



(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)



1-0025672

(51)⁷

G10L 19/22; G10L 19/09; G10L 19/02; (13) B
G10L 19/08

(21) 1-2015-04426

(22) 21/07/2015

(86) PCT/EP2015/066677 21/07/2015

(87) WO 2016/016053 A1 04/02/2016

(30) 14178809.1 28/07/2014 EP

(45) 26/10/2020 391

(43) 25/04/2016 337A

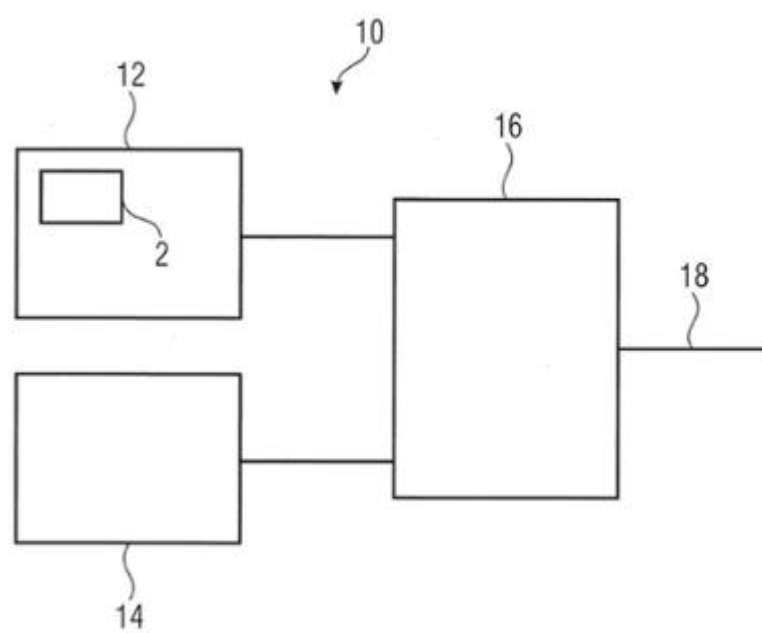
(73) Fraunhofer-Gesellschaft zur Foerderung der angewandten Forschung e.V. (DE)
Hansastrasse 27c, 80686 Muenchen, Germany

(72) RAVELLI, Emmanuel (FR); MULTRUS, Markus (DE); DOEHLA, Stefan (DE);
GRILL, Bernhard (DE); JANDER, Manuel (DE).

(74) CÔNG TY LUẬT TRÁCH NHIỆM HỮU HẠN AMBYS HÀ NỘI (AMBYS
HANOI)

(54) THIẾT BỊ VÀ PHƯƠNG PHÁP LỰA CHỌN MỘT TRONG SỐ THUẬT TOÁN
MÃ HÓA THỨ NHẤT CÓ ĐẶC ĐIỂM THỨ NHẤT VÀ THUẬT TOÁN MÃ HÓA
THỨ HAI CÓ ĐẶC ĐIỂM THỨ HAI, THIẾT BỊ MÃ HÓA PHẦN CỦA TÍN HIỆU
ÂM THANH, HỆ THỐNG MÃ HÓA VÀ GIẢI MÃ

