



(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)



1-0050038

(51)^{2016.01} G10L 19/02; G10L 19/022

(13) B

(21) 1-2024-02090

(22) 10/06/2016

(62) 1-2021-07907

(86) PCT/EP2016/063371 10/06/2016

(87) WO 2016/202701 22/12/2016

(30) 15172282.4 16/06/2015 EP; 15189398.9 12/10/2015 EP

(45) 25/08/2025 449

(43) 25/07/2024 436

(73) Fraunhofer-Gesellschaft zur Foerderung der angewandten Forschung e.V. (DE)
Hansastraße 27c, 80686 Muenchen, Germany

(72) SCHNELL, Markus (DE); LUTZKY, Manfred (DE); FOTOPOULOU, Eleni (GR);
SCHMIDT, Konstantin (DE); BENNDORF, Conrad (DE); TOMASEK, Adrian
(DE); ALBERT, Tobias (DE); SEIDL, Timon (DE).

(74) Công ty Luật TNHH Ambys Hà Nội (AMBYS HANOI)

(54) BỘ GIẢI MÃ ÂM THANH VÀ PHƯƠNG PHÁP GIẢI MÃ TÍN HIỆU ÂM
THANH

(21) 1-2024-02090

(57) Sáng chế đề cập đến bộ giải mã âm thanh và phương pháp giải mã tín hiệu âm thanh. Phiên bản được rút gọn của thủ tục giải mã âm thanh có thể hiệu quả hơn và/hoặc đạt được sự duy trì phù hợp được cải thiện nếu cửa sổ tổng hợp được sử dụng để giải mã âm thanh rút gọn là phiên bản rút gọn của cửa sổ tổng hợp tham chiếu chứa trong thủ tục giải mã âm thanh không được rút gọn bằng cách lấy mẫu giảm bởi thừa số lấy mẫu giảm mà tốc độ lấy mẫu giảm và tốc độ lấy mẫu gốc lệch nhau, và được lấy mẫu giảm sử dụng phép nội suy theo phân đoạn trong các phân đoạn $\frac{1}{4}$ độ dài khung.

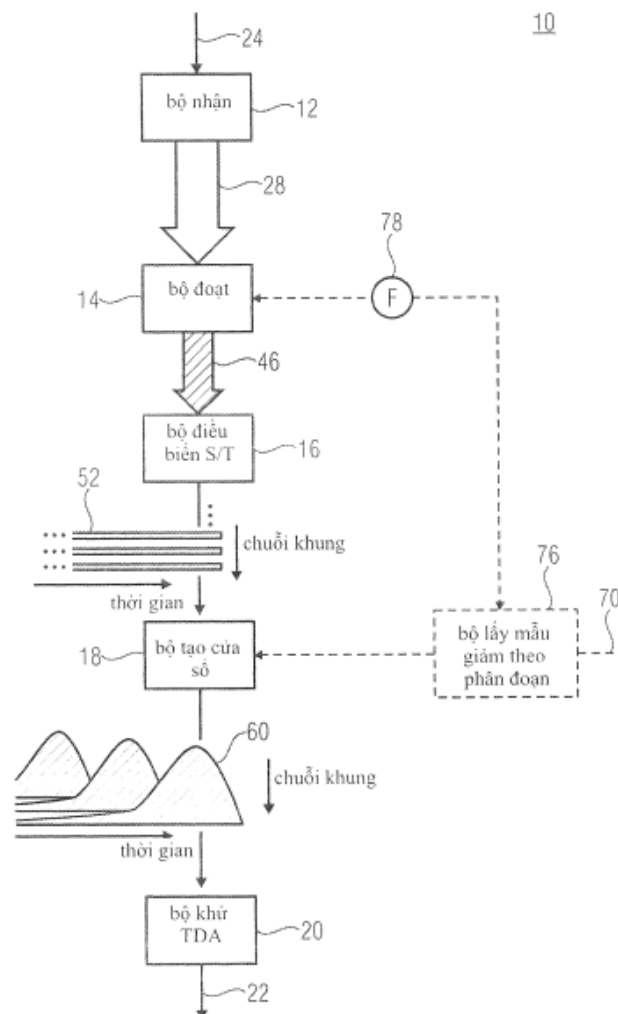


FIG 2

Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến khía cạnh giải mã được rút gọn.

Tình trạng kỹ thuật của sáng chế

Mã hóa âm thanh tiên tiến (Advanced Audio Coding – AAC) độ trễ thấp được tăng cường (Enhanced Low Delay AAC - AAC-ELD) MPEG-4 thường hoạt động ở tốc độ bit lên đến 48kHz, mà dẫn đến độ trễ thuật toán là 15 mili giây (milisecond - ms). Đối với một số ứng dụng, ví dụ sự truyền nhép môi của âm thanh, độ trễ thấp hơn đều nhau là đáng mong muốn. AAC-ELD đã cung cấp lựa chọn này bằng cách hoạt động ở các tốc độ lấy mẫu cao hơn, ví dụ 96kHz, và do đó cung cấp các chế độ hoạt động với độ trễ thấp hơn ngang bằng, ví dụ 7,5ms. Tuy nhiên, chế độ hoạt động này đi kèm với độ phức tạp cao không cần thiết do tốc độ lấy mẫu cao.

Giải pháp cho vấn đề này là áp dụng phiên bản được rút gọn của giàn lọc và do đó, kết xuất tín hiệu âm thanh ở tốc độ lấy mẫu thấp hơn, ví dụ 48kHz thay vì 96kHz. Hoạt động rút gọn đã là một phần của AAC-ELD như được kế thừa từ bộ mã hóa-giải mã MPEG-4 AAC-LD, mà đóng vai trò là cơ sở cho AAC-ELD.

Tuy nhiên, câu hỏi vẫn còn là làm cách nào để tìm kiếm phiên bản được rút gọn của giàn lọc cụ thể. Tức là, sự không chắc chắn duy nhất là cách các hệ số cửa sổ được suy ra trong khi cho phép kiểm tra sự phù hợp rõ ràng của các chế độ hoạt động được rút gọn của bộ giải mã AAC-ELD.

Trong phần sau đây, các nguyên tắc về chế độ hoạt động được rút gọn của các bộ mã hóa-giải mã AAC-(E)LD sẽ được mô tả.

Chế độ hoạt động rút gọn hoặc AAC-LD được mô tả cho AAC-LD trong ISO/IEC 14496-3:2009 trong đoạn 4.6.17.2.7 “Adaptation to systems using lower sampling rates” (Sự thích ứng với các hệ thống sử dụng tốc độ lấy mẫu thấp hơn) như sau:

“Trong các ứng dụng nhất định, có thể cần thiết để tích hợp bộ giải mã độ trễ thấp vào hệ thống âm thanh chạy ở tốc độ lấy mẫu thấp hơn (ví dụ 16kHz) trong khi tốc độ lấy mẫu danh nghĩa của trọng tải dòng bit là cao hơn nhiều (ví dụ 48kHz, tương ứng với độ trễ bộ mã hóa-giải mã thuật toán là xấp xỉ 20ms). Trong các trường hợp này, là thuận lợi để giải mã đầu ra của bộ mã hóa-giải mã độ trễ thấp trực tiếp ở tốc độ

lấy mẫu mục tiêu hơn là sử dụng hoạt động chuyển đổi tốc độ lấy mẫu bổ sung sau khi giải mã.

Điều này có thể được làm gần đúng bởi sự rút gọn thích hợp của cả hai, kích thước khung và tốc độ lấy mẫu, bởi một số thừa số nguyên (ví dụ 2, 3), dẫn đến độ phân giải thời gian/tần số giống nhau của bộ mã hóa-giải mã. Ví dụ, đầu ra của bộ mã hóa-giải mã có thể được tạo ra ở tốc độ lấy mẫu 16kHz thay vì tốc độ lấy mẫu danh nghĩa 48kHz bằng cách giữ lại một phần ba thấp nhất (tức là $480/3 = 160$) của hệ số phổ trước khâu lọc tổng hợp và giảm kích thước biến đổi nghịch đảo tới một phần ba (tức là kích thước của số $960/3 = 320$).

Kết quả là, việc giải mã cho các tốc độ lấy mẫu thấp hơn làm giảm cả yêu cầu về bộ nhớ và yêu cầu tính toán, nhưng có thể không đem lại chính xác đầu ra giống như giải mã bằng thông đầy đủ, theo sau bởi giới hạn băng và chuyển đổi tốc độ mẫu.

Lưu ý rằng việc giải mã ở tốc độ lấy mẫu thấp hơn, như được mô tả ở trên, không ảnh hưởng đến việc giải thích các mức, mà đề cập đến tốc độ lấy mẫu danh nghĩa của trọng tải dòng bit độ trễ thấp AAC.”

Lưu ý rằng AAC-LD làm việc với khung làm việc biến đổi cosin rời rạc cải biên (modified discrete cosine transform – MDCT) tiêu chuẩn và hai dạng cửa sổ, tức là cửa sổ hình sin và cửa sổ chồng lấp thấp. Cả hai cửa sổ được mô tả đầy đủ bởi các công thức và do đó, các hệ số cửa sổ cho các độ dài biến đổi bất kỳ có thể được xác định.

So sánh với AAC-LD, bộ mã hóa-giải mã AAC-ELD thể hiện hai khác biệt chính:

- Cửa sổ MDCT độ trễ thấp (Low Delay MDCT window - LD-MDCT)
- Khả năng sử dụng công cụ sao chép băng phổ (Spectral Band Replication - SBR độ trễ thấp)

Thuật toán biến đổi cosin rời rạc cải biên nghịch đảo (inverse MDCT – IMDCT) sử dụng cửa sổ MDCT độ trễ thấp được mô tả trong 4.6.20.2 trong tài liệu tham khảo [1], mà rất giống phiên bản IMDCT tiêu chuẩn sử dụng, ví dụ cửa sổ hình sin. Các hệ số của các cửa sổ MDCT độ trễ thấp (kích thước khung 480 và 512 mẫu) được cho trong bảng 4.A.15 và 4.A.16 trong tài liệu tham khảo [1]. Lưu ý là các hệ số không thể được xác định bởi công thức, như các hệ số là kết quả của thuật toán tối ưu hóa. Fig.9 thể hiện sơ đồ về hình dạng cửa sổ đối với kích thước khung 512.

Trong trường hợp công cụ SBR độ trễ thấp (low delay SBR - LD-SBR) được sử dụng kết hợp với bộ mã hóa AAC-ELD, các giàn lọc của môđun LD-SBR cũng được rút gọn. Điều này đảm bảo rằng môđun SBR hoạt động với cùng độ phân giải tần số và do đó, không yêu cầu sự thích ứng hơn nào.

Do đó, sự mô tả ở trên bộc lộ rằng có nhu cầu để rút gọn các hoạt động giải mã như, ví dụ, rút gọn phép giải mã tại AAC-ELD. Sẽ là khả thi để tìm ra các hệ số cho hàm cửa sổ tổng hợp rút gọn lần nữa, nhưng đây là nhiệm vụ nặng nề, đòi hỏi lưu trữ bổ sung để lưu trữ phiên bản được rút gọn và kết xuất sự kiểm tra phù hợp giữa phép giải mã không rút gọn và giải mã rút gọn phức tạp hơn hoặc, từ quan điểm khác, không tuân theo cách rút gọn được yêu cầu trong, ví dụ, AAC-ELD. Phụ thuộc vào tỉ lệ rút gọn, tức là tỉ lệ giữa tốc độ lấy mẫu gốc và tốc độ lấy mẫu rút gọn, có thể suy ra hàm cửa sổ tổng hợp rút gọn một cách đơn giản bằng cách lấy mẫu giảm, tức là chọn ra mọi hệ số cửa sổ thứ hai, thứ ba... của hàm cửa sổ tổng hợp gốc, nhưng thủ tục này không dẫn đến sự phù hợp đầy đủ của phép giải mã không rút gọn và giải mã rút gọn, một cách tương ứng. Việc sử dụng các thủ tục thập phân phức tạp hơn được áp dụng cho hàm cửa sổ tổng hợp, dẫn đến độ lệch không chấp nhận được từ hình dạng hàm cửa sổ tổng hợp gốc. Do đó, có nhu cầu trong tình trạng kỹ thuật đối với khái niệm giải mã rút gọn được cải thiện.

Tài liệu tham khảo

[1] ISO/IEC 14496-3:2009

[2] M13958, “Proposal for an Enhanced Low Delay Coding Mode”, 10/2006, Hangzhou, Trung Quốc

Bản chất kỹ thuật của sáng chế

Mục đích của sáng chế là cung cấp sơ đồ giải mã âm thanh mà cho phép giải mã rút gọn được cải thiện.

Mục đích này đạt được bởi đối tượng của các điểm yêu cầu bảo hộ độc lập.

Sáng chế dựa trên việc tìm ra rằng phiên bản được rút gọn của thủ tục giải mã âm thanh có thể hiệu quả hơn và/hoặc đạt được sự duy trì phù hợp được cải thiện nếu cửa sổ tổng hợp được sử dụng để giải mã âm thanh được rút gọn là phiên bản được rút gọn của cửa sổ tổng hợp tham chiếu chứa trong thủ tục giải mã âm thanh không được rút gọn bằng cách lấy mẫu giảm bởi thừa số lấy mẫu giảm mà tốc độ lấy mẫu được lấy

mẫu giảm và tốc độ lấy mẫu gốc lệch nhau, và được lấy mẫu giảm sử dụng phép nội suy theo phân đoạn trong các phân đoạn bằng $1/4$ độ dài khung.

Mô tả vắn tắt các hình vẽ

Các khía cạnh có lợi của sáng chế là đối tượng của các điểm yêu cầu bảo hộ phụ thuộc. Các phương án ưu tiên của sáng chế được mô tả dưới đây dựa vào các hình vẽ đi kèm, trong đó:

Fig.1 là sơ đồ giản lược minh họa các yêu cầu khôi phục hoàn hảo cần được tuân theo khi giải mã rút gọn để bảo toàn sự khôi phục hoàn hảo.

Fig.2 thể hiện sơ đồ khối của bộ giải mã âm thanh để giải mã rút gọn theo phương án của sáng chế;

Fig.3 là sơ đồ giản lược minh họa, trong nửa phía trên, cách mà tín hiệu âm thanh đã được mã hóa tại tốc độ lấy mẫu gốc thành dòng dữ liệu và, trong nửa dưới được phân tách khỏi nửa trên bởi đường nét đứt nằm ngang, hoạt động giải mã rút gọn để khôi phục tín hiệu âm thanh từ dòng dữ liệu tại tốc độ lấy mẫu giảm hoặc rút gọn, để minh họa chế độ hoạt động của bộ giải mã âm thanh của Fig.2;

Fig.4 là sơ đồ giản lược minh họa sự kết hợp của bộ tạo cửa sổ và bộ khử răng cưa miền thời gian của Fig.2;

Fig.5 minh họa việc thực thi có khả năng để đạt được sự khôi phục theo Fig.4 sử dụng phép xử lý đặc biệt của các phần có trọng số bằng 0 của các phần thời gian được điều biến phổ thành thời gian;

Fig.6 là sơ đồ giản lược minh họa phép lấy mẫu giảm để thu được cửa sổ tổng hợp được lấy mẫu giảm;

Fig.7 là sơ đồ khối minh họa hoạt động rút gọn của AAC-ELD gồm công cụ SBR độ trễ thấp;

Fig.8 thể hiện sơ đồ khối của bộ giải mã âm thanh để giải mã rút gọn theo phương án trong đó bộ điều biến, bộ tạo cửa sổ và bộ khử được thực hiện trong phép thực thi nâng; và

Fig.9 là sơ đồ về các hệ số cửa sổ của cửa sổ độ trễ thấp theo AAC-ELD cho kích thước khung 512 mẫu như ví dụ về cửa sổ tổng hợp tham chiếu cần được lấy mẫu giảm.

Mô tả chi tiết sáng chế

Phần mô tả sau đây bắt đầu với sự minh họa của phương án để giải mã rút gọn

đối với bộ mã hóa-giải mã AAC-ELD. Tức là, phần mô tả sau đây bắt đầu với phương án mà có thể tạo ra chế độ được rút gọn cho AAC-ELD. Phần mô tả này đồng thời tạo thành một dạng giải thích về sự thúc đẩy dưới các phương án của sáng chế này. Sau đây, sự mô tả được khái quát, do đó dẫn đến sự mô tả về bộ giải mã âm thanh và phương pháp giải mã âm thanh theo phương án của sáng chế.

Như được mô tả trong phần giới thiệu về đặc điểm kỹ thuật của sáng chế, AAC-ELD sử dụng các cửa sổ MDCT độ trễ thấp. Để tạo ra các phiên bản được rút gọn của chúng, tức là các cửa sổ độ trễ thấp được rút gọn, đề xuất được giải thích sau đây để tạo thành chế độ rút gọn cho AAC-ELD sử dụng thuật toán nội suy hàm nối trực theo phân đoạn mà duy trì đặc tính khôi phục hoàn hảo (perfect reconstruction - PR) của cửa sổ LD-MDCT với độ chính xác cao. Do đó, thuật toán cho phép tạo ra các hệ số cửa sổ ở dạng trực tiếp, như được mô tả trong ISO/IEC 14496-3:2009, cũng như dạng nâng, như được mô tả trong tài liệu tham khảo [2], theo cách tương thích. Điều này có nghĩa là cả hai cách thực hiện tạo ra đầu ra 16 bit phù hợp.

Phép nội suy của cửa sổ MDCT độ trễ thấp được thực hiện như sau:

Nói chung, phép nội suy hàm nối trực sẽ được sử dụng để tạo ra các hệ số cửa sổ được rút gọn để duy trì đáp ứng tần số và hầu hết đặc tính khôi phục hoàn hảo (quanh 170dB SNR). Phép nội suy cần có sự ràng buộc trong các phân đoạn nhất định để duy trì đặc tính khôi phục hoàn hảo. Đối với các hệ số cửa sổ c bao hàm nhân DCT của phép biến đổi (xem Fig.1, $c(1024) \dots c(2048)$), yêu cầu ràng buộc sau đây,

$$1 = |(\text{sgn} \cdot c(i) \cdot c(2N - 1 - i) + c(N + i) \cdot c(N - 1 - i))| \quad \text{với } i = 0 \dots N/2 - 1 \quad (1)$$

trong đó N là kích thước khung. Một số cách thực hiện có thể sử dụng các ký hiệu khác để tối ưu hóa độ phức tạp, ở đây, được ký hiệu bởi sgn . Yêu cầu trong (1) có thể được minh họa bởi Fig.1. Nên nhớ rằng đơn giản là ngay cả trong trường hợp $F=2$, tức là lấy nửa tốc độ mẫu, để lại mọi hệ số cửa sổ thứ hai của cửa sổ tổng hợp tham chiếu để thu được cửa sổ tổng hợp được rút gọn không đáp ứng yêu cầu.

