}^T(\boldsymbol{\Omega}_{x_i}(l)) + \Sigma_A(l). \quad (131)$$

Dựa vào phương trình (131) có thể thấy rằng $B(l)$ bao gồm hai thành phần bổ sung có thể quy về thành phần HOA có hướng và thành phần HOA xung quanh. Giá trị xấp xỉ hạng $\mathcal{I}(l)$ $B_I(l)$ của nó tạo ra giá trị xấp xỉ của thành phần HOA có hướng, tức là $B_I(l) \approx \sum_{i=1}^I \bar{\sigma}_{x_i}^2(l) \mathbf{S}(\boldsymbol{\Omega}_{x_i}(l)) \mathbf{S}^T(\boldsymbol{\Omega}_{x_i}(l))$, (132)

mà tiếp theo phương trình (126) dựa trên tỷ lệ công suất có hướng trên công suất xung quanh.

Tuy nhiên, cần phải nhấn mạnh rằng một số phần của $\Sigma_A(l)$ chắc chắn sẽ lọt vào $B_I(l)$, vì $\Sigma_A(l)$ có đầy đủ bậc và do đó, không gian con được mở rộng bởi các cột của các ma trận $\sum_{i=1}^I \bar{\sigma}_{x_i}^2(l) \mathbf{S}(\boldsymbol{\Omega}_{x_i}(l)) \mathbf{S}^T(\boldsymbol{\Omega}_{x_i}(l))$ và $\Sigma_A(l)$ không vuông góc với nhau. Với phương trình (132), vector $\sigma^2(l)$ trong phương trình (77) dùng để tìm kiếm các hướng chính có thể được biểu diễn bằng $\sigma^2(l) = \text{diag}(\mathbf{E}^T B_I(l) \mathbf{E})$ (133)

$$\begin{aligned}
 &= \text{diag} \left(\begin{bmatrix} \mathbf{S}^T & (\boldsymbol{\Omega}_1) \mathbf{B}_J(l) \mathbf{S}(\boldsymbol{\Omega}_1) & & \mathbf{S}^T & (\boldsymbol{\Omega}_1) \mathbf{B}_J(l) \mathbf{S}(\boldsymbol{\Omega}_Q) \\ \vdots & & \ddots & \vdots & \\ \mathbf{S}^T & (\boldsymbol{\Omega}_Q) \mathbf{B}_J(l) \mathbf{S}(\boldsymbol{\Omega}_1) & & \mathbf{S}^T & (\boldsymbol{\Omega}_Q) \mathbf{B}_J(l) \mathbf{S}(\boldsymbol{\Omega}_Q) \end{bmatrix} \right) \approx \quad (134) \\
 &\text{diag} \left(\begin{bmatrix} \sum_{i=1}^I \bar{\sigma}_{x_i}^2(l) v_N^2(\angle(\boldsymbol{\Omega}_1, \boldsymbol{\Omega}_{x_i})) & & \sum_{i=1}^I \bar{\sigma}_{x_i}^2(l) v_N(\angle(\boldsymbol{\Omega}_1, \boldsymbol{\Omega}_{x_i})) v_N(\angle(\boldsymbol{\Omega}_{x_i}, \boldsymbol{\Omega}_Q)) \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^I \bar{\sigma}_{x_i}^2(l) v_N(\angle(\boldsymbol{\Omega}_Q, \boldsymbol{\Omega}_{x_i})) v_N(\angle(\boldsymbol{\Omega}_{x_i}, \boldsymbol{\Omega}_1)) & & \sum_{i=1}^I \bar{\sigma}_{x_i}^2(l) v_N^2(\angle(\boldsymbol{\Omega}_Q, \boldsymbol{\Omega}_{x_i})) \end{bmatrix} \right) \\
 &= \left[\sum_{i=1}^I \bar{\sigma}_{x_i}^2(l) v_N^2(\angle(\boldsymbol{\Omega}_1, \boldsymbol{\Omega}_{x_i})) \quad \dots \quad \sum_{i=1}^I \bar{\sigma}_{x_i}^2(l) v_N^2(\angle(\boldsymbol{\Omega}_Q, \boldsymbol{\Omega}_{x_i})) \right]^T. \quad (136)
 \end{aligned}$$

Trong phương trình (135), tính chất sau đây của hàm điều hòa cầu thể hiện trong phương trình (47) được sử dụng: $\mathbf{S}^T(\boldsymbol{\Omega}_q) \mathbf{S}(\boldsymbol{\Omega}_{q'}) = v_N(\angle(\boldsymbol{\Omega}_q, \boldsymbol{\Omega}_{q'}))$. (137)

Phương trình (136) thể hiện rằng các thành phần $\sigma_q^2(l)$ của $\sigma^2(l)$ là các giá trị xấp xỉ của công suất tín hiệu đến từ các hướng thử nghiệm $\boldsymbol{\Omega}_q$, $1 \leq q \leq Q$.