(57) Sáng chế đề cập đến thiết bị và phương pháp lựa chọn một trong số thuật toán mã hóa thứ nhất có đặc điểm thứ nhất và thuật toán mã hóa thứ hai có đặc điểm thứ hai, thiết bị mã hóa phần của tín hiệu âm thanh, hệ thống mã hóa và giải mã. Thiết bị lựa chọn một trong số thuật toán mã hóa thứ nhất có đặc điểm thứ nhất và thuật toán mã hóa thứ hai có đặc điểm thứ hai để mã hóa phần của tín hiệu âm thanh để thu được phiên bản được mã hóa của phần của tín hiệu âm thanh, bao gồm bộ lọc được tạo cấu hình để nhận tín hiệu âm thanh, để giảm biên độ của các sóng hài trong tín hiệu âm thanh và để xuất ra phiên bản được lọc của tín hiệu âm thanh. Bộ ước lượng thứ nhất được đề xuất để sử dụng phiên bản được lọc của tín hiệu âm thanh trong phép ước lượng SNR hoặc SNR được phân đoạn của phần của tín hiệu âm thanh như số đo chất lượng thứ nhất cho một phần của tín hiệu âm thanh, mà được kết hợp với thuật toán mã hóa thứ nhất, mà không thực sự mã hóa và giải mã phần của tín hiệu âm thanh sử dụng thuật toán mã hóa thứ nhất. Bộ ước lượng thứ hai được đề xuất để ước lượng SNR hoặc SNR phân đoạn như số đo chất lượng thứ hai cho phần của tín hiệu âm thanh, mà được kết hợp với thuật toán mã hóa thứ hai, mà không thực sự mã hóa và giải mã phần của tín hiệu âm thanh sử dụng thuật toán mã hóa thứ hai. Thiết bị bao gồm bộ điều khiển để lựa chọn thuật toán mã hóa thứ nhất hoặc thuật toán mã hóa thứ hai dựa trên phép so sánh giữa số đo chất lượng thứ nhất và số đo chất lượng thứ hai.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến việc mã hóa âm thanh và, cụ thể, đề cập đến mã hóa âm thanh được chuyển đổi, mà, đối với các phần khác nhau của tín hiệu âm thanh, tín hiệu được mã hóa được tạo ra sử dụng các thuật toán mã hóa khác nhau.

Tình trạng kỹ thuật của sáng chế

Các bộ mã hóa âm thanh được chuyển đổi mà xác định các thuật toán mã hóa khác nhau cho các phần khác nhau của tín hiệu âm thanh được biết đến. Nói chung, các bộ mã hóa âm thanh được chuyển đổi cung cấp để chuyển đổi giữa hai chế độ khác nhau, tức là các thuật toán, như dự báo tuyến tính xuất phát từ mã đại số (ACELP - Algebraic Code Excited Linear Prediction) và kích thích được mã hóa biến đổi (TCX - Transform Coded Excitation).

Chế độ LPD của mã hóa âm thanh tiếng nói hợp nhất MPEG (MPEG Unified Speech Audio Coding - MPEG USAC) được dựa trên hai chế độ khác nhau ACELP và TCX. ACELP cung cấp chất lượng tốt hơn cho tín hiệu tương tự tiếng nói và tín hiệu tương tự chuyển tiếp. TCX cung cấp chất lượng tốt hơn cho tín hiệu tương tự âm nhạc và tín hiệu tương tự nhiễu âm. Bộ mã hóa quyết định sử dụng chế độ nào trên cơ sở khung cạnh khung. Sự quyết định tạo ra bởi bộ mã hóa có tính quyết định cho chất lượng bộ mã hóa-giải mã. Một quyết định sai lầm có thể tạo ra giả âm mạnh, cụ thể tại các tốc độ bit thấp.

Phương pháp dễ dàng nhất để quyết định chế độ nào được sử dụng là sự lựa chọn chế độ chu trình đóng, tức là để hoàn thành việc mã hóa/giải mã cả hai chế độ, sau đó tính toán tiêu chuẩn lựa chọn (ví dụ, SNR phân đoạn) cho cả hai chế độ dựa trên tín hiệu âm thanh và các tín hiệu âm thanh được mã hóa/giải mã, và cuối cùng chọn chế độ dựa trên tiêu chuẩn lựa chọn. Phương pháp này thường tạo ra quyết định ổn định và

chống nhiễu. Tuy nhiên, điều này cũng đòi hỏi sự phức tạp đáng kể, bởi vì cả hai chế độ phải được chạy tại mỗi khung.