Các hệ số $c(0) \dots c(2N - 1)$ được liệt kê dọc theo dạng hình thoi. $N/4$ số không trong các hệ số cửa sổ, mà đáp ứng được cho sự giảm độ trễ của giàn lọc, được đánh dấu sử dụng mũi tên nét đậm. Fig.1 là sự phụ thuộc của các hệ số do sự tạo nếp chứa trong MDCT và cả các điểm mà phép nội suy cần ràng buộc để tránh sự phụ thuộc không mong muốn bất kỳ.

- Mọi $N/2$ hệ số, phép nội suy cần ngừng để duy trì (1)
- Ngoài ra, thuật toán nội suy cần ngừng mọi $N/4$ hệ số do các số không được chèn. Điều này đảm bảo rằng các số không được duy trì và lỗi nội suy không phổ biến mà duy trì PR.

Ràng buộc thứ hai không chỉ yêu cầu cho phân đoạn chứa các số không mà còn yêu cầu cho các phân đoạn khác. Biết rằng một số hệ số trong nhân DCT không được xác định bởi thuật toán tối ưu hóa mà được xác định bởi công thức (1) để cho phép PR, một vài điểm gián đoạn trong dạng cửa sổ có thể được giải thích, ví dụ quanh $c(1536+128)$ trên Fig.1. Để giảm thiểu lỗi PR, phép nội suy cần ngừng ở các điểm này, mà xuất hiện trong lưới $N/4$.

Vì lý do này, kích thước phân đoạn $N/4$ được chọn cho phép nội suy hàm nối trực theo phân đoạn để tạo ra các hệ số cửa sổ được rút gọn. Các hệ số cửa sổ nguồn luôn được cho bởi các hệ số được sử dụng cho $N = 512$, cũng cho các hoạt động rút gọn dẫn đến các kích thước khung là $N = 240$ hoặc $N = 120$. Thuật toán cơ bản được tóm tắt một cách rất ngắn gọn trong phần sau đây như mã MATLAB:

```
FAC = Downscaling factor    % e.g. 0.5
sb = 128;                  % segment size of source window
w_down = [];               % downsampled window
nSegments = length(W)/(sb); % number of segments; W=LD window
                           coefficients for N=512

xn = ((0:(FAC*sb-1))+0.5)/FAC-0.5; % spline init
for i=1:nSegments,
    w_down=[w_down,spline([0:(sb-1)],W((i-1)*sb+(1:(sb))),xn)];
end;
```

Vì hàm nối trực có thể không hoàn toàn tắt định, thuật toán hoàn chỉnh được cụ thể hóa một cách chính xác trong phần sau đây, mà có thể được nêu ra trong ISO/IEC 14496-3:2009, để tạo thành chế độ rút gọn được cải thiện trong AAC-ELD.

Nói cách khác, phần sau đây cung cấp đề xuất về cách để ý tưởng được nêu ở trên có thể được áp dụng cho ER AAC ELD, tức là về cách để bộ giải mã độ phức tạp thấp có thể giải mã dòng bit ER AAC ELD được mã hóa ở tốc độ dữ liệu thứ nhất tại tốc độ dữ liệu thứ hai thấp hơn tốc độ dữ liệu thứ nhất. Tuy nhiên, nhấn mạnh là việc xác định N như được sử dụng trong phần sau đây tuân thủ tiêu chuẩn. Ở đây, N tương ứng với chiều dài của nhân DCT trong khi ở phần trước, trong các điểm yêu cầu bảo

hệ, và các phương án tổng quát được mô tả sau đây, N tương ứng với độ dài khung, cụ thể là độ dài chồng lấp lẫn nhau của các nhân DCT, tức là nửa độ dài nhân DCT. Theo đó, trong khi N được biểu thị là 512 ở trên đây, ví dụ, sẽ được biểu thị là 1024 trong phần sau đây.

Các đoạn sau đây được đề xuất để bao gồm 14496-3:2009 thông qua sửa đổi.

A.0 Thích ứng với các hệ thống sử dụng các tốc độ lấy mẫu thấp hơn

Đối với các ứng dụng nhất định, ER AAC LD có thể thay đổi tốc độ mẫu phát lại để tránh các bước lấy mẫu bổ sung (xem 4.6.17.2.7). ER AAC ELD có thể áp dụng các bước rút gọn tương tự sử dụng của số MDCT độ trễ thấp và công cụ LD-SBR. Trong trường hợp AAC-ELD hoạt động với công cụ LD-SBR, thừa số rút gọn được giới hạn tới các bội số của 2. Không có LD-SBR, kích thước khung rút gọn cần là số nguyên.

A.1 Rút gọn của số MDCT độ trễ thấp

Cửa sổ LD-MDCT w_{LD} đối với $N=1024$ được rút gọn bởi thừa số F sử dụng phép nội suy hàm nội trục theo phân đoạn. Số lượng của các số không dẫn trước trong các hệ số cửa sổ, tức là $N/8$, xác định kích thước phân đoạn. Các hệ số cửa sổ được rút gọn w_{LD_d} được sử dụng cho MDCT nghịch đảo như được mô tả trong 4.6.20.2 nhưng với độ dài cửa sổ được rút gọn $N_d = N/F$. Lưu ý là thuật toán cũng có thể tạo ra các hệ số nâng được rút gọn của LD-MDCT.

```
fs_window_size = 2048; /* Number of fullscale window coefficients.
According to ISO/IEC 14496-3:2009,
        use 2048. For lifting implemenations, please adjust
this variable accordingly */
ds_window_size = N * fs_window_size / (1024 * F); /* downscaled
window coefficients; N determines the
        transformation length according to 4.6.20.2 */
fs_segment_size = 128;
num_segments    = fs_window_size / fs_segment_size;
ds_segment_size = ds_window_size / num_segments;
tmp[128], y[128]; /* temporary buffers */

/* loop over segments */
for (b = 0; b < num_segments; b++) {
    /* copy current segment to tmp */
    copy(&W_LD[b * fs_segment_size], tmp, fs_segment_size);

    /* apply cubic spline interpolation for downscaling */
    /* calculate interpolating phase */
    phase = (fs_window_size - ds_window_size) / (2 *
ds_window_size);
```

```

/* calculate the coefficients c of the cubic spline given tmp */
/* array of precalculated constants */
m = {0.166666672, 0.25, 0.266666681, 0.267857134,
      0.267942578, 0.267948717, 0.267949164};
n = fs_segment_size; /* for simplicity */

/* calculate vector r needed to calculate the coefficients c */
for (i = n - 3; i >= 0; i--)
    r[i] = 3 * ((tmp[i + 2] - tmp[i + 1]) - (tmp[i + 1] -
tmp[i]));
for (i = 1; i < 7; i++)
    r[i] -= m[i - 1] * r[i - 1];
for(i = 7; i < n - 4; i++)
    r[i] -= 0.267949194 * r[i - 1];

/* calculate coefficients c */
c[n - 2] = r[n - 3] / 6;
c[n - 3] = (r[n - 4] - c[n - 2]) * 0.25;
for (i = n - 4; i > 7; i--)
    c[i] = (r[i - 1] - c[i + 1]) * 0.267949194;
for (i = 7; i > 1; i--)
    c[i] = (r[i - 1] - c[i + 1]) * m[i - 1];
c[1] = r[0] * m[0];
c[0] = 2 * c[1] - c[2];
c[n - 1] = 2 * c[n - 2] - c[n - 3];

/* keep original samples in temp buffer y because samples of
tmp will be replaced with interpolated samples */
copy(tmp, y, fs_segment_size);

/* generate downscaled points and do interpolation */
for (k = 0; k < ds_segment_size; k++) {
    step = phase + k * fs_segment_size / ds_segment_size;
    idx = floor(step);
    diff = step - idx;
    di = (c[idx + 1] - c[idx]) / 3;
    bi = (y[idx + 1] - y[idx]) - (c[idx + 1] + 2 * c[idx]) / 3;
    /* calculate downscaled values and store in tmp */
    tmp[k] = y[idx] + diff * (bi + diff * (c[idx] + diff * di));
}

/* assemble downscaled window */
copy(tmp, &W_LD_d[b * ds_segment_size], ds_segment_size);
}

```

A.2 Sự rút gọn của công cụ SBR độ trễ thấp

Trong trường hợp công cụ SBR độ trễ thấp được sử dụng kết hợp với ELD, công cụ này có thể được rút gọn tới tốc độ mẫu thấp hơn, ít nhất là để rút gọn các thừa số là bội số của 2. Thừa số rút gọn F điều khiển số lượng các băng được sử dụng cho giàn lọc phân tích và tổng hợp CLDFB. Hai đoạn sau đây mô tả giàn lọc phân tích và tổng hợp CLDFB được rút gọn, xem 4.6.19.4.

4.6.20.5.2.1 Giàn lọc CLDFB phân tích được rút gọn

- Xác định số lượng băng CLDFB rút gọn $B = 32/F$.

Dịch chuyển các mẫu trong mảng x bởi B vị trí. B mẫu cũ nhất bị loại bỏ và B mẫu mới được lưu trữ trong các vị trí từ 0 đến $B-1$.

Nhân các mẫu của mảng x bởi hệ số cửa sổ c_i để có được mảng z . Các hệ số cửa sổ c_i thu được bởi phép nội suy tuyến tính của các hệ số c , tức là qua phương trình:

$$c_i(i) = \frac{1}{2} [c(2F \cdot i + 1 + p) + c(2F \cdot i + p)], \quad 0 \leq i < (10B), \quad p = \text{int}\left(\frac{64}{2B} - 0.5\right).$$

Các hệ số cửa sổ của c có thể được tìm thấy trong Bảng 4.A.90.

Tính tổng các mẫu để tạo ra mảng phân tử $2B$ u :

$$u(n) = z(n) + z(n + 2B) + z(n + 4B) + z(n + 6B) + z(n + 8B), \quad 0 \leq n < (2B).$$

Tính toán B mẫu bằng con mới bởi phép toán ma trận Mu , trong đó:

$$M(k, n) = 2 \cdot \exp\left(\frac{j \cdot \pi \cdot (k + 0.5) \cdot (2n - (3B - 1))}{2B}\right), \quad \begin{cases} 0 \leq k < B \\ 0 \leq n < 2B \end{cases}$$

Trong phương trình, $\exp(\cdot)$ biểu thị hàm mũ phức và j là đơn vị ảo.

4.6.20.5.2.2 Giàn lọc CLDFB tổng hợp được rút gọn

Xác định số lượng băng CLDFB rút gọn $B = 64/F$.

Dịch chuyển các mẫu trong mảng v bởi $2B$ vị trí. $2B$ mẫu cũ nhất bị loại bỏ.

B mẫu bằng con có giá trị phức được nhân với ma trận N , trong đó:

$$N(k, n) = \frac{1}{64} \cdot \exp\left(\frac{j \cdot \pi \cdot (k + 0.5) \cdot (2n - (B - 1))}{2B}\right), \quad \begin{cases} 0 \leq k < B \\ 0 \leq n < 2B \end{cases}$$

Trong phương trình, $\exp(\cdot)$ biểu thị hàm mũ phức và j là đơn vị ảo. Phần thực của đầu ra từ phép toán này được lưu trữ trong các vị trí 0 đến $2B - 1$ của mảng v .

Trích ra các mẫu từ v để tạo ra mảng phân tử $10B$ g .

$$\begin{aligned} g(2B \cdot n + k) &= v(4B \cdot n + k) \\ g(2B \cdot n + B + k) &= v(4B \cdot n + 3B + k) \end{aligned} \quad \begin{cases} 0 \leq n \leq 4 \\ 0 \leq k < B \end{cases}$$

Nhân các mẫu của mảng g bởi hệ số cửa sổ c_i để có được mảng w . Các hệ số cửa sổ c_i thu được bởi phép nội suy tuyến tính của các hệ số c , tức là qua phương trình:

$$c_i(i) = \frac{1}{2} [c(2F \cdot i + 1 + p) + c(2F \cdot i + p)], \quad 0 \leq i < (10B), p = \text{int}\left(\frac{64}{2B} - 0,5\right).$$

Các hệ số cửa sổ c có thể được tìm thấy trong Bảng 4.A.90.

Tính toán B mẫu đầu ra mới bởi phép tính tổng của các mẫu từ mảng w theo

$$\text{output}(n) = \sum_{i=0}^{i \leq 9} w(Bi + n), \quad 0 \leq n < B.$$

Lưu ý rằng việc thiết lập $F = 2$ cung cấp giàn lọc tổng hợp được rút gọn theo 4.6.19.4.3. Do đó, để xử lý dòng bit LD-SBR được rút gọn với thừa số rút gọn bổ sung F , F cần được nhân với 2.

4.6.20.5.2.3 giàn lọc CLDFB có giá trị thực được rút gọn

Việc rút gọn CLDFB có thể cũng được áp dụng cho các phiên bản có giá trị thực của chế độ SBR công suất thấp. Để minh họa, vui lòng xem xét 4.6.19.5.

Đối với giàn lọc phân tích và tổng hợp có giá trị thực được rút gọn, theo dõi phần mô tả trong 4.6.20.5.2.1 và 4.6.20.2.2 và trao đổi bộ điều biến $\exp()$ trong M bởi bộ điều biến $\cos()$.

A.3 Phân tích MDCT độ trễ thấp

Phần phụ này mô tả giàn lọc MDCT độ trễ thấp được sử dụng trong bộ mã hóa AAC ELD. Thuật toán MDCT lõi hầu như không thay đổi, nhưng với cửa sổ dài hơn, sao cho n được chạy từ $-N$ đến $N-1$ (thay vì từ 0 đến $N-1$)

Hệ số phổ, $X_{i,k}$, được xác định như sau:

$$X_{i,k} = -2 \cdot \sum_{n=-N}^{N-1} z_{i,n} \cos\left(\frac{2\pi}{N}(n + n_0) \left(k + \frac{1}{2}\right)\right) \quad \text{với } 0 \leq k < N/2$$

trong đó:

$z_{i,n}$ = chuỗi đầu vào được tạo cửa sổ
 N = chỉ số mẫu

K = chỉ số hệ số phổ
 I = chỉ số khối
 N = độ dài cửa sổ
 n_0 = $(-N / 2 + 1) / 2$

Độ dài cửa sổ N (dựa trên cửa sổ hình sin) là 1024 hoặc 960.

Độ dài cửa sổ của cửa sổ độ trễ thấp là $2 \cdot N$. Việc tạo cửa sổ được mở rộng về trước theo cách sau:

$$z_{i,n} = w_{LD}(N - 1 - n) \cdot x'_{i,n}$$

đối với $n = -N, \dots, N-1$, với cửa sổ tổng hợp w được sử dụng như cửa sổ phân tích bằng cách nghịch đảo thứ tự.

A.4 Tổng hợp MDCT độ trễ thấp

Giàn lọc tổng hợp được biến đổi so với thuật toán IMDCT tiêu chuẩn sử dụng cửa sổ hình sin theo thứ tự để chấp nhận giàn lọc độ trễ thấp. Thuật toán IMDCT lõi hầu như không thay đổi, nhưng với cửa sổ dài hơn, sao cho n chạy lên đến $2N-1$ (hơn là lên đến $N-1$).

$$x_{\text{im}} = -\frac{2}{N} \sum_{k=0}^{\frac{N}{2}-1} \text{spec}[i][k] \cos\left(\frac{2\pi}{N}(n+n_0)\left(k+\frac{1}{2}\right)\right) \quad \text{với } 0 \leq n < 2N$$

trong đó:

n = chỉ số mẫu
 i = chỉ số cửa sổ
 k = chỉ số hệ số phổ
 N = độ dài cửa sổ/hai lần độ dài khung
 n_0 = $(-N / 2 + 1) / 2$

với $N = 960$ hoặc 1024 .

Việc tạo cửa sổ và chồng lấp-cộng được tiến hành theo cách sau:

Cửa sổ có độ dài N được thay thế bởi cửa sổ có độ dài $2N$ với nhiều chồng lấp hơn ở trước, và ít chồng lấp hơn về sau (các giá trị $N/8$ thực tế là không).

Tạo cửa sổ cho cửa sổ độ trễ thấp:

$$z_{i,n} = w_{LD}(n) \cdot x_{i,n}$$

Trong đó cửa sổ hiện tại có độ dài là $2N$, do đó $n=0, \dots, 2N-1$.

Chồng lắp và cộng:

$$out_{i,n} = z_{i,n} + z_{i-1,n+\frac{N}{2}} + z_{i-2,n+N} + z_{i-3,n+N+\frac{N}{2}}$$

với $0 \leq n < N/2$

Ở đây, các đoạn được đề xuất để được nêu trong 14496-3:2009 thông qua đầu sửa đổi.

Đương nhiên, phần mô tả ở trên về chế độ rút gọn có khả năng cho AAC-ELD chỉ biểu diễn một phương án của sáng chế và một số cải biến là khả thi. Nói chung, các phương án của sáng chế không được giới hạn đến bộ giải mã âm thanh thực hiện phiên bản được rút gọn của giải mã AAC-ELD. Nói cách khác, các phương án của sáng chế có thể, ví dụ, được suy ra bằng cách thực hiện bộ giải mã âm thanh có thể thực hiện quy trình biến đổi nghịch đảo chỉ theo cách được rút gọn mà không mang hoặc sử dụng các nhiệm vụ khác cụ thể của AAC-ELD khác nhau như, ví dụ, truyền dựa trên hệ số tỉ lệ đường bao phổ, lọc tạo hình nhiễu âm theo thời gian (temporal noise shaping - TNS), sao chép băng phổ (spectral band replication - SBR) hoặc tương tự.

Sau đó, phương án tổng quát hơn cho bộ giải mã âm thanh được mô tả. Ví dụ được nêu ở trên cho bộ giải mã âm thanh AAC-ELD hỗ trợ chế độ được rút gọn được mô tả ở trên, do đó có thể biểu diễn sự thực hiện của bộ giải mã âm thanh được mô tả sau đó. Cụ thể, bộ giải mã được giải thích sau đó được thể hiện trên Fig.2 trong khi Fig.3 minh họa các bước thực hiện bởi bộ giải mã của Fig.2.

Bộ giải mã âm thanh của Fig.2, mà thường được biểu thị sử dụng ký hiệu tham chiếu 10, bao gồm bộ nhận 12, bộ đoạt 14, bộ điều biến phổ thành thời gian 16, bộ tạo cửa sổ 18 và bộ khử răng cưa miền thời gian 20, tất cả được nối theo chuỗi với nhau theo thứ tự theo dõi của chúng. Sự tương tác và chức năng của các khối 12 đến 20 của bộ giải mã âm thanh 10 được mô tả trong phần sau đây dựa vào Fig.3. Như được mô tả ở đầu của phần mô tả sáng chế, các khối từ 12 đến 20 có thể được thực hiện trong phần mềm, phần cứng có thể lập trình hoặc phần cứng như ở dạng chương trình máy tính, FPGA hoặc máy tính được lập trình thích hợp, các bộ vi xử lý được lập trình hoặc mạch tích hợp ứng dụng cụ thể với các khối 12 đến 20 biểu diễn các thủ tục con tương ứng, các đường dẫn mạch hoặc tương tự.