YÊU CẦU BẢO HỘ

1. Phương pháp giải nén tín hiệu ambisonic bậc cao (Higher Order Ambisonics - HOA) đã được nén mà bao gồm tín hiệu có hướng được mã hóa và tín hiệu xung quanh được mã hóa, phương pháp này bao gồm các bước:

nhận tín hiệu HOA đã nén;

giải mã cảm quan tín hiệu HOA đã nén để tạo ra tín hiệu HOA có hướng được giải mã và tín hiệu HOA xung quanh được giải mã;

thực hiện mở rộng bậc trên tín hiệu HOA xung quanh được giải mã để thu được dạng biểu diễn của tín hiệu HOA xung quanh được giải mã;

biên soạn lại dạng biểu diễn tín hiệu HOA được giải mã từ dạng biểu diễn của tín hiệu HOA xung quanh được giải mã và tín hiệu HOA có hướng được giải mã.

2. Phương pháp theo điểm 1, trong đó dạng biểu diễn tín hiệu HOA được giải mã có bậc lớn hơn một.

3. Phương pháp theo điểm 2, trong đó bậc của tín hiệu HOA xung quanh được giải mã nhỏ hơn bậc của dạng biểu diễn tín hiệu HOA được giải mã.

4. Thiết bị giải nén tín hiệu ambisonic bậc cao (Higher Order Ambisonics - HOA) đã được nén mà bao gồm tín hiệu có hướng được mã hóa và tín hiệu xung quanh được mã hóa, thiết bị này bao gồm:

giao diện đầu vào để nhận tín hiệu HOA đã nén;

bộ giải mã âm thanh mà giải mã cảm quan tín hiệu HOA đã nén để tạo ra tín hiệu HOA có hướng được giải mã và tín hiệu HOA xung quanh được giải mã;

bộ xử lý để thực hiện mở rộng bậc trên tín hiệu HOA xung quanh được giải mã để thu được dạng biểu diễn của tín hiệu HOA xung quanh được giải mã;

bộ tổng hợp để biên soạn lại dạng biểu diễn tín hiệu HOA được giải mã từ dạng biểu diễn của tín hiệu HOA xung quanh được giải mã và tín hiệu HOA có hướng được giải mã.

5. Thiết bị theo điểm 4, trong đó dạng biểu diễn tín hiệu HOA được giải mã có bậc lớn hơn một.
6. Thiết bị theo điểm 5, trong đó bậc của tín hiệu HOA xung quanh được giải mã nhỏ hơn bậc của dạng biểu diễn tín hiệu HOA được giải mã.
7. Vật ghi bất biến đọc được bằng máy tính chứa các lệnh mà khi được thực thi bởi bộ xử lý thực hiện phương pháp theo điểm 1.