Để giảm độ phức tạp, phương pháp thay thế là sự lựa chọn chế độ chu trình mở. Sự lựa chọn chu trình mở bao gồm không thực hiện sự mã hóa/giải mã hoàn thiện cả hai chế độ nhưng thay vì chọn một chế độ sử dụng tiêu chuẩn lựa chọn được tính toán với sự phức tạp thấp. Sự phức tạp trường hợp xấu nhất sau đó được giảm bởi sự phức tạp của chế độ ít phức tạp nhất (thường là TCX), trừ đi sự phức tạp cần để tính toán tiêu chuẩn lựa chọn. Sự lưu theo độ phức tạp thường đáng kể, điều này khiến loại phương pháp này thu hút khi độ phức tạp trường hợp xấu nhất bộ mã hóa-giải mã chịu ràng buộc.

Tiêu chuẩn AMR-WB+ (được định nghĩa trong tiêu chuẩn quốc tế 3GPP TS 26.290 V6.1.0 2004-12) bao gồm sự lựa chọn chế độ chu trình mở, được sử dụng để quyết định giữa tất cả các sự tổ hợp của ACELP/TCX20/TCX40/TCX80 trong khung 80ms. Điều này được mô tả trong Phần 5.2.4 của 3GPP TS 26.290. Điều này cũng được mô tả trong bài hội thảo “Low Complex Audio Encoding for Mobile, Multimedia, VTC 2006, Makinen et al.” và các bằng sáng chế US 7,747,430 B2 và US 7,739,120 B2 phản hồi lại tác giả của bài hội thảo này.

Patent US 7,747,430 B2 bộc lộ sự lựa chọn chế độ chu trình mở dựa trên sự phân tích các tham số dự báo dài hạn. Patent US 7,739,120 B2 bộc lộ sự lựa chọn chế độ chu trình mở dựa trên các đặc điểm tín hiệu biểu thị loại nội dung âm thanh trong các phản tương ứng của tín hiệu âm thanh, trong đó, nếu sự lựa chọn như vậy không thể tồn tại, sự lựa chọn còn dựa trên phép ước lượng bằng cách thống kê được thực hiện cho các phản xung quanh tương ứng.

Sự lựa chọn chế độ chu trình mở AMR-WB+ có thể được mô tả trong hai bước chính. Trong bước chính thứ nhất, một số đặc tính được tính toán trên tín hiệu âm thanh, như độ lệch tiêu chuẩn của các mức năng lượng, mối tương quan năng lượng tần số thấp/tần số cao, tổng năng lượng, khoảng cách cặp phổ điện vào (immittance spectral pair - ISP), độ trễ bước và độ khuếch đại, độ nghiêng phổ. Các đặc tính này sau đó được sử dụng để tạo ra lựa chọn giữa ACELP và TCX, sử dụng bộ phân loại trên cơ sở ngưỡng đơn giản. Nếu TCX được lựa chọn trong bước chính thứ nhất, sau

đó bước chính thứ hai quyết định giữa các tổ hợp khả thi TCX20/TCX40/TCX80 theo cách chu trình đóng.

Công bố đơn sáng chế quốc tế WO 2012/110448 A1 bộc lộ phương pháp để quyết định giữa hai thuật toán mã hóa có các đặc điểm khác nhau dựa trên kết quả phát hiện chuyển tiếp và kết quả chất lượng của tín hiệu âm thanh. Ngoài ra, việc áp dụng độ trễ được bộc lộ, trong đó độ trễ phụ thuộc vào sự lựa chọn được thực hiện trước đó, tức là, đối với các phần sớm hơn của tín hiệu âm thanh.

Trong bài thảo luận “Low Complex Audio Encoding for Mobile, Multimedia, VTC 2006, Makinen et al.”, sự lựa chọn chế độ chu trình đóng và chu trình mở của AMR-WB+ được so sánh. Các bài kiểm tra nghe chủ quan biểu thị rằng sự lựa chọn chế độ chu trình mở thực hiện kém hơn đáng kể so với sự lựa chọn chế độ chu trình đóng. Nhưng cũng được chỉ ra rằng sự lựa chọn chế độ chu trình mở giảm độ phức tạp trường hợp xấu nhất 40%.