Theo cách được nêu chi tiết hơn dưới đây, bộ giải mã âm thanh 10 của Fig.2 được tạo cấu hình để, và các phần tử của bộ giải mã âm thanh 10 được tạo cấu hình để

kết hợp một cách thích hợp để giải mã tín hiệu âm thanh 22 từ dòng dữ liệu 24 với lưu ý là bộ giải mã âm thanh 10 giải mã tín hiệu 22 tại tốc độ lấy mẫu là $1/F$ tốc độ lấy mẫu mà tại đó tín hiệu âm thanh 22 đã được mã hóa biến đổi 24 tại phía mã hóa. Ví dụ, F có thể là số hữu tỉ lớn hơn một. Bộ giải mã âm thanh có thể được tạo kết cấu để hoạt động ở các thừa số rút gọn khác nhau hoặc biến thiên F hoặc ở thừa số cố định. Các cách khác được mô tả chi tiết hơn dưới đây.

Cách mà trong đó tín hiệu âm thanh 22 được mã hóa biến đổi ở tốc độ mã hóa hoặc tốc độ lấy mẫu gốc thành dòng dữ liệu được minh họa trên Fig.3 ở nửa phía trên. Tại 26 của Fig.3 minh họa các hệ số phổ sử dụng các hộp hoặc hình vuông nhỏ 28 được sắp xếp theo kiểu thời gian phổ dọc theo trục thời gian 30 mà lần lượt chạy theo chiều ngang trên Fig.3, và trục tần số 32 mà chạy theo chiều dọc trên Fig.3. Các hệ số phổ 28 được truyền trong dòng dữ liệu 24. Cách mà trong đó các hệ số phổ 28 đã thu được, và do đó cách mà thông qua đó các hệ số phổ 28 biểu diễn tín hiệu âm thanh 22, được minh họa trên Fig.3 tại 34, mà minh họa cho một phần trục thời gian 30 cách mà các hệ số phổ 28 thuộc về, hoặc biểu diễn phần thời gian tương ứng, đã thu được từ tín hiệu âm thanh.

Cụ thể, các hệ số 28 như được truyền trong dòng dữ liệu 24 là các hệ số của phép biến đổi được chồng của tín hiệu âm thanh 22 sao cho tín hiệu âm thanh 22, được lấy mẫu ở tốc độ lấy mẫu gốc hoặc tốc độ lấy mẫu mã hóa, được phân chia thành các khung liên tiếp hoặc không chồng lấp thời gian trực tiếp có độ dài được định trước N , trong đó các hệ số phổ N được truyền trong dòng dữ liệu 24 cho từng khung 36. Tức là, các hệ số biến đổi 28 thu được từ tín hiệu âm thanh 22 sử dụng phép biến đổi chồng được lấy mẫu tới hạn. Trong sự biểu diễn ảnh phổ theo thời gian phổ 26, từng cột của chuỗi theo thời gian trong số các cột của các hệ số phổ 28 phù hợp với một khung tương ứng trong số các khung 36 của chuỗi các khung. N hệ số phổ 28 thu được cho khung tương ứng 36 bởi phép biến đổi phân tích theo phổ hoặc phép điều biến thời gian thành phổ, tuy nhiên các hàm điều biến mà kéo dài theo thời gian không chỉ bắt chéo khung 36 mà thuộc về các hệ số phổ thu được 28, mà còn bắt chéo $E + 1$ khung đứng trước, trong đó E có thể là số nguyên bất kỳ hoặc số nguyên được đánh số chẵn bất kỳ lớn hơn không. Tức là, các hệ số phổ 28 của một cột của ảnh phổ tại 26 mà thuộc về khung 36 nhất định thu được bằng cách áp dụng phép biến đổi lên cửa sổ biến đổi, mà ngoài khung tương ứng bao gồm $E + 1$ khung nằm trong trước đây tương ứng

với khung hiện thời. Phép khai triển phổ của các mẫu của tín hiệu âm thanh trong cửa sổ biến đổi 38 này, mà được minh họa trên Fig.3 cho cột các hệ số biến đổi 28 thuộc về khung giữa 36 của phần được thể hiện tại 34 đạt được sử dụng hàm cửa sổ phân tích đơn thức độ trễ thấp 40 mà các mẫu phổ trong cửa sổ biến đổi 38 được gán trọng số trước khi cùng tùy thuộc vào MDCT hoặc MDST hoặc phép biến đổi khai triển phổ khác. Để giảm độ trễ phía bộ mã hóa, cửa sổ phân tích 40 bao gồm khoảng bằng 0 42 tại đầu dẫn trước theo thời gian của nó sao cho bộ mã hóa không cần đợi phản tương ứng của các mẫu mới nhất trong khung hiện thời 36 để tính toán các hệ số phổ 28 cho khung hiện thời 36 này. Tức là, trong khoảng bằng 0 42 hàm cửa sổ độ trễ thấp 40 là không hoặc có các hệ số cửa sổ bằng 0 sao cho các mẫu âm thanh cùng được định vị của khung hiện thời 36, do trọng số cửa sổ 40, không đóng góp vào các hệ số biến đổi 28 được truyền cho khung đó và dòng dữ liệu 24. Tức là, tóm tắt các vấn đề ở trên, các hệ số biến đổi 28 thuộc về khung hiện thời 36 thu được bằng cách tạo cửa sổ và khai triển phổ của các mẫu của tín hiệu âm thanh trong cửa sổ biến đổi 38 mà bao gồm khung hiện thời cũng như các khung đứng trước theo thời gian và chồng lấp theo thời gian với các cửa sổ biến đổi tương ứng được sử dụng để xác định các hệ số phổ 28 thuộc về các khung lân cận theo thời gian.

Trước khi bắt đầu lại việc mô tả bộ giải mã âm thanh 10, cần được lưu ý rằng sự mô tả của việc truyền các hệ số phổ 28 trong dòng dữ liệu 24 như được cung cấp đến nay đã được đơn giản hóa với cách mà trong đó các hệ số phổ 28 được lượng tử hóa hoặc được mã hóa thành dòng dữ liệu 24 và/hoặc cách mà trong đó tín hiệu âm thanh 22 đã được xử lý trước trước khi đưa tín hiệu âm thanh sang phép biến đổi chồng. Ví dụ, bộ mã hóa âm thanh có tín hiệu âm thanh đã mã hóa biến đổi tín hiệu âm thanh 22 thành dòng dữ liệu 24 có thể được điều khiển thông qua mô hình tâm thính học hoặc có thể sử dụng mô hình tâm thính học để giữ nhiều âm lượng tử hóa và lượng tử hóa các hệ số phổ 28 không thể nhận biết cho người nghe và/hoặc dưới hàm ngưỡng che, do đó xác định các thừa số định tỉ lệ cho các băng phổ mà sử dụng các hệ số phổ được lượng tử hóa và được truyền 28 được định tỉ lệ. Các thừa số định tỉ lệ cũng sẽ được báo hiệu trong dòng dữ liệu 24. Theo cách khác, bộ mã hóa âm thanh có thể là loại kích thích mã hóa biến đổi (TCX - transform coded excitation) của bộ mã hóa. Sau đó, tín hiệu âm thanh sẽ phải trải qua phép lọc phân tích dự báo tuyến tính trước khi tạo thành sự biểu diễn thời gian phổ 26 của các hệ số phổ 28 bằng cách áp dụng phép biến

đổi chồng lên tín hiệu kích thích, tức là tín hiệu dự dự báo tuyến tính. Ví dụ, các hệ số dự báo tuyến tính cũng có thể được báo hiệu trong dòng dữ liệu 24, và phép lượng tử hóa đồng đều có thể được áp dụng để thu được các hệ số phổ 28.

Hơn nữa, phần mô tả được đưa ra cho đến nay cũng đã được đơn giản hóa đối với độ dài khung của khung 36 và/hoặc đối với hàm cửa sổ độ trễ thấp 40. Trong thực tế, tín hiệu âm thanh 22 có thể đã được mã hóa thành dòng dữ liệu 24 theo cách sử dụng kích thước khung biến đổi và/hoặc các cửa sổ 40 khác nhau. Tuy nhiên, phần mô tả được đưa ra trong phần sau đây tập trung trên một cửa sổ 40 và một độ dài khung, dù phần mô tả sau đây có thể dễ dàng được mở rộng tới trường hợp trong đó bộ mã hóa entropy thay đổi các tham số này trong suốt quá trình mã hóa tín hiệu âm thanh thành dòng dữ liệu.

Quay lại bộ giải mã âm thanh 10 của Fig.2 và phần mô tả của bộ giải mã âm thanh 10, bộ nhận 12 nhận dòng dữ liệu 24 và do đó nhận, cho từng khung 36, N hệ số phổ 28, tức là cột tương ứng của các hệ số phổ 28 được thể hiện trên Fig.3. Cần được nhắc lại rằng độ dài theo thời gian của các khung 36, được đo trong các mẫu của tín hiệu hoặc tốc độ lấy mẫu gốc hoặc tốc độ lấy mẫu mã hóa, là N, như được biểu thị trên Fig.3 tại 34, nhưng bộ giải mã âm thanh 10 của Fig.2 được tạo cấu hình để giải mã tín hiệu âm thanh 22 tại tốc độ lấy mẫu giảm. Bộ giải mã âm thanh 10 hỗ trợ, ví dụ, chỉ chức năng giải mã rút gọn được mô tả sau đây. Theo cách khác, bộ giải mã âm thanh 10 sẽ có khả năng khôi phục tín hiệu âm thanh tại tốc độ lấy mẫu gốc hoặc tốc độ lấy mẫu mã hóa, nhưng có thể được chuyển đổi giữa chế độ giải mã rút gọn và chế độ giải mã không rút gọn với chế độ giải mã rút gọn trùng khớp với chế độ của bộ giải mã âm thanh 10 của phép vận hành như được giải thích sau đây. Ví dụ, bộ mã hóa âm thanh 10 có thể được chuyển đổi thành chế độ giải mã rút gọn trong trường hợp mức âm lượng, khả năng môi trường tái tạo giảm hoặc tương tự. Bất cứ khi nào trường hợp thay đổi bộ giải mã âm thanh 10 có thể, ví dụ, chuyển đổi trở lại từ chế độ giải mã rút gọn sang chế độ giải mã không rút gọn. Trong bất kỳ trường hợp nào, theo quy trình giải mã rút gọn của bộ giải mã 10 như được mô tả sau đây, tín hiệu âm thanh 22 được khôi phục tại tốc độ lấy mẫu mà tại đó khung 36 có, ở tốc độ lấy mẫu giảm, độ dài ngắn hơn được đo trong các mẫu của tốc độ lấy mẫu giảm này, cụ thể là độ dài của N/F mẫu tại tốc độ lấy mẫu giảm.

Đầu ra của bộ nhận 12 là chuỗi của N hệ số phổ, cụ thể là một tập hợp của N hệ số phổ, tức là một cột trên Fig.3, trên mỗi khung 36. Đã được nêu ra từ phần mô tả vắn tắt ở trên của quy trình mã hóa biến đổi để tạo thành dòng dữ liệu 24 mà bộ nhận 12 có thể áp dụng các nhiệm vụ khác nhau trong việc thu được N hệ số phổ trên mỗi khung 36. Ví dụ, bộ nhận 12 có thể sử dụng phép giải mã entropy để đọc các hệ số phổ 28 từ dòng dữ liệu 24. Bộ nhận 12 cũng có thể tạo hình theo phổ các hệ số phổ được đọc từ dòng dữ liệu với các thừa số định tỉ lệ được cung cấp trong dòng dữ liệu và/hoặc các thừa số định tỉ lệ được suy ra bởi các hệ số dự báo tuyến tính được truyền trong dòng dữ liệu 24. Ví dụ, bộ nhận 12 có thể thu được các thừa số định tỉ lệ từ dòng dữ liệu 24, cụ thể là trên cơ sở mỗi khung và mỗi khung con, và sử dụng các thừa số định tỉ lệ này để định tỉ lệ các thừa số định tỉ lệ được truyền trong dòng dữ liệu 24. Ngoài ra, bộ nhận 12 có thể suy ra các thừa số định tỉ lệ từ các hệ số dự báo tuyến tính được truyền trong dòng dữ liệu 24, cho từng khung 36, và sử dụng các thừa số định tỉ lệ này để định tỉ lệ các hệ số phổ được truyền 28. Một cách tùy chọn, bộ nhận 12 có thể thực hiện việc điền đầy khoảng trống để điền đầy theo cách tổng hợp các phần được lượng tử hóa bằng 0 trong các tập hợp của N hệ số phổ 18 trên mỗi khung. Ngoài ra hoặc theo cách khác, bộ nhận 12 có thể áp dụng bộ lọc tổng hợp TNS lên hệ số bộ lọc TNS được truyền trên mỗi khung để hỗ trợ việc khôi phục các hệ số phổ 28 từ dòng dữ liệu với các hệ số TNS cũng được truyền trong dòng dữ liệu 24. Chỉ các nhiệm vụ có khả năng được nêu của bộ nhận 12 sẽ được hiểu như danh sách không loại trừ của các số đo có khả năng và bộ nhận 12 có thể thực hiện nhiệm vụ nữa và nhiệm vụ khác có liên quan đến việc đọc các hệ số phổ 28 từ dòng dữ liệu 24.

Do đó bộ đoạt 14 nhận từ bộ nhận 12 ảnh phổ 26 của các hệ số phổ 28 và đoạt, cho từng khung 36, đoạn tần số thấp 44 của N hệ số phổ của khung 36 tương ứng, cụ thể là N/F hệ số phổ tần số thấp nhất.

Tức là, bộ điều biến phổ sang thời gian 16 nhận từ bộ đoạt 14 dòng hoặc chuỗi 46 của N/F hệ số phổ 28 trên mỗi khung 36, tương ứng với lát tần số thấp ngoài ảnh phổ 26, được ghi theo phổ vào các hệ số phổ tần số thấp nhất được minh họa sử dụng chỉ số "0" trên Fig.3 và mở rộng cho đến các hệ số phổ của chỉ số $N/F - 1$.

Bộ điều biến phổ sang thời gian 16 đưa, đối với mỗi khung 36, đoạn tần số thấp 44 tương ứng của các hệ số phổ 28 tới phép biến đổi nghịch đảo 48 có các hàm điều biến có độ dài $(E+2) \cdot N/F$ kéo dài theo thời gian trên khung tương ứng và E+1 khung

đứng trước như được định họa tại 50 trên Fig.3, bằng cách đó thu được phần theo thời gian có độ dài $(E+2) \cdot N/F$, tức là phân đoạn thời gian chưa được tạo cửa sổ 52. Tức là, bộ điều biến phổ sang thời gian có thể thu được phân đoạn thời gian theo thời gian là $(E+2) \cdot N/F$ mẫu của tốc độ lấy mẫu giảm bởi hàm điều biến gán trọng số và tổng với cùng độ dài sử dụng, ví dụ, công thức thứ nhất của đoạn thay thế được đề xuất A.4 được biểu thị ở trên. N/F mẫu mới nhất của phân đoạn thời gian 52 thuộc về khung hiện thời 36. Các hàm điều biến có thể, như được biểu thị, ví dụ, là các hàm cosin trong trường hợp phép biến đổi nghịch đảo là MDCT nghịch đảo, hoặc các hàm sin trong trường hợp phép biến đổi nghịch đảo là MDCT nghịch đảo.

Do đó, bộ tạo cửa sổ 52 nhận, đối với từng khung, phần theo thời gian 52, N/F mẫu tại đầu dẫn trước của nó tương ứng theo thời gian với khung tương ứng trong khi các mẫu khác của phần theo thời gian 52 tương ứng thuộc về các khung đứng trước theo thời gian tương ứng. Bộ tạo cửa sổ 18 tạo cửa sổ, cho từng khung 36, phần theo thời gian 52 sử dụng cửa sổ tổng hợp đơn thức 54 có độ dài $(E+2) \cdot N/F$ bao gồm phần bằng 0 56 có độ dài $1/4 \cdot N/F$ tại đầu dẫn trước của nó, tức là $1/F \cdot N/F$ hệ số cửa sổ có giá trị bằng 0, và có đỉnh 58 trong khoảng theo thời gian kế tiếp, theo thời gian, phần bằng 0 56, tức là khoảng theo thời gian của phần theo thời gian 52 không được bao hàm bởi phần bằng không 52. Khoảng theo thời gian sau đó có thể được gọi là phần khác không của cửa sổ 58 và có độ dài $7/4 \cdot N/F$ được đo trong các mẫu có tốc độ lấy mẫu giảm, tức là $7/4 \cdot N/F$ hệ số cửa sổ. Bộ tạo cửa sổ 18 gán trọng số, ví dụ, phần theo thời gian 53, sử dụng cửa sổ 58. Phép gán trọng số hoặc phép nhân 58 này của từng phần theo thời gian với cửa sổ 54 dẫn đến phần theo thời gian được tạo cửa sổ 60, một phần cho từng khung 36 và trùng khớp với phần theo thời gian tương ứng 52 theo như độ phủ theo thời gian được quan tâm. Trong đoạn A.4 được đề xuất ở trên, việc xử lý phép tạo cửa sổ mà có thể được sử dụng bởi cửa sổ 18 được mô tả bởi công thức liên quan $z_{i,n}$ với $x_{i,n}$, trong đó $x_{i,n}$ tương ứng với các phần theo thời gian 52 được nêu trên chưa được tạo cửa sổ và $z_{i,n}$ tương ứng với các phần theo thời gian được tạo cửa sổ 60 với i biểu thị chuỗi khung/cửa sổ và n biểu thị, trong từng phần theo thời gian 52/60, các mẫu hoặc giá trị của các phần tương ứng 52/60 theo tốc độ lấy mẫu giảm.

Do đó, bộ khử răng cưa miền thời gian 20 nhận từ bộ tạo cửa sổ 18 chuỗi các phần theo thời gian được tạo cửa sổ 60, cụ thể là một chuỗi trên mỗi khung 36. Bộ khử 20 đưa các phần theo thời gian được tạo cửa sổ 60 của các khung 36 tới quy trình

chồng lắp-cộng 62 bằng cách ghi từng phần theo thời gian được tạo cửa sổ 60 với các giá trị N/F dẫn trước của nó để trùng khớp với khung 36 tương ứng. Bằng phép đo này, đoạn ở đầu kế tiếp có độ dài $(E + 1)/(E + 2)$ của phần theo thời gian được tạo cửa sổ 60 của khung hiện thời, tức là phần còn lại có độ dài $(E + 1) \cdot N/F$, chồng lắp với phần đầu bằng dài bằng tương ứng của phần theo thời gian của khung liền trước. Trong công thức, bộ khử răng cửa miền thời gian 20 có thể hoạt động như được thể hiện trong công thức cuối cùng của phiên bản được đề xuất ở trên của đoạn A.4, trong đó $out_{i,n}$ tương ứng với các mẫu âm thanh của tín hiệu âm thanh được khôi phục 22 tại tốc độ lấy mẫu giảm.