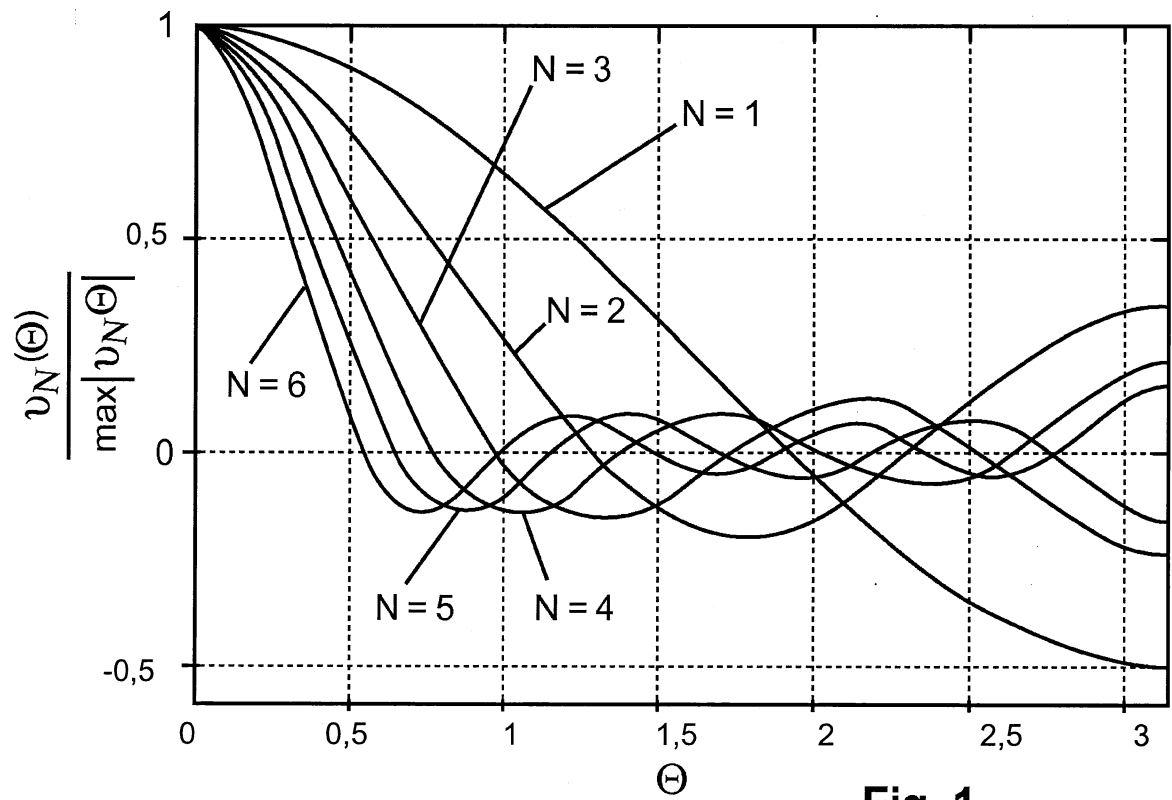


Fig. 1

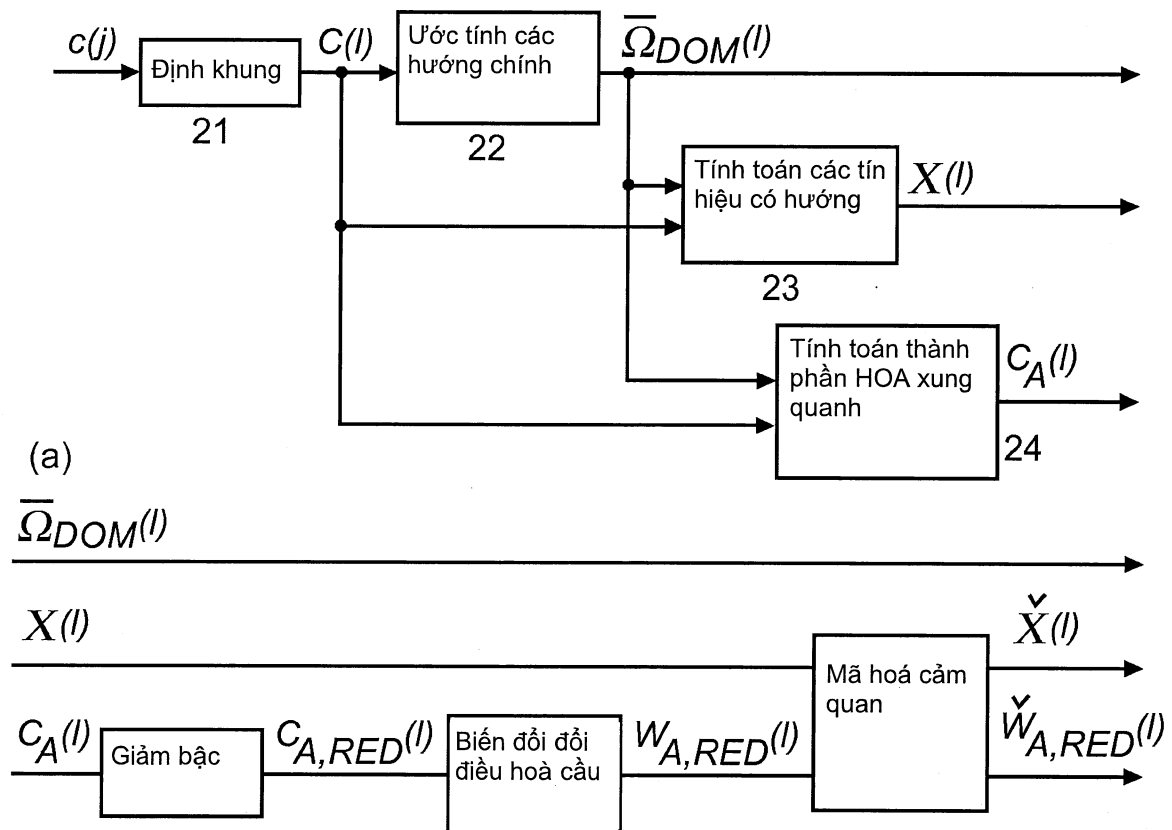