Bản chất kỹ thuật của sáng chế

Mục đích của sáng chế là đề xuất phương pháp được cải thiện mà cho phép sự lựa chọn giữa thuật toán mã hóa thứ nhất và thuật toán mã hóa thứ hai với sự thực hiện tốt và độ phức tạp giảm.

Mục đích của sáng chế đạt được bởi thiết bị theo điểm 1, bởi phương pháp mã hóa theo điểm 18 và bởi chương trình máy tính theo điểm 19.

Các phương án của sáng chế đề xuất thiết bị để lựa chọn một trong số thuật toán mã hóa thứ nhất có đặc điểm thứ nhất và thuật toán mã hóa thứ hai có đặc điểm thứ hai để mã hóa một phần của tín hiệu âm thanh để thu được phiên bản được mã hóa của phần của tín hiệu âm thanh, thiết bị bao gồm:

bộ lọc được tạo cấu hình để nhận tín hiệu âm thanh, để giảm biên độ của các sóng hài trong tín hiệu âm thanh và để phát ra phiên bản được lọc của tín hiệu âm thanh;

bộ ước lượng thứ nhất để sử dụng phiên bản được lọc của tín hiệu âm thanh trong phép ước lượng tỉ số tín hiệu trên nhiễu (signal to noise ratio - SNR) hoặc SNR phân đoạn của phần của tín hiệu âm thanh như số đo chất lượng thứ nhất cho một phần của

tín hiệu âm thanh, mà được kết hợp với thuật toán mã hóa thứ nhất, mà không thực sự mã hóa và giải mã phần của tín hiệu âm thanh sử dụng thuật toán mã hóa thứ nhất;

bộ ước lượng thứ hai để ước lượng SNR hoặc SNR phân đoạn như số đo chất lượng thứ hai cho phần của tín hiệu âm thanh, mà được kết hợp với thuật toán mã hóa thứ hai, mà không thực sự mã hóa và giải mã phần của tín hiệu âm thanh sử dụng thuật toán mã hóa thứ hai; và

bộ điều khiển để lựa chọn thuật toán mã hóa thứ nhất hoặc thuật toán mã hóa thứ hai dựa trên phép so sánh giữa số đo chất lượng thứ nhất và số đo chất lượng thứ hai.

Các phương án của sáng chế đề xuất phương pháp lựa chọn một trong số thuật toán mã hóa thứ nhất có đặc điểm thứ nhất và thuật toán mã hóa thứ hai có đặc điểm thứ hai để mã hóa một phần của tín hiệu âm thanh để thu được phiên bản được mã hóa của phần của tín hiệu âm thanh, phương pháp bao gồm:

lọc tín hiệu âm thanh để giảm biên độ của các sóng hài trong tín hiệu âm thanh và để phát ra phiên bản được lọc của tín hiệu âm thanh;

sử dụng phiên bản được lọc của tín hiệu âm thanh trong phép ước lượng SNR hoặc SNR phân đoạn của phần của tín hiệu âm thanh như số đo chất lượng thứ nhất cho một phần của tín hiệu âm thanh, mà được kết hợp với thuật toán mã hóa thứ nhất, mà không thực sự mã hóa và giải mã phần của tín hiệu âm thanh sử dụng thuật toán mã hóa thứ nhất;

ước lượng số đo chất lượng thứ hai cho phần của tín hiệu âm thanh, mà được kết hợp với thuật toán mã hóa thứ hai, mà không thực sự mã hóa và giải mã phần của tín hiệu âm thanh sử dụng thuật toán mã hóa thứ hai; và

lựa chọn thuật toán mã hóa thứ nhất hoặc thuật toán mã hóa thứ hai dựa trên phép so sánh giữa số đo chất lượng thứ nhất và số đo chất lượng thứ hai.