Quy trình tạo cửa sổ 58 và chồng lắp-cộng 62 như được thực hiện bởi bộ tạo cửa sổ 18 và bộ khử răng cửa miền thời gian 20 được minh họa chi tiết hơn dưới đây dựa vào Fig.4. Fig.4 sử dụng cả thuật ngữ được áp dụng trong đoạn A.4 được đề xuất ở trên và các ký hiệu tham chiếu được áp dụng trên Fig.3 và Fig.4. $x_{0,0}$ to $x_{0,(E+2) \cdot N/F - 1}$ biểu diễn phần theo phần thời gian thứ không 52 thu được bởi bộ điều biến thời gian sang phổ 16 cho khung thứ không 36. Chỉ số thứ nhất của x biểu thị các khung 36 dọc theo thứ tự theo thời gian, và chỉ số thứ hai của x chỉ dẫn các mẫu của khung dọc theo thứ tự theo thời gian, bước liên mẫu thuộc về tốc độ mẫu giảm. Sau đó, trên Fig.4, w_0 đến $w_{(E+2) \cdot N/F - 1}$ biểu thị các hệ số cửa sổ của cửa sổ 54. Tương tự chỉ số thứ hai của x , tức là phần theo thời gian 52 như được xuất ra bởi bộ điều biến 16, chỉ số của w là sao cho chỉ số 0 tương ứng với giá trị mẫu cũ nhất và chỉ số $(E+2) \cdot N/F - 1$ tương ứng với giá trị mẫu mới nhất khi cửa sổ 54 được áp dụng cho phần theo thời gian 52 tương ứng. Bộ tạo cửa sổ 18 tạo cửa sổ phần theo thời gian 52 sử dụng cửa sổ 54 để thu được phần theo thời gian được tạo cửa sổ 60 sao cho $z_{0,0}$ to $z_{0,(E+2) \cdot N/F - 1}$, mà biểu thị phần theo thời gian được tạo cửa sổ 60 cho khung thứ 0, thu được theo $z_{0,0} = x_{0,0} \cdot w_0, \dots, z_{0,(E+2) \cdot N/F - 1} = x_{0,(E+2) \cdot N/F - 1} \cdot w_{(E+2) \cdot N/F - 1}$. Các chỉ số của z có cùng nghĩa như đối với x . Theo cách này, bộ điều biến 16 và bộ tạo cửa sổ 18 có tác dụng cho từng khung được chỉ báo bởi chỉ số thứ nhất là x và z . Bộ khử 20 tính tổng $E + 2$ phần theo thời gian được tạo cửa sổ 60 của $E + 2$ khung liền nhau với độ lệch các mẫu của các phần theo thời gian được tạo cửa sổ 60 tương ứng với nhau bởi một khung, tức là bởi số lượng các mẫu trên mỗi khung 36, cụ thể là N/F , để thu được các mẫu u của một khung hiện thời, ở đây là $u_{-(E+1),0} \dots u_{-(E+1),N/F-1}$. Ngoài ra, ở đây, chỉ số thứ nhất của u biểu thị số lượng khung và chỉ số thứ hai chỉ dẫn các mẫu của khung này dọc theo thứ tự theo thời

gian. Bộ khử ghép các khung được khôi phục thu được sao cho các mẫu của tín hiệu âm thanh được khôi phục 22 trong các khung liên tiếp 36 theo sau nhau theo $u_{-(E+1),0} \dots u_{-(E+1),N/F-1}$, $u_{-E,0}$, $\dots$ $u_{-E,N/F-1}$, $u_{-(E-1),0}$, $\dots$. Bộ khử 22 tính toán từng mẫu của tín hiệu âm thanh 22 trong khung thứ $-(E+1)$ theo $u_{-(E+1),0} = Z_{0,0} + Z_{-1,N/F} + \dots Z_{-(E+1),(E+1) \cdot N/F}$, $\dots$, $u_{-(E+1),N/F-1} = Z_{0,N/F-1} + Z_{-1,2 \cdot N/F-1} + \dots + Z_{-(E+1),(E+2) \cdot N/F-1}$, tức là tính tổng $(e+2)$ số hạng trên các mẫu u của khung hiện thời.

Fig.5 minh họa sự khai thác có khả năng của thực tế là, trong số các mẫu được tạo cửa sổ đóng góp vào các mẫu âm thanh u của khung $-(E+1)$, một khung tương ứng với, hoặc đã được tạo cửa sổ sử dụng, phần bằng không 56 của cửa sổ 54, cụ thể là $Z_{-(E+1),(E+7/4) \cdot N/F} \dots Z_{-(E+1),(E+2) \cdot N/F-1}$ được định giá trị là không. Do đó, thay vì thu được tất cả N/F mẫu trong khung thứ $-(E+1)$ 36 của tín hiệu âm thanh u sử dụng $E+2$ số hạng, bộ khử 20 có thể tính toán đầu một phần tư dẫn trước của nó, cụ thể là $u_{-(E+1),(E+7/4) \cdot N/F} \dots u_{-(E+1),(E+2) \cdot N/F-1}$ chỉ sử dụng $E+1$ số hạng theo $u_{-(E+1),(E+7/4) \cdot N/F} = Z_{0,3/4 \cdot N/F} + Z_{-1,7/4 \cdot N/F} + \dots + Z_{-E,(E+3/4) \cdot N/F}$, $\dots$, $u_{-(E+1),(E+2) \cdot N/F-1} = Z_{0,N/F-1} + Z_{-1,2 \cdot N/F-1} + \dots + Z_{-E,(E+1) \cdot N/F-1}$. Theo cách này, bộ tạo cửa sổ thậm chí có thể bỏ sót, một cách hiệu quả, việc thực hiện của phép gán trọng số 58 đối với phần bằng không 56. Do đó, các mẫu $u_{-(E+1),(E+7/4) \cdot N/F} \dots u_{-(E+1),(E+2) \cdot N/F-1}$ của khung thứ $-(E+1)$ hiện thời, thu được sử dụng chỉ $E+1$ số hạng, trong khi $u_{-(E+1),(E+1) \cdot N/F} \dots u_{-(E+1),(E+7/4) \cdot N/F-1}$ sẽ thu được sử dụng $E+2$ số hạng.

Do đó, theo cách được nêu ở trên, bộ giải mã âm thanh 10 của Fig.2 tái tạo, theo cách được rút gọn, tín hiệu âm thanh được mã hóa thành dòng dữ liệu 24. Để đạt được mục đích này, bộ giải mã âm thanh 10 sử dụng hàm cửa sổ 54 mà bản thân nó là phiên bản được lấy mẫu giảm của cửa sổ tổng hợp tham chiếu có độ dài $(E+2) \cdot N$. Như được giải thích đối với Fig.6, phiên bản được lấy mẫu giảm này, tức là cửa sổ 54, thu được bằng cách lấy mẫu giảm cửa sổ tổng hợp tham chiếu bởi thừa số F , tức là thừa số lấy mẫu giảm, sử dụng phép nội suy theo phân đoạn, cụ thể là trong các phân đoạn có độ dài $1/4 \cdot N$ khi được đo trong chế độ chưa được rút gọn, trong các phân đoạn có độ dài $1/4 \cdot N/F$ trong chế độ được lấy mẫu giảm, trong các phân đoạn của các phần tư của độ dài khung của khung 36, được đo theo thời gian và được biểu diễn độc lập với tốc độ lấy mẫu. Trong $4 \cdot (E+2)$ phép nội suy, do đó, được thực hiện, do đó mang lại $4 \cdot (E+2)$ lần $1/4 \cdot N/F$ phân đoạn dài mà, được ghép, biểu diễn phiên bản được lấy mẫu giảm của cửa sổ tổng hợp tham chiếu có độ dài $(E+2) \cdot N$. Xem Fig.6 để minh họa. Fig.6 thể hiện

cửa sổ tổng hợp 54 mà là đơn thức và được sử dụng bởi bộ giải mã âm thanh 10 phù hợp với thủ tục giải mã âm thanh được lấy mẫu giảm dưới cửa sổ tổng hợp tham chiếu 70 mà cửa sổ này có độ dài $(E+2) \cdot N$. Tức là, bằng thủ tục lấy mẫu giảm 72 dẫn từ cửa sổ tổng hợp tham chiếu 70 đến cửa sổ tổng hợp 54 được sử dụng thực sự bởi bộ giải mã âm thanh 10 cho việc giải mã được lấy mẫu giảm, số lượng các hệ số cửa sổ được giảm bởi thừa số F . Trên Fig.6, thuật ngữ của Fig.5 và Fig.6 đã được giữ nguyên, tức là w được sử dụng để biểu thị cửa sổ phiên bản được lấy mẫu giảm 54, trong khi w' đã được sử dụng để biểu thị các hệ số cửa sổ của cửa sổ tổng hợp tham chiếu 70.

Như vừa được đề cập, để thực hiện phép lấy mẫu giảm 72, cửa sổ tổng hợp tham chiếu 70 được xử lý trong các phân đoạn 74 có cùng độ dài. Về số lượng, có $(E+2) \cdot 4$ phân đoạn 74 này. Được đo trong tốc độ lấy mẫu gốc, tức là số lượng các hệ số cửa sổ của cửa sổ tổng hợp tham chiếu 70, từng phân đoạn 74 là $1/4 \cdot N$ hệ số cửa sổ w' dài, và được đo trong tốc độ lấy mẫu giảm và tốc độ lấy mẫu được lấy mẫu giảm, từng phân đoạn 74 là $1/4 \cdot N/F$ hệ số cửa sổ w dài.

Đương nhiên là, có thể thực hiện phép lấy mẫu giảm 72 cho từng hệ số cửa sổ được lấy mẫu giảm w_i trùng khớp ngẫu nhiên với bất kỳ hệ số cửa sổ w'_j nào của cửa sổ tổng hợp tham chiếu 70 bằng cách thiết lập đơn giản $w_i = w'_j$ với thời gian mẫu của w_i trùng với thời gian mẫu của w'_j , và/hoặc bởi phép nội suy tuyến tính các hệ số cửa sổ w_i bất kỳ hiện có, theo thời gian, giữa hai hệ số cửa sổ w'_j và w'_{j+2} bởi phép nội suy tuyến tính, nhưng thủ tục này sẽ dẫn đến phép gần đúng kém của cửa sổ tổng hợp tham chiếu 70, tức là cửa sổ tổng hợp 54 được sử dụng bởi bộ giải mã âm thanh 10 để phép giải mã được lấy mẫu giảm sẽ biểu diễn phép gần đúng kém của cửa sổ tổng hợp tham chiếu 70, do đó không thỏa mãn yêu cầu đảm bảo sự kiểm tra phù hợp của phép giải mã được rút gọn tương ứng với phép giải mã không được rút gọn của tín hiệu âm thanh từ dòng dữ liệu 24. Do đó, phép lấy mẫu giảm 72 gồm thủ tục nội suy mà theo đó phần lớn hệ số w_i của cửa sổ được lấy mẫu giảm 54, cụ thể là các hệ số được định vị lệch khỏi biên của các phân đoạn 74, phụ thuộc vào cách của thủ tục lấy mẫu giảm 72 trên nhiều hơn hai hệ số cửa sổ w' của cửa sổ tham chiếu 70. Cụ thể, trong khi phần lớn hệ số cửa sổ w_i của cửa sổ được lấy mẫu giảm 54 phụ thuộc vào nhiều hơn hai hệ số cửa sổ w'_j của cửa sổ tham chiếu 70 để làm tăng chất lượng của kết quả nội suy/lấy mẫu giảm, tức là chất lượng gần đúng, cho mọi hệ số cửa sổ w_i của phiên bản được lấy

mẫu giảm 54 nó giữ đúng rằng hệ số cửa sổ w_i không phụ thuộc vào các hệ số cửa sổ w_j thuộc về các phân đoạn 74 khác nhau. Hơn nữa, thủ tục lấy mẫu giảm 72 là thủ tục nội suy theo phân đoạn.

Ví dụ, cửa sổ tổng hợp 54 có thể là sự nối chuỗi của các hàm nội trực có độ dài $1/4 \cdot N/F$. Các hàm nội trực bậc ba có thể được sử dụng. Ví dụ này đã được nêu ở trên trong đoạn A.1 trong đó vòng lặp for-next ngoài được lặp theo chuỗi qua các phân đoạn 74 trong đó, trong từng phân đoạn 74, phép lấy mẫu giảm hoặc phép nội suy 72 gồm tổ hợp toán học của các hệ số cửa sổ liên tiếp w' trong phân đoạn 74 hiện thời tại, ví dụ, hệ số thứ nhất cho mệnh đề tiếp theo trong đoạn “tính toán vector r cần để tính các hệ số c ”. Tuy nhiên, phép nội suy được áp dụng trong các phân đoạn có thể cũng được chọn theo cách khác nhau. Tức là, phép nội suy không bị giới hạn ở các hàm nội trực hoặc các hàm nội trực bậc ba. Hơn nữa, phép nội suy tuyến tính hoặc phương pháp nội suy khác bất kỳ cũng có thể được sử dụng. Trong trường hợp bất kỳ, việc thực hiện theo phân đoạn của phép nội suy sẽ dẫn đến việc tính toán các mẫu của cửa sổ tổng hợp được rút gọn, tức là các mẫu ngoài cùng của các phân đoạn của cửa sổ tổng hợp được rút gọn, lân cận phân đoạn khác, để không phụ thuộc vào các hệ số cửa sổ của cửa sổ tổng hợp tham chiếu hiện có trong các phân đoạn khác nhau.

Có thể là bộ tạo cửa sổ 18 thu được cửa sổ tổng hợp được lấy mẫu giảm 54 từ việc lưu trữ trong đó các hệ số cửa sổ w_i của cửa sổ tổng hợp được lấy mẫu giảm 54 này đã được lưu trữ sau khi đã thu được sử dụng phép lấy mẫu giảm 72. Ngoài ra, như được minh họa trên Fig.2, bộ giải mã âm thanh 10 có thể bao gồm bộ lấy mẫu giảm theo phân đoạn 76 thực hiện phép lấy mẫu giảm 72 của Fig.6 trên cơ sở cửa sổ tổng hợp tham chiếu 70.

Cần được lưu ý rằng bộ giải mã âm thanh 10 của Fig.2 có thể được tạo cấu hình để mang chỉ một thừa số lấy mẫu giảm cố định F hoặc có thể mang các giá trị khác. Trong trường hợp đó, bộ giải mã âm thanh 10 có thể tương ứng với giá trị đầu vào cho F như được minh họa trên Fig.2 tại 78. Bộ đo 14, ví dụ, có thể tương ứng với giá trị F này để đo được, như được đề cập ở trên, N/F giá trị phổ trên mỗi phổ của khung. Theo cách tương tự, bộ lấy mẫu giảm theo phân đoạn tùy chọn 76 cũng có thể tương ứng với giá trị F này hoạt động như được biểu thị ở trên. Bộ điều biến S/T 16 có thể tương ứng với F để, ví dụ, suy ra theo cách tính toán các phiên bản được rút gọn/được lấy mẫu giảm của các hàm điều biến, phiên bản được rút gọn/được lấy mẫu giảm

tương ứng với các phiên bản được sử dụng trong chế độ hoạt động không được rút gọn trong đó phép khôi phục dẫn đến tốc độ mẫu âm thanh đầy đủ.

Đương nhiên, bộ điều biến 16 cũng sẽ đáp ứng với đầu vào F 78, như bộ điều biến 16 sẽ sử dụng các phiên bản được lấy mẫu giảm thích hợp của các hàm điều biến và chúng giữ đúng cho bộ tạo cửa sổ 18 và bộ khử 20 đối với sự thích ứng của độ dài thực tế của các khung trong tốc độ lấy mẫu giảm hoặc tốc độ lấy mẫu được lấy mẫu giảm.

Ví dụ, F có thể nằm trong khoảng 1,5 và 10, bao gồm cả 1,5 và 10.

Cần được lưu ý rằng bộ giải mã của Fig.2 và Fig.3 hoặc phép điều biến bất kỳ của chúng được nêu ở đây, có thể được thực thi để thực hiện phép chuyển tiếp phổ thành thời gian sử dụng phép thực thi nâng của MDCT độ trễ thấp như được dẫn dắt trong, ví dụ, EP 2 378 516 B1.

Fig.8 minh họa việc thực thi của bộ giải mã sử dụng khái niệm nâng. Bộ điều biến S/T 16 thực hiện theo ví dụ DCT-IV nghịch đảo và được thể hiện như ở sau đây bởi các khối biểu diễn sự ghép chuỗi của bộ tạo cửa sổ 18 và bộ khử răng cưa miền thời gian 20. Trong ví dụ của Fig.8, E là 2, tức là E=2.

Bộ điều biến 16 bao gồm bộ chuyển đổi tần số/thời gian biến đổi cosin rời rạc loại-iv nghịch đảo. Thay vì xuất ra các chuỗi $(E+2)N/F$ phần theo thời gian dài 52 thì chỉ xuất ra các phần theo thời gian 52 có độ dài $2 \cdot N/F$, tất cả được suy ra từ chuỗi N/F phổ dài 46, các phần được rút ngắn 52 này tương ứng với nhân DCT, tức là $2 \cdot N/F$ mẫu mới nhất của các phần được mô tả trước.

Bộ tạo cửa sổ 18 có tác dụng như được mô tả trước và tạo ra phần theo thời gian được tạo cửa sổ 60 cho từng phần theo thời gian 52, nhưng nó chỉ hoạt động trên nhân DCT. Để đạt được mục đích này, bộ tạo cửa sổ 18 sử dụng hàm cửa sổ ω_i với $i=0 \dots 2N/F-1$, có kích thước nhân. Mỗi quan hệ giữa w_i với $i=0 \dots (E+2) \cdot N/F-1$ được mô tả sau đây, chỉ như là mối quan hệ giữa các hệ số nâng được đề cập sau đây và w_i với $i=0 \dots (E+2) \cdot N/F-1$.

Việc sử dụng thuật ngữ được áp dụng ở trên, quy trình được mô tả cho đến nay mang lại:

$$Z_{k,n} = \omega_n \cdot X_{k,n} \quad \text{với } n = 0, \dots, 2M-1,$$

với việc xác định lại $M = N/F$, sao cho M tương ứng với kích thước khung được biểu diễn trong miền được rút gọn và sử dụng thuật ngữ của các hình vẽ từ Fig.2 đến

Fig.6, tuy nhiên, $z_{k,n}$ và $x_{k,n}$ sẽ chỉ chứa các mẫu của phần theo thời gian được tạo cửa sổ và phần theo thời gian chưa được tạo cửa sổ trong nhân DCT có kích thước $2 \cdot M$ và tương ứng theo thời gian với các mẫu $E \cdot N/F \dots (E+2) \cdot N/F - 1$ trên Fig.4. Tức là, n là số nguyên biểu thị chỉ số mẫu và ω_n là hệ số hàm cửa sổ giá trị thực tương ứng với chỉ số mẫu n .