Fig. 2

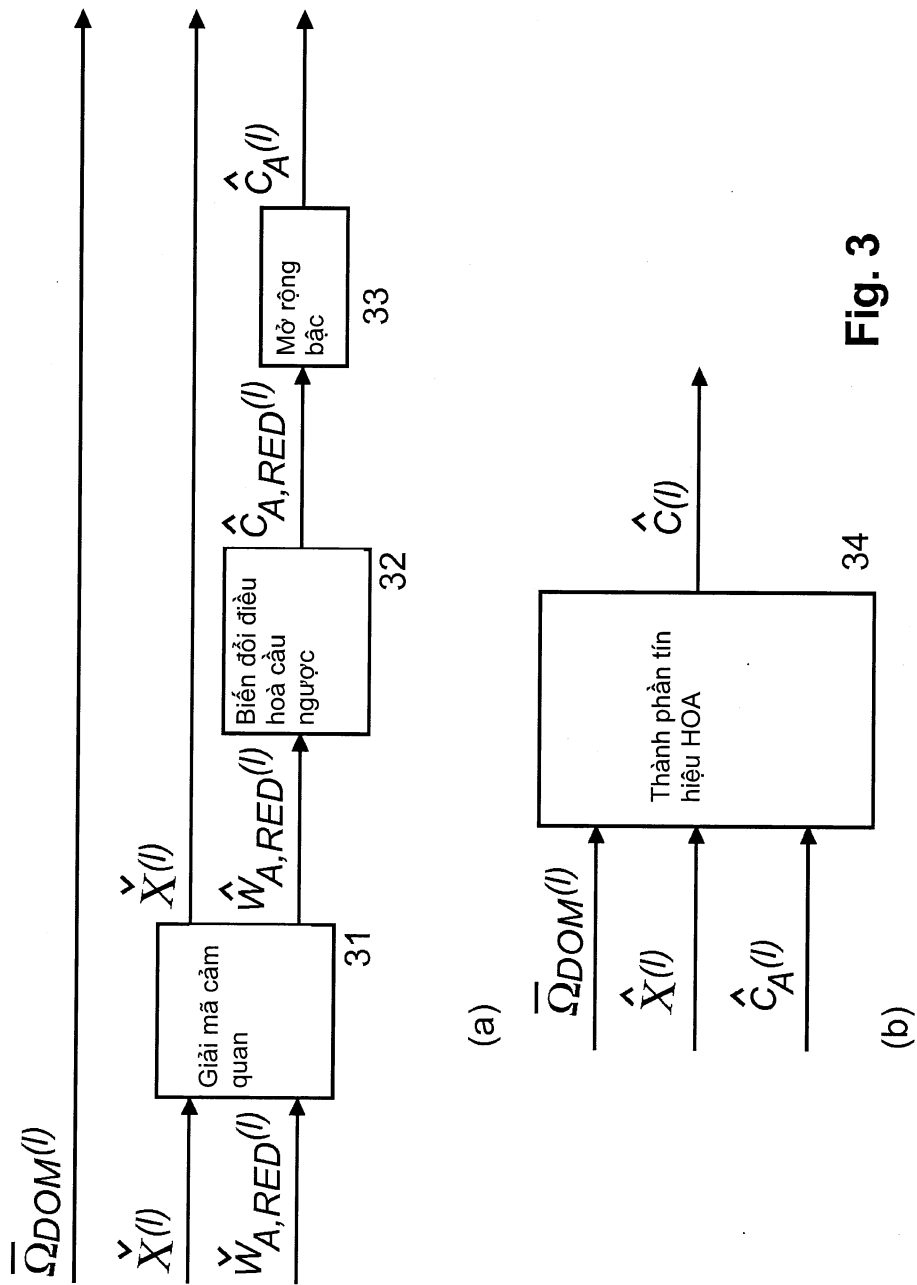


Fig. 3