Quy trình chồng lấp/cộng của bộ khử 20 hoạt động theo cách khác so với mô tả ở trên. Quy trình này tạo ra các phần theo thời gian trung gian $m_k(0), \dots, m_k(M-1)$ dựa trên phương trình hoặc biểu thức

$$m_{k,n} = z_{k,n} + z_{k-1,n+M} \quad \text{với } n = 0, \dots, M-1$$

Trong việc thực thi của Fig.8, thiết bị còn bao gồm bộ nâng 80 mà có thể được nội suy như một phần của bộ điều biến 16 và bộ tạo cửa sổ 18 vì bộ nâng 80 bù cho việc là bộ điều biến và bộ tạo cửa sổ giới hạn việc xử lý của chúng tới nhân DCT thay vì xử lý sự mở rộng của các hàm điều biến và cửa sổ tổng hợp vượt quá nhân về phía trước mà sự mở rộng được đưa vào để bù cho phần bằng không 56. Bằng cách sử dụng khung làm việc của các bộ làm trễ và các bộ nhận 82 và các bộ cộng 84, bộ nâng 80 tạo ra các phần theo thời gian được khôi phục cuối cùng của các khung có độ dài M trong các cặp của các khung liền nhau dựa trên phương trình hoặc biểu thức

$$u_{k,n} = m_{k,n} + l_{n-M/2} \cdot m_{k-1,M-1-n} \quad \text{với } n = M/2, \dots, M-1,$$

và

$$u_{k,n} = m_{k,n} + l_{M-1-n} \cdot \text{out}_{k-1,M-1-n} \quad \text{với } n=0, \dots, M/2-1,$$

trong đó l_n với $n = 0 \dots M-1$ là hệ số có giá trị thực tương ứng với cửa sổ tổng hợp được rút gọn theo cách được mô tả chi tiết hơn dưới đây.

Nói cách khác, đối với sự chồng lấp được mở rộng của E khung vào trong các khung trước, chỉ M phép toán nhân-cộng bổ sung được yêu cầu, như có thể được thấy trong khung làm việc của các bộ lọc 80. Các phép toán này đôi khi cũng được tham chiếu đến như "các ma trận độ trễ bằng 0". Đôi khi, các phép toán này cũng được biết đến như "các bước nâng". Việc thực thi hiệu quả được thể hiện trên Fig.8 có thể dưới một số trường hợp hiệu quả hơn như việc thực thi đơn giản. Để rõ ràng hơn, phụ thuộc vào việc thực thi cụ thể, việc thực thi hiệu quả hơn này có thể dẫn đến tiết kiệm M phép toán, như trường hợp thực thi đơn giản cho M phép toán, có thể nên thực hiện, như việc thực thi được thể hiện trên Fig.9, yêu cầu theo nguyên tắc, $2M$ phép toán

trong khung làm việc của môđun 820 và M phép toán trong khung làm việc của bộ lọc 830.

Đối với sự phụ thuộc của ω_n với $n=0\dots 2M-1$ và l_n với $n = 0\dots M-1$ vào cửa sổ tổng hợp w_i với $i = 0\dots (E+2)M-1$ (được nhắc lại là ở đây $E=2$), công thức sau đây mô tả mối quan hệ giữa chúng với việc chuyển vị các chỉ số ký hiệu được sử dụng cho đến nay vào trong ngoặc đơn sau biến số tương ứng:

$$\begin{aligned}w(i) &= l\left(\frac{M}{2} - 1 - n\right) \cdot l(M - 1 - n) \cdot \omega(M + n) \\w\left(\frac{M}{2} + i\right) &= l(n) \cdot l\left(\frac{M}{2} + n\right) \cdot \omega\left(\frac{3M}{2} + n\right) \\w(M + i) &= l\left(\frac{M}{2} - 1 - n\right) \cdot \omega(M + n) \\w\left(\frac{3M}{2} + i\right) &= -l(n) \cdot \omega\left(\frac{3M}{2} + n\right) \\w(2M + i) &= -\omega(M + n) - l(M - 1 - n) \cdot \omega(n) \\w\left(\frac{5M}{2} + i\right) &= -\omega\left(\frac{3M}{2} + n\right) - l\left(\frac{M}{2} + n\right) \cdot \omega\left(\frac{M}{2} + n\right) \\w(3M + i) &= -\omega(n) \\w\left(\frac{7M}{2} + i\right) &= \omega(M + n) \\&\text{với } i, n = 0 \dots \frac{M}{2} - 1\end{aligned}$$

Lưu ý là cửa sổ w_i chứa các giá trị định trên phía bên phải trong công thức này, tức là giữa các chỉ số $2M$ và $4M - 1$. Công thức trên liên quan đến các hệ số l_n với $n = 0\dots M-1$ và ω_n $n = 0, \dots, 2M-1$ với các hệ số w_n với $n = 0\dots (E+2)M-1$ của cửa sổ tổng hợp được rút gọn. Như có thể thấy, l_n với $n = 0\dots M-1$ thực tế chỉ phụ thuộc vào $\frac{3}{4}$ hệ số của cửa sổ tổng hợp được rút gọn, cụ thể là phụ thuộc vào w_n với $n = 0\dots (E+1)M-1$, trong khi ω_n $n = 0, \dots, 2M-1$ phụ thuộc vào tất cả w_n với $n = 0\dots (E+2)M-1$.

Như được nêu trên, có thể là bộ tạo cửa sổ 18 thu được cửa sổ tổng hợp được lấy mẫu giảm 54 w_n với $n = 0\dots (E+2)M-1$ từ việc lưu trữ trong đó các hệ số cửa sổ w_i của cửa sổ tổng hợp được lấy mẫu giảm 54 này đã được lưu trữ sau khi đã thu được sử dụng phép lấy mẫu giảm 72, và trong đó các hệ số cửa sổ được đọc để tính toán các hệ số l_n với $n = 0\dots M-1$ và ω_n $n = 0, \dots, 2M-1$ sử dụng mối tương quan trên, nhưng theo cách khác, bộ tạo cửa sổ 18 có thể truy tìm các hệ số l_n với $n = 0 \dots M-1$ và ω_n $n = 0, \dots, 2M-1$, do đó được tính toán từ cửa sổ tổng hợp được lấy mẫu giảm trước, từ việc

lưu trữ một cách trực tiếp. Ngoài ra, như được nêu ở trên, bộ giải mã âm thanh 10 có thể bao gồm bộ lấy mẫu giảm theo phân đoạn 76 thực hiện phép lấy mẫu giảm 72 của Fig.6 trên cơ sở của cửa sổ tổng hợp tham chiếu 70, bằng cách đó thu được w_n với $n = 0 \dots (E+2)M-1$ trên cơ sở mà bộ tạo cửa sổ 18 tính toán các hệ số l_n với $n = 0 \dots M-1$ và ω_n $n = 0, \dots, 2M-1$ sử dụng mối tương quan/công thức nêu trên. Ngay cả khi sử dụng phép thực thi nâng, nhiều hơn một giá trị cho F có thể được chấp nhận.

Tóm tắt ngắn gọn phép thực thi nâng, cùng thu được bộ giải mã âm thanh 10 được tạo cấu hình để giải mã tín hiệu âm thanh 22 tại tốc độ lấy mẫu thứ nhất từ dòng dữ liệu 24 mà tín hiệu âm thanh được mã hóa biến đổi thành tại tốc độ lấy mẫu thứ hai, tốc độ lấy mẫu thứ nhất là $1/F$ tốc độ lấy mẫu thứ hai, bộ giải mã âm thanh 10 bao gồm bộ nhận 12 mà nhận, trên mỗi khung có độ dài N của tín hiệu âm thanh, N hệ số phổ 28, bộ đoạt 14 mà đoạt được cho mỗi khung, đoạn tần số thấp có độ dài N/F ngoài các hệ số phổ $N/28$, bộ điều biến phổ sang thời gian 16 được tạo cấu hình để đưa, đối với từng khung 36, đoạn tần số thấp tới phép biến đổi nghịch đảo có các hàm điều biến có độ dài $2 \cdot N/F$ kéo dài theo thời gian qua khung tương ứng và khung đứng trước để thu được phần theo thời gian có độ dài $2 \cdot N/F$, và bộ tạo cửa sổ 18 mà tạo cửa sổ, cho từng khung 36, phần theo thời gian $x_{k,n}$ theo $z_{k,n} = \omega_n \cdot x_{k,n}$ đối với $n = 0, \dots, 2M-1$ để thu được phần theo thời gian được tạo cửa sổ $z_{k,n}$ với $n = 0 \dots 2M-1$. Bộ khử răng cưa miền thời gian 20 tạo ra các phần theo thời gian trung gian $m_k(0), \dots, m_k(M-1)$ theo $m_{k,n} = z_{k,n} + z_{k-1,n+M}$ với $n = 0, \dots, M-1$. Cuối cùng, bộ nâng 80 tính toán các khung $u_{k,n}$ của tín hiệu âm thanh với $n = 0 \dots M-1$ theo $u_{k,n} = m_{k,n} + l_{n-M/2} \cdot m_{k-1,M-1-n}$ với $n = M/2, \dots, M-1$, và $u_{k,n} = m_{k,n} + l_{M-1-n} \cdot \text{out}_{k-1,M-1-n}$ với $n=0, \dots, M/2-1$, trong đó l_n với $n = 0 \dots M-1$ là các hệ số nâng trong đó phép biến đổi nghịch đảo là MDCT nghịch đảo hoặc MDST nghịch đảo, và trong đó l_n với $n = 0 \dots M-1$ và ω_n $n = 0, \dots, 2M-1$ phụ thuộc vào các hệ số w_n với $n = 0 \dots (E+2)M-1$ của cửa sổ tổng hợp, và cửa sổ tổng hợp là phiên bản được lấy mẫu giảm của cửa sổ tổng hợp tham chiếu có độ dài $4 \cdot N$, được lấy mẫu giảm bởi thừa số F bởi phép nội suy theo phân đoạn trong các phân đoạn có độ dài $1/4 \cdot N$.

Đã được nêu ra từ phần mô tả ở trên của sự đề xuất cho việc mở rộng của AAC-ELD đối với chế độ giải mã rút gọn mà bộ giải mã âm thanh của Fig.2 có thể được kèm theo với công cụ SBR độ trễ thấp. Ví dụ, phần sau đây nêu cách mà bộ mã hóa AAC-ELD mở rộng để hỗ trợ chế độ rút gọn được đề xuất ở trên, sẽ hoạt động khi sử

dụng công cụ SBR độ trễ thấp. Như đã được đề cập trong phần giới thiệu của bản mô tả sáng chế, trong trường hợp công cụ SBR độ trễ thấp được sử dụng kết hợp với bộ mã hóa AAC-ELD, các giàn lọc của môđun SBR độ trễ thấp cũng được rút gọn. Điều này đảm bảo rằng môđun SBR hoạt động với cùng độ phân giải tần số và do đó, không yêu cầu sự thích ứng hơn nào. Fig.7 tóm tắt đường truyền tín hiệu của bộ giải mã AAC-ELD hoạt động ở 96 kHz, với kích thước khung là 480 mẫu, trong chế độ SBR được lấy mẫu giảm và với thừa số rút gọn F là 2.

Trên Fig.7, dòng bit tới như được xử lý bởi chuỗi các khối, cụ thể là bộ giải mã AAC, khối LD-MDCT nghịch đảo, khối phân tích CLDFB, bộ giải mã SBR và khối tổng hợp giàn lọc độ trễ thấp phức tạp (complex low delay filter bank - CLDFB). Dòng bit bằng với dòng dữ liệu 24 được mô tả trước đối với các hình vẽ từ Fig.3 đến Fig.6, nhưng được kèm thêm bởi các dữ liệu SBR theo tham số hỗ trợ việc định dạng phổ của sự sao chép phổ của băng mở rộng phổ mở rộng tần số phổ của tín hiệu âm thanh thu được bằng việc giải mã âm thanh được rút gọn tại đầu ra của khối MDCT độ trễ thấp nghịch đảo, việc tạo hình phổ được thực hiện bởi bộ giải mã SBR. Cụ thể, bộ giải mã AAC tìm kiếm tất cả các phần tử cú pháp cần thiết bằng cách phân tích và giải mã entropy thích hợp. Bộ giải mã AAC có thể trùng khớp một phần với bộ nhận 12 của bộ giải mã âm thanh 10 mà, trên Fig.7, được thể hiện bởi khối MDCT độ trễ thấp nghịch đảo. Trên Fig.7, ví dụ là F bằng 2. Tức là, khối MDCT độ trễ thấp nghịch đảo của Fig.7 xuất ra như ví dụ cho tín hiệu âm thanh được khôi phục 22 của Fig.2, tín hiệu thời gian 48kHz được lấy mẫu giảm ở nửa tốc độ mà tại đó tín hiệu âm thanh được mã hóa trước tiên thành dòng bit tới. Khối phân tích CLDFB chia nhỏ tín hiệu thời gian 48-kHz này, tức là tín hiệu âm thanh thu được bởi phép giải mã âm thanh được rút gọn, thành N băng, ở đây $N = 16$, và bộ giải mã SBR tính toán các hệ số định dạng lại cho các băng này, theo đó định dạng lại N băng được điều khiển thông qua dữ liệu SBR trong dòng bit đầu vào đến đầu vào của bộ giải mã AAC, và khối tổng hợp CLDFB chuyển tiếp lại từ miền phổ sang miền thời gian với việc thu được, do đó, tín hiệu mở rộng tần số cao sẽ được thêm vào các tín hiệu âm thanh được giải mã ban đầu được xuất ra với khối MDCT độ trễ thấp nghịch đảo.

Lưu ý là sự hoạt động chuẩn của SBR sử dụng CLDFB 32 băng. Thuật toán nội suy cho các hệ số cửa sổ CLDFB 32 băng $c_{i_{32}}$ đã được cho trong 4.6.19.4.1 trong tài liệu tham khảo [1],

$$c_{i_{32}}(i) = \frac{1}{2}[c_{64}(2i+1) + c_{64}(2i)], \quad 0 \leq i < 320,$$

trong đó c_{64} là các hệ số cửa sổ của cửa sổ 64 băng đã cho trong bảng 4.A.90 trong tài liệu tham khảo [1]. Công thức này có thể còn được tổng quát hóa để cũng định nghĩa các hệ số cửa sổ cho số lượng băng ít hơn B

$$c_{i_B}(i) = \frac{1}{2}[c_{64}(2F \cdot i + 1 + p) + c_{64}(2F \cdot i + p)], \quad 0 \leq i < (10B), p = \text{int}\left(\frac{64}{2B} - 0.5\right)$$

trong đó F biểu thị thừa số rút gọn là $F = 32/B$. Với định nghĩa này về các hệ số cửa sổ, giàn lọc phân tích và tổng hợp CLDFB có thể hoàn toàn được mô tả như được nêu trong đoạn ví dụ A.2 nêu trên.

Do đó, các ví dụ trên cung cấp một số định nghĩa còn thiếu cho bộ mã hóa-giải mã AAC-ELD để làm thích ứng bộ mã hóa-giải mã với các hệ thống có tốc độ mẫu thấp hơn. Các định nghĩa này có thể được nêu trong tiêu chuẩn ISO/IEC 14496-3:2009.

Do đó, trong phần mô tả ở trên, không kể đến những điều khác, đã được mô tả là:

Bộ giải mã âm thanh có thể được tạo cấu hình để giải mã tín hiệu âm thanh tại tốc độ lấy mẫu thứ nhất từ dòng dữ liệu mà tín hiệu âm thanh được mã hóa biến đổi thành tại tốc độ lấy mẫu thứ hai, tốc độ lấy mẫu thứ nhất là $1/F$ tốc độ lấy mẫu thứ hai, bộ giải mã âm thanh bao gồm: bộ nhận được tạo cấu hình để nhận, trên mỗi khung có độ dài N của tín hiệu âm thanh, N hệ số phổ; bộ đoạt được tạo cấu hình để đoạt được cho từng khung, đoạn tần số thấp có độ dài N/F trong số N hệ số phổ; bộ điều biến phổ sang thời gian được tạo cấu hình để đưa, đối với từng khung, đoạn tần số thấp tới phép biến đổi nghịch đảo có các hàm điều biến có độ dài $(E+2) \cdot N/F$ kéo dài theo thời gian qua khung tương ứng và $E+1$ khung đứng trước để thu được phần theo thời gian có độ dài $(E+2) \cdot N/F$; bộ tạo cửa sổ được tạo cấu hình để tạo cửa sổ, cho từng khung, phần theo thời gian sử dụng cửa sổ tổng hợp đơn thức có độ dài $(E+2) \cdot N/F$ bao gồm phần bằng 0 có độ dài $1/4 \cdot N/F$ tại đầu dẫn trước của nó và có đỉnh trong khoảng theo thời gian của cửa sổ tổng hợp đơn thức, khoảng theo thời gian nối tiếp phần bằng 0 và có

độ dài $7/4 \cdot N/F$ để bộ tạo cửa sổ thu được phần theo thời gian được tạo cửa sổ có độ dài $(E+2) \cdot N/F$; và bộ khử răng cưa miền thời gian được tạo cấu hình để đưa phần theo thời gian được tạo cửa sổ của các khung sang quy trình chồng lấp và cộng để đoạn ở đầu kế tiếp có độ dài $(E+1)/(E+2)$ của phần theo thời gian được tạo cửa sổ của khung hiện thời chồng lấp đầu dẫn trước có độ dài $(E+1)/(E+2)$ của phần theo thời gian được tạo cửa sổ của khung đứng trước, trong đó phép biến đổi nghịch đảo là MDCT nghịch đảo hoặc MDST nghịch đảo, và trong đó cửa sổ tổng hợp đơn thức là phiên bản được lấy mẫu giảm của cửa sổ tổng hợp đơn thức tham chiếu có độ dài $(E+2) \cdot N$, được tạo cửa sổ bởi thừa số F bằng phép nội suy theo phân đoạn trong các phân đoạn có độ dài $1/4 \cdot N/F$.

Bộ giải mã âm thanh theo phương án, trong đó cửa sổ tổng hợp đơn thức là sự nối chuỗi của các hàm nối trục có độ dài $1/4 \cdot N/F$.

Bộ giải mã âm thanh theo phương án, trong đó cửa sổ tổng hợp đơn thức là sự nối chuỗi của các hàm nối trục bậc ba có độ dài $1/4 \cdot N/F$.

Bộ giải mã âm thanh theo phương án bất kỳ trong số các phương án nêu trên, trong đó $E=2$.

Bộ giải mã âm thanh theo phương án bất kỳ trong số các phương án nêu trên, trong đó phép biến đổi nghịch đảo là MDCT nghịch đảo.

Bộ giải mã âm thanh theo phương án bất kỳ trong số các phương án nêu trên, trong đó nhiều hơn 80% khối lượng cửa sổ tổng hợp đơn thức được chứa trong khoảng theo thời gian nối tiếp phần bằng 0 và có độ dài $7/4 \cdot N/F$.

Bộ giải mã âm thanh theo phương án bất kỳ trong số các phương án nêu trên, trong đó bộ giải mã âm thanh được tạo cấu hình để thực hiện phép nội suy hoặc để suy ra cửa sổ tổng hợp đơn thức từ việc lưu trữ.

Bộ giải mã âm thanh theo phương án bất kỳ trong số các phương án nêu trên, trong đó bộ giải mã âm thanh được tạo cấu hình để mang các giá trị khác nhau cho F .

Bộ giải mã âm thanh theo phương án bất kỳ trong số các phương án nêu trên, trong đó F nằm trong khoảng 1,5 và 10, bao gồm cả 1,5 và 10.

Phương pháp được thực hiện bởi bộ giải mã âm thanh theo phương án bất kỳ trong số các phương án nêu trên.

Chương trình máy tính có mã chương trình để thực hiện, khi chạy trên máy tính, phương pháp theo phương án.

Xét đến thuật ngữ "có chiều dài ...", cần lưu ý rằng thuật ngữ này được hiểu là phép đo chiều dài của mẫu. Xét đến độ dài của phần bằng 0 và các phân đoạn, cần lưu ý rằng độ dài có thể có giá trị nguyên. Ngoài ra, độ dài có thể là giá trị không nguyên.

Đối với khoảng theo thời gian mà trong đó đỉnh được định vị trí, được lưu ý rằng Fig.1 thể hiện đỉnh này cũng như khoảng theo thời gian để minh họa cho ví dụ của cửa sổ tổng hợp đơn thức tham chiếu với $E=2$ và $N=512$: đỉnh có cực đại của nó ở mẫu gần đúng số 1408 và khoảng theo thời gian kéo dài từ mẫu số 1024 đến mẫu số 1920. Do đó, khoảng theo thời gian là $7/8$ của độ dài nhân DCT.

Đối với thuật ngữ "phiên bản được lấy mẫu giảm", được lưu ý rằng trong phần mô tả ở trên, thay vì thuật ngữ này thì có thể sử dụng thuật ngữ "phiên bản được rút gọn" để thay thế.

Đối với thuật ngữ "khối lượng của hàm trong khoảng nhất định", được lưu ý là khối lượng của hàm sẽ biểu thị tích phân xác định của hàm tương ứng trong khoảng tương ứng.

Trong trường hợp bộ giải mã âm thanh mang các giá trị khác nhau cho F , bộ giải mã âm thanh có thể bao gồm phần lưu trữ có các phiên bản được nội suy theo phân đoạn của cửa sổ tổng hợp đơn thức tham chiếu hoặc có thể thực hiện phép nội suy theo phân đoạn cho giá trị hoạt động hiện thời của F . Các phiên bản được nội suy theo phân đoạn khác nhau có điểm chung là phép nội suy không ảnh hưởng tiêu cực đến tính gián đoạn tại các đường biên phân đoạn. Chúng có thể, như được mô tả ở trên, là các hàm nối trục.

Bằng cách suy ra cửa sổ tổng hợp đơn thức bởi phép nội suy theo phân đoạn từ cửa sổ tổng hợp đơn thức tham chiếu như cửa sổ được thể hiện trên Fig.1 ở trên, các phân đoạn $4 \cdot (E+2)$ có thể được thực hiện bởi phép gần đúng của hàm nối trục như bởi các hàm nối trục bậc ba và bất kể phép nội suy, tính không liên tiếp mà có mặt trong cửa sổ tổng hợp đơn thức tại bước $1/4 \cdot N/F$ do phần bằng 0 được đưa vào theo cách tổng hợp như phương tiện làm giảm độ trễ được bảo toàn.

Các phương án khác

Phương án 1. Bộ giải mã âm thanh (10) được tạo cấu hình để giải mã tín hiệu âm thanh (22) ở tốc độ lấy mẫu thứ nhất từ dòng dữ liệu (24) mà tín hiệu âm thanh được mã hóa biến đổi thành ở tốc độ lấy mẫu thứ hai, tốc độ lấy mẫu thứ nhất bằng $1/F$ tốc độ lấy mẫu thứ hai, bộ giải mã âm thanh (10) bao gồm:

bộ nhận (12) được tạo cấu hình để nhận N hệ số phổ (28) trên mỗi khung có độ dài N của tín hiệu âm thanh;

bộ đoạt (14) được tạo cấu hình để đoạt được đoạn có tần số thấp có độ dài N/F trong số N hệ số phổ (28) cho từng khung;

bộ điều biến phổ sang thời gian (16) được tạo cấu hình để đưa, đối với từng khung (36), đoạn có tần số thấp sang phép biến đổi nghịch đảo có các hàm điều biến có độ dài $(E+2) \cdot N/F$ kéo dài theo thời gian qua khung tương ứng và $E+1$ khung trước để thu được phần theo thời gian có độ dài $(E+2) \cdot N/F$;

bộ tạo cửa sổ (18) được tạo cấu hình để tạo cửa sổ, cho từng khung (36), phần theo thời gian sử dụng cửa sổ tổng hợp có độ dài $(E+2) \cdot N/F$ bao gồm phần bằng 0 có độ dài $1/4 \cdot N/F$ tại đầu dẫn trước của nó và có đỉnh nằm trong khoảng theo thời gian của cửa sổ tổng hợp, khoảng theo thời gian nối tiếp phần bằng 0 và có độ dài $7/4 \cdot N/F$ sao cho bộ tạo cửa sổ thu được phần theo thời gian được tạo cửa sổ có độ dài $(E+2) \cdot N/F$; và

bộ khử răng cửa miền thời gian (20) được tạo cấu hình để đưa phần theo thời gian được tạo cửa sổ của các khung sang quy trình chồng lấp và cộng sao cho đoạn ở đầu kế tiếp có độ dài $(E+1)/(E+2)$ của phần theo thời gian được tạo cửa sổ của khung hiện thời chồng lấp đầu dẫn trước có độ dài $(E+1)/(E+2)$ của phần theo thời gian được tạo cửa sổ của khung đứng trước,

trong đó phép biến đổi nghịch đảo là MDCT nghịch đảo hoặc MDST nghịch đảo, và

trong đó cửa sổ tổng hợp là phiên bản được lấy mẫu giảm của cửa sổ tổng hợp tham chiếu có độ dài $(E+2) \cdot N$, được lấy mẫu giảm bởi thừa số F bởi phép nội suy theo phân đoạn trong các phân đoạn có độ dài $1/4 \cdot N$.

Phương án 2. Bộ giải mã âm thanh (10) theo điểm 1, trong đó cửa sổ tổng hợp là sự nối chuỗi của các hàm nối trục có độ dài $1/4 \cdot N/F$.

Phương án 3. Bộ giải mã âm thanh (10) theo điểm 1 hoặc điểm 2, trong đó cửa sổ tổng hợp là sự nối chuỗi của các hàm nối trục bậc ba có độ dài $1/4 \cdot N/F$.

Phương án 4. Bộ giải mã âm thanh (10) theo điểm bất kì trong số các điểm nêu trên, trong đó $E=2$.

Phương án 5. Bộ giải mã âm thanh (10) theo điểm bất kì trong số các điểm nêu trên, trong đó phép biến đổi nghịch đảo là MDCT nghịch đảo.

Phương án 6. Bộ giải mã âm thanh (10) theo điểm bất kì trong số các điểm nêu trên, trong đó nhiều hơn 80% khối lượng của số tổng hợp được chứa trong khoảng theo thời gian nối tiếp phần bằng 0 và có độ dài $7/4 \cdot N/F$.

Phương án 7. Bộ giải mã âm thanh (10) theo điểm bất kì trong số các điểm nêu trên, trong đó bộ giải mã âm thanh (10) được tạo cấu hình để thực hiện phép nội suy hoặc để suy ra cửa sổ tổng hợp từ việc lưu trữ.

Phương án 8. Bộ giải mã âm thanh (10) theo điểm bất kì trong số các điểm nêu trên, trong đó bộ giải mã âm thanh (10) được tạo cấu hình để mang các giá trị khác nhau cho F.

Phương án 9. Bộ giải mã âm thanh (10) theo điểm bất kì trong số các điểm nêu trên, trong đó F nằm trong khoảng giữa 1,5 và 10, bao gồm cả 1,5 và 10.

Phương án 10. Bộ giải mã âm thanh (10) theo điểm bất kì trong số các điểm nêu trên, trong đó cửa sổ tổng hợp tham chiếu là đơn thức.

Phương án 11. Bộ giải mã âm thanh (10) theo điểm bất kì trong số các điểm nêu trên, trong đó bộ giải mã âm thanh (10) được tạo cấu hình để thực hiện phép nội suy theo cách mà phần lớn các hệ số của cửa sổ tổng hợp phụ thuộc vào nhiều hơn hai hệ số của cửa sổ tổng hợp tham chiếu.

Phương án 12. Bộ giải mã âm thanh (10) theo điểm bất kì trong số các điểm nêu trên, trong đó bộ giải mã âm thanh (10) được tạo cấu hình để thực hiện phép nội suy theo cách mà từng hệ số của cửa sổ tổng hợp được phân tách bởi nhiều hơn hai hệ số từ các đường biên phân đoạn phụ thuộc vào nhiều hơn hai hệ số của cửa sổ tổng hợp tham chiếu.

Phương án 13. Bộ giải mã âm thanh (10) theo điểm bất kì trong số các điểm nêu trên, trong đó bộ tạo cửa sổ (18) và bộ khử răng cưa miền thời gian kết hợp sao cho bộ tạo cửa sổ bỏ qua phần bằng 0 trong việc gán trọng số phần theo thời gian sử dụng cửa sổ tổng hợp và bộ khử răng cưa miền thời gian (20) bỏ qua phần không được gán trọng số tương ứng của phần theo thời gian được tạo cửa sổ trong quy trình chồng lấp và cộng sao cho chỉ E+1 phần theo thời gian được tạo cửa sổ được tính tổng để dẫn đến phần không được gán trọng số tương ứng của khung tương ứng và E+2 phần được tạo cửa sổ được tính tổng trong trình nhấc của khung tương ứng.

Phương án 14. Bộ giải mã âm thanh để tạo ra phiên bản được rút gọn của cửa sổ tổng hợp của bộ giải mã âm thanh (10) theo điểm bất kì trong số các điểm nêu trên,

trong đó $E=2$ để hàm cửa sổ tổng hợp bao gồm nhân liên quan đến nửa độ dài $2 \cdot N/F$ được đặt trước bởi trình nhấc nửa độ dài $2 \cdot N/F$ và trong đó bộ điều biến phổ thành thời gian (16), bộ tạo cửa sổ (18) và bộ khử răng cưa miền thời gian (20) được thực thi để kết hợp trong phép thực thi nâng mà theo đó

bộ điều biến phổ thành thời gian (16) hạn chế việc đưa, đối với từng khung (36), đoạn có tần số thấp sang phép biến đổi nghịch đảo có các hàm điều biến có độ dài $(E+2) \cdot N/F$ kéo dài theo thời gian qua khung tương ứng và $E+1$ khung đứng trước, tới nhân biến đổi trùng khớp với khung tương ứng và một khung đứng trước để thu được phần theo thời gian $x_{k,n}$ với $n = 0 \dots 2M-1$ với $M=N/F$ là chỉ số mẫu và k là chỉ số khung;

bộ tạo cửa sổ (18) tạo cửa sổ, cho từng khung (36), phần theo thời gian $x_{k,n}$ theo $z_{k,n} = \omega_n \cdot x_{k,n}$ với $n = 0, \dots, 2M-1$ để thu được phần theo thời gian được tạo cửa sổ $z_{k,n}$ với $n = 0 \dots 2M-1$;

bộ khử răng cưa miền thời gian (20) tạo ra các phần theo thời gian trung gian $m_k(0), \dots, m_k(M-1)$ theo $m_{k,n} = z_{k,n} + z_{k-1,n+M}$ với $n = 0, \dots, M-1$, và

bộ giải mã âm thanh bao gồm bộ nâng (80) được tạo cấu hình để thu được các khung $u_{k,n}$ với $n = 0 \dots M-1$ theo

$$u_{k,n} = m_{k,n} + l_{n-M/2} \cdot m_{k-1,M-1-n} \quad \text{với } n = M/2, \dots, M-1,$$

và

$$u_{k,n} = m_{k,n} + l_{M-1-n} \cdot \text{out}_{k-1,M-1-n} \quad \text{với } n=0, \dots, M/2-1,$$

trong đó l_n với $n = 0 \dots M-1$ là các hệ số nâng, và trong đó l_n với $n = 0 \dots M-1$ và ω_n với $n = 0, \dots, 2M-1$ phụ thuộc vào các hệ số w_n với $n = 0 \dots (E+2)M-1$ của cửa sổ tổng hợp.

Phương án 15. Bộ giải mã âm thanh (10) được tạo cấu hình để giải mã tín hiệu âm thanh (22) với tốc độ lấy mẫu thứ nhất từ dòng dữ liệu (24) mà tín hiệu âm thanh được mã hóa biến đổi thành ở tốc độ lấy mẫu thứ hai, tốc độ lấy mẫu thứ nhất bằng $1/F$ tốc độ lấy mẫu thứ hai, bộ giải mã âm thanh (10) bao gồm:

bộ nhận (12) được tạo cấu hình để nhận N hệ số phổ (28) trên mỗi khung có độ dài N của tín hiệu âm thanh;

bộ đoạt (14) được tạo cấu hình để đoạt được đoạn có tần số thấp có độ dài N/F trong số N hệ số phổ (28) cho từng khung;

bộ điều biến phổ thành thời gian (16) được tạo cấu hình để đưa, đối với từng khung (36), đoạn tần số thấp sang phép biến đổi nghịch đảo có các hàm điều biến có độ dài $2 \cdot N/F$ kéo dài theo thời gian qua khung tương ứng và khung đứng trước để thu được phân theo thời gian có độ dài $2 \cdot N/F$;

bộ tạo cửa sổ (18) được tạo cấu hình để tạo cửa sổ, cho từng khung (36), phân theo thời gian $x_{k,n}$ theo $z_{k,n} = \omega_n \cdot x_{k,n}$ với $n = 0, \dots, 2M-1$ để thu được phân theo thời gian được tạo cửa sổ $z_{k,n}$ với $n = 0 \dots 2M-1$;

bộ khử răng cưa miền thời gian (20) được tạo cấu hình để tạo ra các phân theo thời gian trung gian $m_k(0), \dots, m_k(M-1)$ theo $m_{k,n} = z_{k,n} + z_{k-1,n+M}$ với $n = 0, \dots, M-1$, và

bộ nâng (80) được tạo cấu hình để thu được các khung $u_{k,n}$ của tín hiệu âm thanh với $n = 0 \dots M-1$ theo

$$u_{k,n} = m_{k,n} + l_{n-M/2} \cdot m_{k-1,M-1-n} \quad \text{với } n = M/2, \dots, M-1,$$

và

$$u_{k,n} = m_{k,n} + l_{M-1-n} \cdot \text{out}_{k-1,M-1-n} \quad \text{với } n=0, \dots, M/2-1,$$

trong đó l_n với $n = 0 \dots M-1$ là các hệ số nâng,

trong đó phép biến đổi nghịch đảo là MDCT nghịch đảo hoặc MDST nghịch đảo, và

trong đó l_n với $n = 0 \dots M-1$ và ω_n với $n = 0, \dots, 2M-1$ phụ thuộc vào các hệ số w_n với $n = 0 \dots (E+2)M-1$ của cửa sổ tổng hợp, và cửa sổ tổng hợp là phiên bản được lấy mẫu giảm của cửa sổ tổng hợp tham chiếu có độ dài $4 \cdot N$, được lấy mẫu giảm bởi thừa số F bởi phép nội suy theo phân đoạn trong các phân đoạn có độ dài $1/4 \cdot N$.

Phương án 16. Thiết bị tạo ra phiên bản được rút gọn của cửa sổ tổng hợp của bộ giải mã âm thanh (10) theo điểm bất kì trong số các điểm nêu trên, trong đó thiết bị được tạo cấu hình để lấy mẫu giảm cửa sổ tổng hợp tham chiếu có độ dài $(E+2) \cdot N$ bởi thừa số F bởi phép nội suy theo phân đoạn trong $4 \cdot (E+2)$ phân đoạn có cùng độ dài.

Phương án 17. Phương pháp tạo ra phiên bản được rút gọn của cửa sổ tổng hợp của bộ giải mã âm thanh (10) theo điểm bất kì trong số các điểm nêu trên, trong đó phương pháp bao gồm bước lấy mẫu giảm cửa sổ tổng hợp tham chiếu có độ dài $(E+2) \cdot N$ bởi thừa số F bởi phép nội suy theo phân đoạn trong $4 \cdot (E+2)$ phân đoạn có cùng độ dài.

Phương án 18. Phương pháp giải mã tín hiệu âm thanh (22) với tốc độ lấy mẫu thứ nhất từ dòng dữ liệu (24) mà tín hiệu âm thanh mà được mã hóa biến đổi thành ở tốc độ lấy mẫu thứ hai, tốc độ lấy mẫu thứ nhất bằng $1/F$ tốc độ lấy mẫu thứ hai, phương pháp bao gồm:

nhận, trên mỗi khung có độ dài N của tín hiệu âm thanh, N hệ số phổ (28);

đoạt được đoạn có tần số thấp có độ dài N/F trong số N hệ số phổ (28) cho từng khung;

thực hiện phép điều biến phổ sang thời gian bằng cách đưa, đối với từng khung (36), đoạn có tần số thấp sang phép biến đổi nghịch đảo có các hàm điều biến có độ dài $(E+2) \cdot N/F$ kéo dài theo thời gian qua khung tương ứng và $E+1$ khung đứng trước để thu được phần theo thời gian có độ dài $(E+2) \cdot N/F$;

tạo cửa sổ, cho từng khung (36), phần theo thời gian sử dụng cửa sổ tổng hợp có độ dài $(E+2) \cdot N/F$ bao gồm phần bằng 0 có độ dài $1/4 \cdot N/F$ tại đầu dẫn trước của nó và có đỉnh nằm trong khoảng theo thời gian của cửa sổ tổng hợp, khoảng theo thời gian nối tiếp phần bằng 0 và có độ dài $7/4 \cdot N/F$ sao cho bộ tạo cửa sổ thu được phần theo thời gian được tạo cửa sổ có độ dài $(E+2) \cdot N/F$; và

thực hiện việc khử răng cưa miền thời gian bằng cách đưa phần theo thời gian được tạo cửa sổ của các khung sang quy trình chồng lấp và cộng sao cho đoạn ở đầu kế tiếp có độ dài $(E+1)/(E+2)$ của phần theo thời gian được tạo cửa sổ của khung hiện thời chồng lấp đầu dẫn trước có độ dài $(E+1)/(E+2)$ của phần theo thời gian được tạo cửa sổ của khung đứng trước,

trong đó phép biến đổi nghịch đảo là MDCT nghịch đảo hoặc MDST nghịch đảo, và

trong đó cửa sổ tổng hợp là phiên bản được lấy mẫu giảm của cửa sổ tổng hợp tham chiếu có độ dài $(E+2) \cdot N$, được lấy mẫu giảm bởi thừa số F bởi phép nội suy theo phân đoạn trong phân đoạn có độ dài $1/4 \cdot N$.

Phương án 19. Chương trình máy tính có mã chương trình để thực hiện, khi chạy trên máy tính, phương pháp theo phương án 17 hoặc phương án 18.

YÊU CẦU BẢO HỘ

1. Bộ giải mã âm thanh bao gồm:

bộ nhận được tạo cấu hình để nhận, đối với từng khung trong số các khung của tín hiệu âm thanh, phổ mà tạo thành sự khai triển phổ của phần theo thời gian bao gồm khung tương ứng và $N-1$ khung đứng trước, với N là số nguyên, trong đó bộ nhận được tạo cấu hình để thực hiện điền đầy khoảng trống để điền đầy một cách tổng hợp các phần được lượng tử hóa bằng 0 trong phổ;

bộ đoạt, được tạo cấu hình để đoạt, cho từng khung, phần tần số thấp bằng $1/F$, theo độ dài của phổ;

bộ điều biến phổ thành thời gian được tạo cấu hình để đưa, đối với từng khung, phần tần số thấp sang phép biến đổi nghịch đảo để thu được sự biểu diễn theo thời gian của phần theo thời gian;

bộ tạo cửa sổ được tạo cấu hình để tạo cửa sổ, cho từng khung, sự biểu diễn theo thời gian của phần theo thời gian sử dụng cửa sổ tổng hợp bao gồm phần bằng 0 tại đầu dẫn trước của nó và có đỉnh nằm trong khoảng theo thời gian của cửa sổ tổng hợp, mà nối tiếp phần bằng 0, sao cho bộ tạo cửa sổ thu được sự biểu diễn theo thời gian được tạo cửa sổ của phần theo thời gian; và

bộ khử răng cưa miền thời gian được tạo cấu hình để đưa sự biểu diễn theo thời gian được tạo cửa sổ của phần theo thời gian của các khung sang quy trình chồng lấp-cộng ở khoảng cách liên khung với nhau tương ứng với độ dài khung,

trong đó phép biến đổi nghịch đảo là biến đổi cosin rời rạc cải biên (modified discrete cosine transform – MDCT) nghịch đảo hoặc biến đổi sin rời rạc cải biên (modified discrete sine transform – MDST) nghịch đảo, và

trong đó cửa sổ tổng hợp là phiên bản được lấy mẫu giảm của cửa sổ tổng hợp tham chiếu, được lấy mẫu giảm bởi thừa số F bởi phép nội suy theo phân đoạn trong $4 \cdot N$ phân đoạn có độ dài phân đoạn bằng nhau.

2. Bộ giải mã âm thanh theo điểm 1, trong đó cửa sổ tổng hợp là sự nối chuỗi của hàm nối trục cho từng phân đoạn trong số $4 \cdot N$ phân đoạn.

3. Bộ giải mã âm thanh theo điểm 1, trong đó cửa sổ tổng hợp là sự nối chuỗi của một hàm nối trục cho từng phân đoạn trong số $4 \cdot N$ phân đoạn.

4. Bộ giải mã âm thanh theo điểm 1, trong đó $N = 4$.

5. Bộ giải mã âm thanh theo điểm 1, trong đó phép biến đổi nghịch đảo là MDCT nghịch đảo.
6. Bộ giải mã âm thanh theo điểm 1, trong đó hơn 80% khối lượng của cửa sổ tổng hợp được bao gồm trong khoảng theo thời gian tiếp theo phần bằng 0 và khoảng theo thời gian tiếp theo phần bằng 0 bằng $7/4$ lần chiều dài khung.
7. Bộ giải mã âm thanh theo điểm 1, trong đó bộ giải mã âm thanh được tạo cấu hình để thực hiện phép nội suy hoặc để suy ra cửa sổ tổng hợp từ việc lưu trữ.
8. Bộ giải mã âm thanh theo điểm 1, trong đó bộ giải mã âm thanh được tạo cấu hình để mang các giá trị khác nhau cho F .
9. Bộ giải mã âm thanh theo điểm 1, trong đó F nằm trong khoảng giữa 1,5 và 10, bao gồm cả 1,5 và 10.
10. Bộ giải mã âm thanh theo điểm 1, trong đó cửa sổ tổng hợp tham chiếu là đơn thức.
11. Bộ giải mã âm thanh theo điểm 1, trong đó bộ giải mã âm thanh được tạo cấu hình để thực hiện phép nội suy theo cách mà phần lớn các hệ số của cửa sổ tổng hợp phụ thuộc vào nhiều hơn hai hệ số của cửa sổ tổng hợp tham chiếu.
12. Bộ giải mã âm thanh theo điểm 1, trong đó bộ giải mã âm thanh được tạo cấu hình để thực hiện phép nội suy theo cách mà từng hệ số của cửa sổ tổng hợp được phân tách bởi nhiều hơn hai hệ số từ các đường biên phân đoạn phụ thuộc vào nhiều hơn hai hệ số của cửa sổ tổng hợp tham chiếu.
13. Bộ giải mã âm thanh theo điểm 1, trong đó bộ tạo cửa sổ và bộ khử răng cưa miền thời gian kết hợp sao cho bộ tạo cửa sổ bỏ qua phần bằng 0 trong việc gán trọng số phần theo thời gian bằng cách sử dụng cửa sổ tổng hợp và bộ khử răng cưa miền thời gian bỏ qua phần không được gán trọng số tương ứng của phần theo thời gian được tạo cửa sổ trong quy trình chồng lấp-cộng.
14. Phương pháp giải mã tín hiệu âm thanh, phương pháp bao gồm các bước:
nhận, đối với từng khung trong số các khung của tín hiệu âm thanh, phổ mà tạo thành sự khai triển phổ của phần theo thời gian bao gồm khung tương ứng và $N-1$ khung đứng trước, với N là số nguyên;
thực hiện điền đầy khoảng trống để điền đầy một cách tổng hợp các phần được lượng tử hóa bằng 0 trong phổ;
đoạt cho từng khung, phần tần số thấp bằng $1/F$, theo độ dài, của phổ;

thực hiện điều biến phổ thành thời gian bằng cách đưa, đối với từng khung, phần tần số thấp sang phép biến đổi nghịch đảo để thu được sự biểu diễn theo thời gian của phần theo thời gian;

tạo cửa sổ, cho từng khung, sự biểu diễn theo thời gian của phần theo thời gian sử dụng cửa sổ tổng hợp bao gồm phần bằng 0 tại đầu dẫn trước của nó và có đỉnh nằm trong khoảng theo thời gian của cửa sổ tổng hợp, mà nối tiếp phần bằng 0, để thu được sự biểu diễn theo thời gian được tạo cửa sổ của phần theo thời gian; và

thực hiện khử răng cưa miền thời gian bằng cách đưa sự biểu diễn theo thời gian được tạo cửa sổ của phần theo thời gian của các khung sang quy trình chồng lấp-cộng ở khoảng cách liên khung với nhau tương ứng với độ dài khung,

trong đó phép biến đổi nghịch đảo là MDCT nghịch đảo hoặc MDST nghịch đảo, và

trong đó cửa sổ tổng hợp là phiên bản được lấy mẫu giảm của cửa sổ tổng hợp tham chiếu, được lấy mẫu giảm bởi thừa số F bởi phép nội suy theo phân đoạn trong $4 \cdot N$ phân đoạn có độ dài phân đoạn bằng nhau.

15. Vật ghi lưu trữ số không chuyển tiếp đã lưu trữ trên đó chương trình máy tính để thực hiện phương pháp giải mã tín hiệu âm thanh theo điểm 14,

khi chương trình máy tính đã nêu được chạy bởi máy tính.

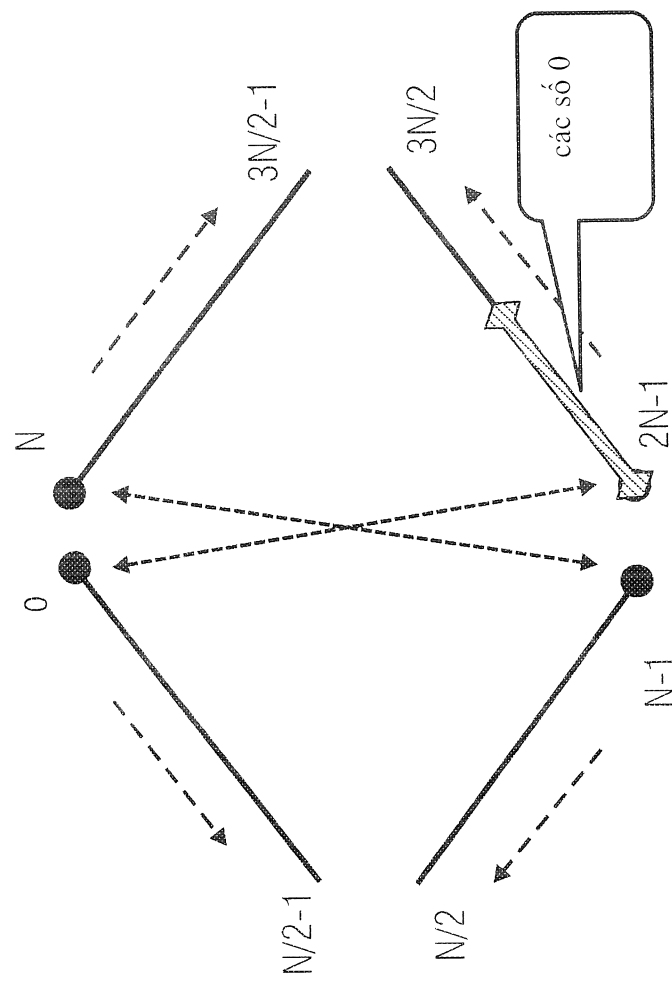


FIG 1

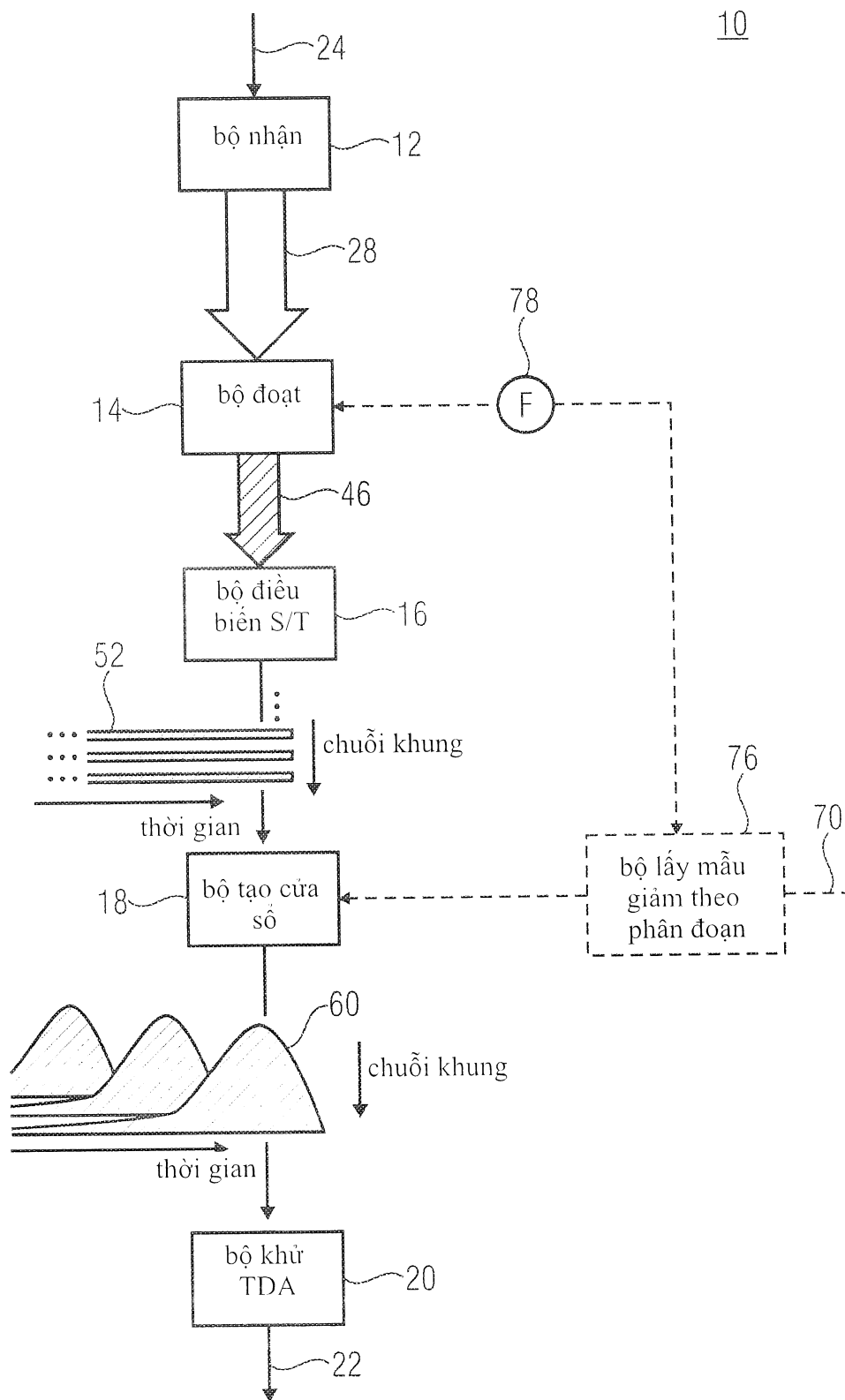


FIG 2

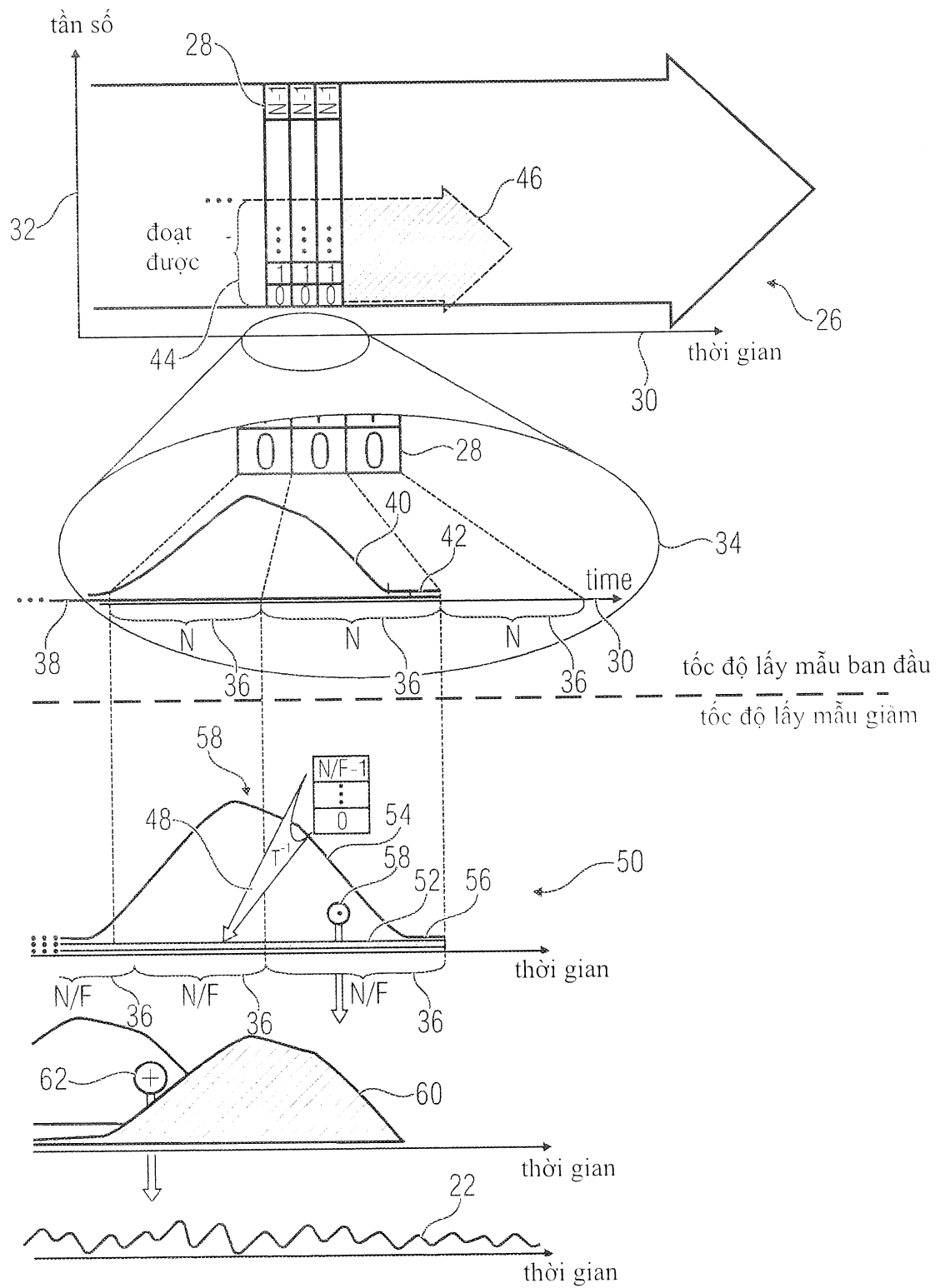


FIG 3

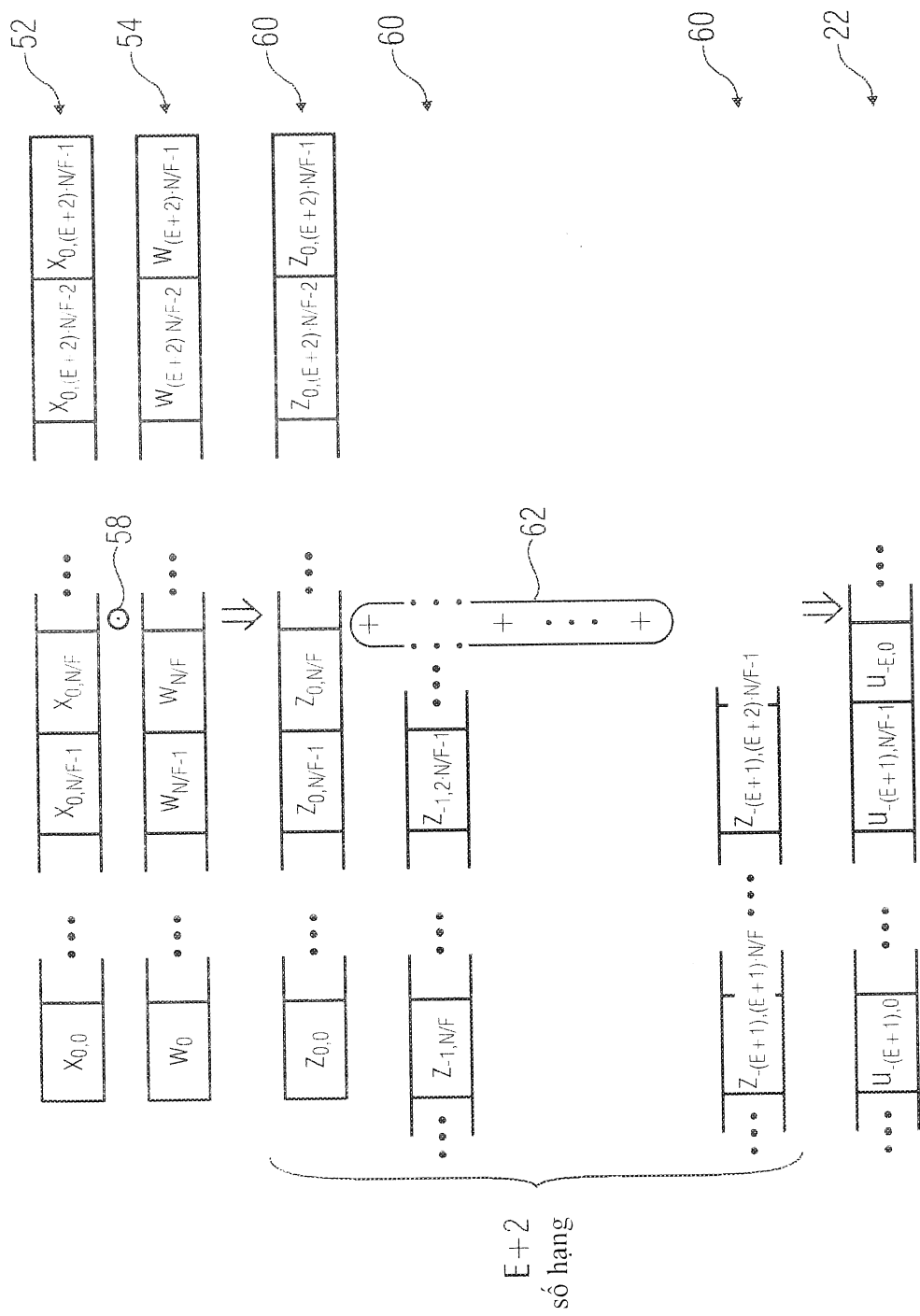


FIG 4

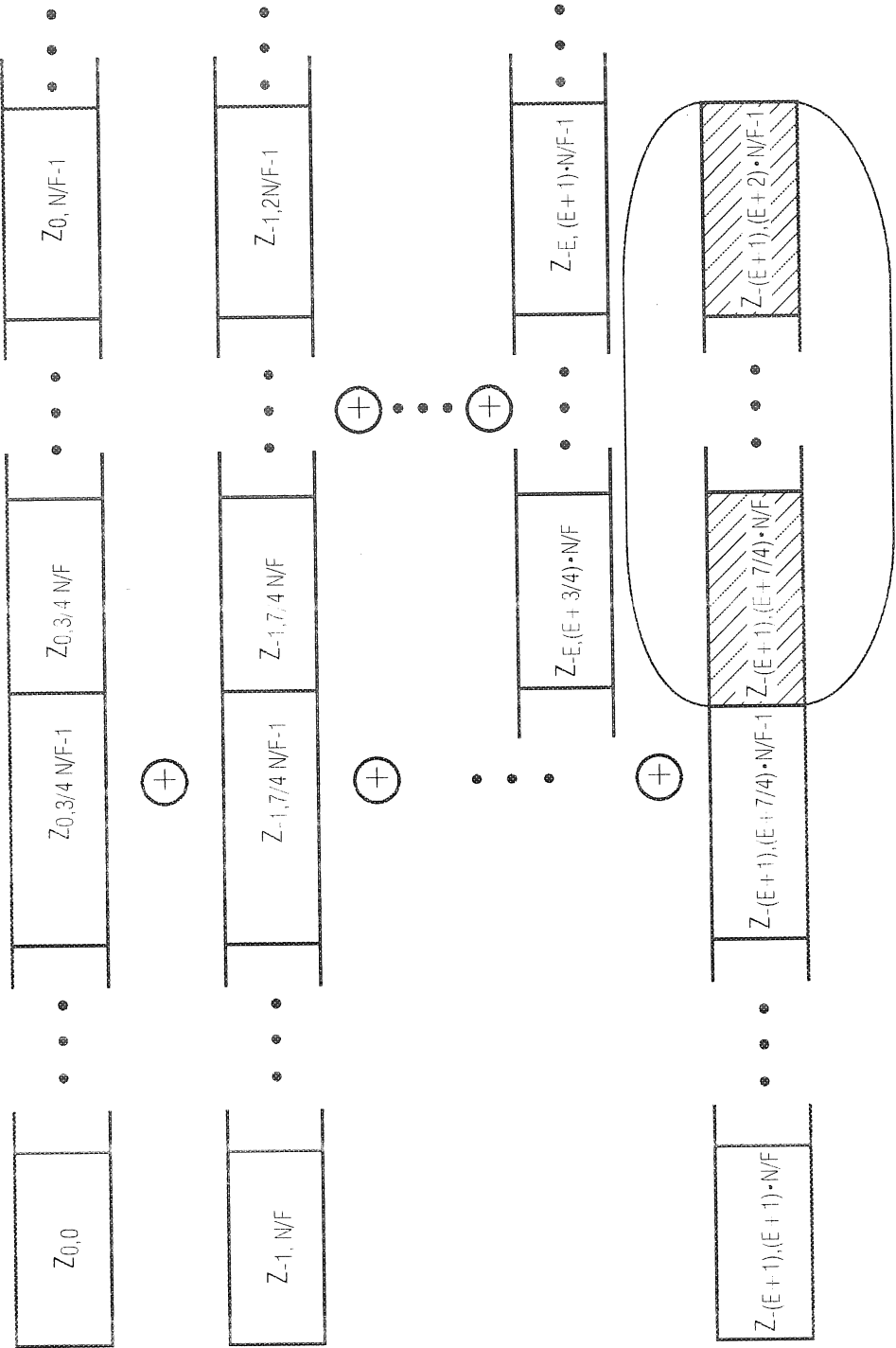


FIG 5

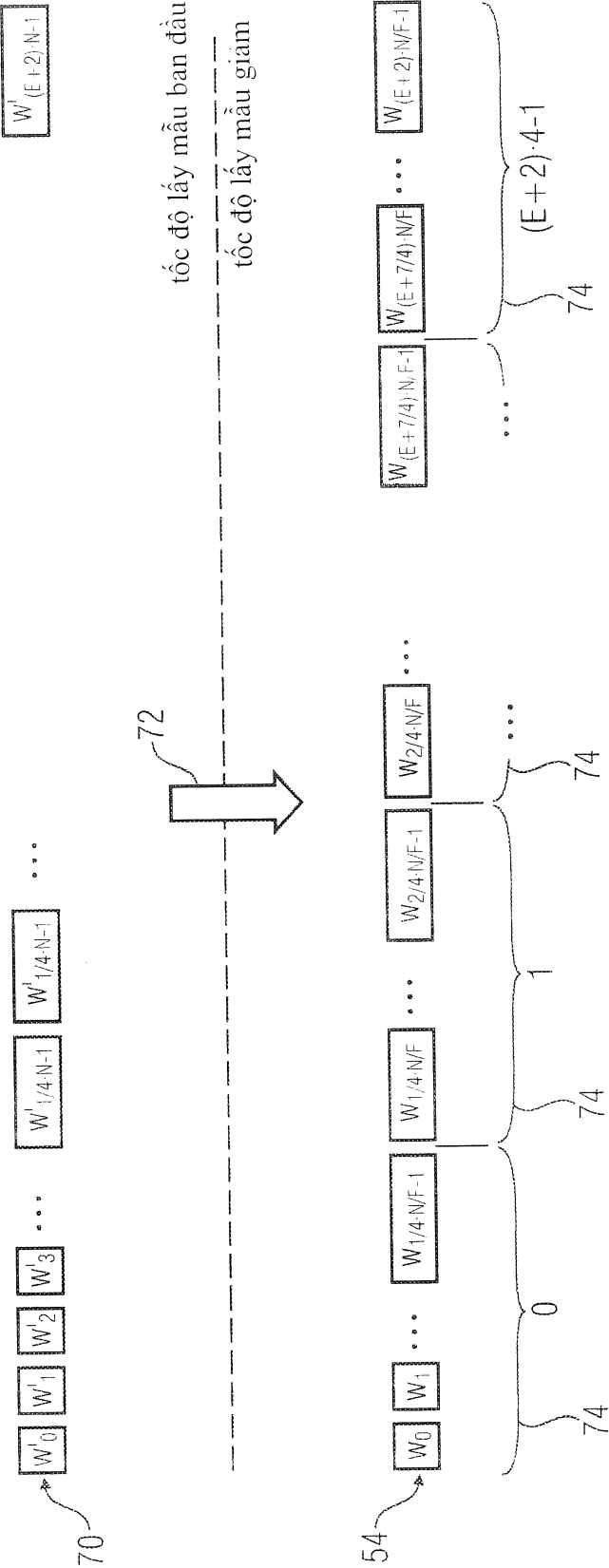


FIG 6

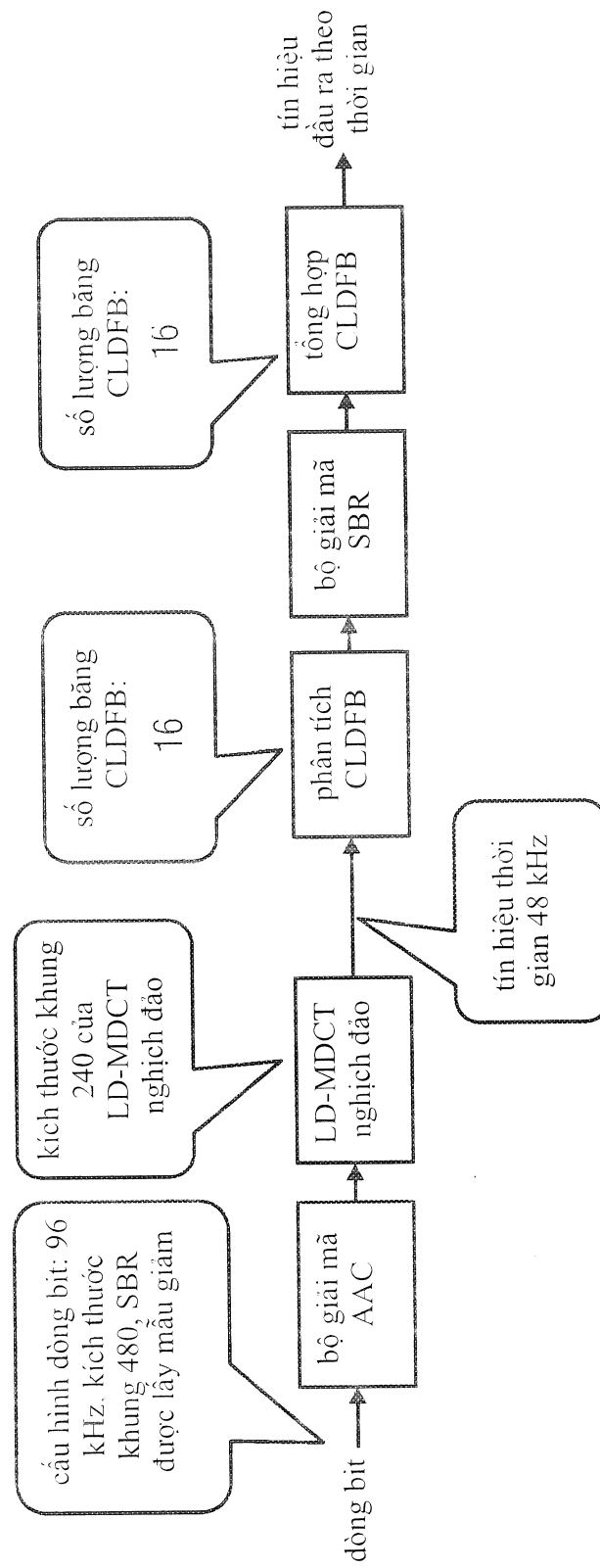


FIG 7

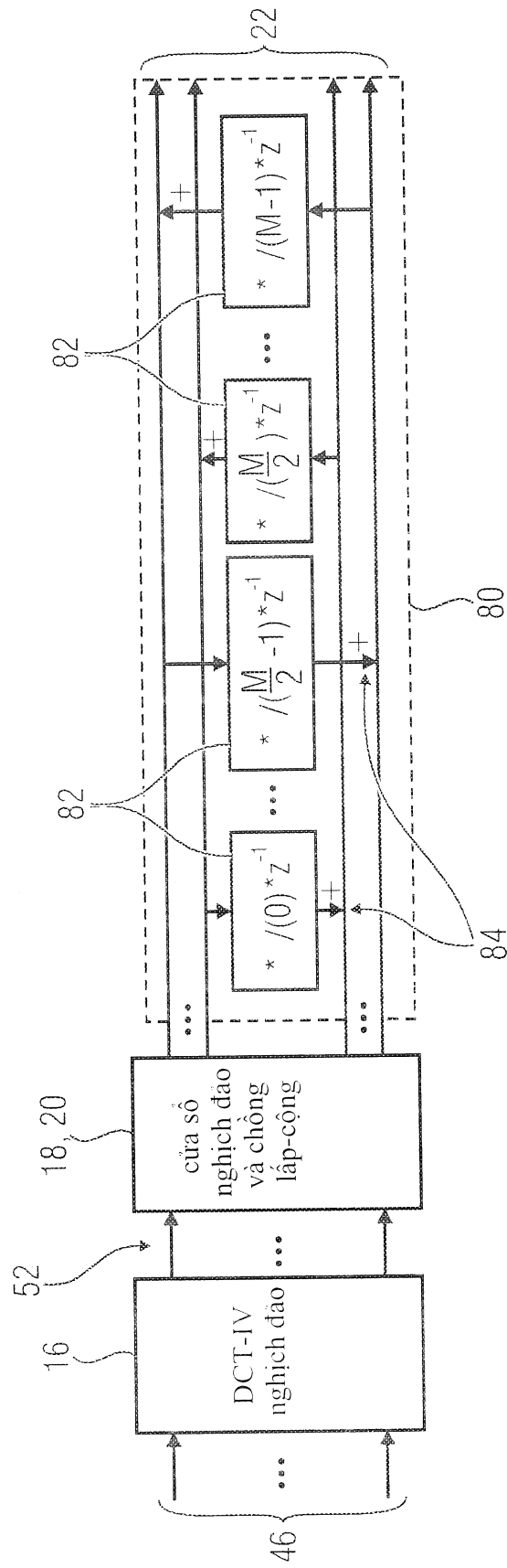


FIG 8

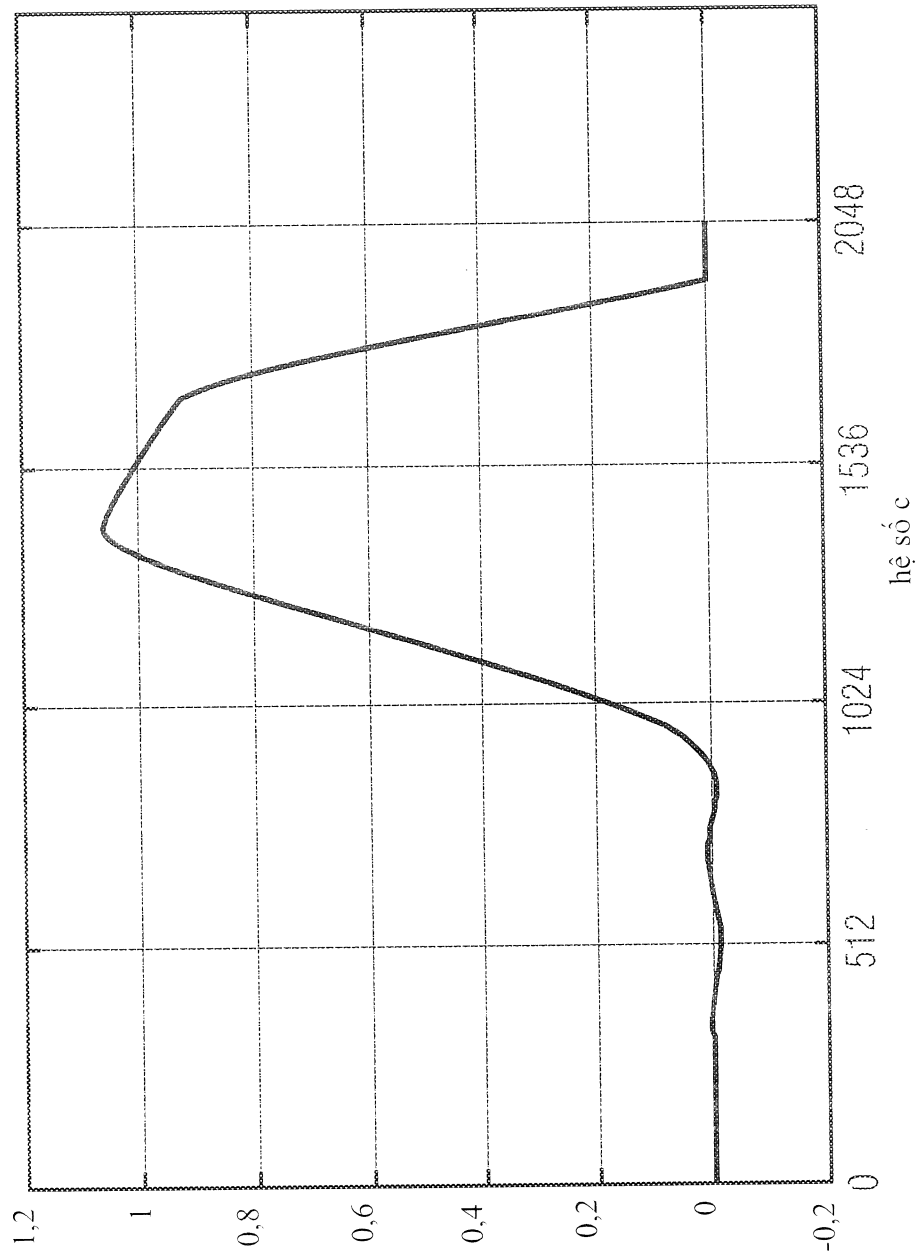


FIG 9