Các phương án của sáng chế dựa trên sự công nhận rằng sự lựa chọn chu trình mở với hiệu quả được cải thiện có thể được thực hiện bằng cách ước lượng số đo chất lượng cho mỗi thuật toán trong số thuật toán mã hóa thứ nhất và thứ hai và lựa chọn một trong số các thuật toán mã hóa dựa trên phép so sánh giữa các số đo chất lượng thứ nhất và thứ hai. Các số đo chất lượng được ước lượng, tức là, tín hiệu âm thanh

không thực sự được mã hóa và được giải mã để thu được các số đo chất lượng. Do đó, các số đo chất lượng có thể được thu với độ phức tạp giảm. Sự lựa chọn chế độ sau đó có thể được thực hiện sử dụng các số đo chất lượng được ước lượng có thể so sánh với sự lựa chọn chế độ chu trình đóng. Hơn nữa, sáng chế được dựa trên sự công nhận rằng sự lựa chọn chế độ được cải thiện có thể thu được nếu phép ước lượng số đo chất lượng thứ nhất sử dụng phiên bản được lọc của phần của tín hiệu âm thanh, mà trong đó các sóng hài được giảm khi so với phiên bản không lọc của tín hiệu âm thanh.

Trong các phương án của sáng chế, sự lựa chọn chế độ chu trình mở trong đó SNR phân đoạn của ACELP và TCX được ước lượng đầu tiên với độ phức tạp thấp được thực hiện. Và sau đó sự lựa chọn chế độ được thực hiện sử dụng các giá trị SNR phân đoạn được ước lượng, như trong sự lựa chọn chế độ chu trình đóng.

Các phương án của sáng chế không tận dụng các đặc tính cổ điển + phương pháp phân loại như nó được thực hiện trong sự lựa chọn chế độ chu trình mở của AMR-WB+. Nhưng thay vào đó, các phương án của sáng chế cố để ước lượng số đo chất lượng của mỗi chế độ và lựa chọn chế độ có chất lượng tốt nhất.

Mô tả vắn tắt các hình vẽ

Các phương án của sáng chế bây giờ sẽ được mô tả chi tiết với sự viện dẫn đến các hình vẽ kèm theo, trong đó:

Fig.1 thể hiện sơ đồ giản lược của phương án của thiết bị để lựa chọn một trong số thuật toán mã hóa thứ nhất và thuật toán mã hóa thứ hai;

Fig.2 thể hiện sơ đồ giản lược của phương án của thiết bị mã hóa tín hiệu âm thanh;

Fig.3 thể hiện sơ đồ giản lược của phương án của thiết bị để lựa chọn một trong số thuật toán mã hóa thứ nhất và thuật toán mã hóa thứ hai;

Fig.4a và Fig.4b là sự biểu diễn khả thi của SNR và SNR phân đoạn.

Mô tả chi tiết sáng chế

Trong phần mô tả dưới đây, các phần tử/các bước giống nhau trong các hình vẽ khác nhau đề cập đến bởi các ký hiệu tham chiếu giống nhau. Lưu ý rằng các dấu hiệu

hình vẽ, như các liên kết tín hiệu và tương tự, mà không cần thiết trong việc hiểu sáng chế đã được loại bỏ.

Fig.1 thể hiện thiết bị 10 để lựa chọn một trong số thuật toán mã hóa thứ nhất, như thuật toán TCX và thuật toán mã hóa thứ hai, như thuật toán ACELP, như bộ mã hóa để mã hóa một phần của tín hiệu âm thanh. Thiết bị 10 bao gồm bộ ước lượng thứ nhất 12 để ước lượng SNR hoặc SNR phân đoạn của phần của tín hiệu âm thanh như số đo chất lượng thứ nhất cho phần tín hiệu được đề xuất. Số đo chất lượng thứ nhất được kết hợp với thuật toán mã hóa thứ nhất. Thiết bị 10 bao gồm bộ lọc 2 được tạo cấu hình để nhận tín hiệu âm thanh, để giảm biên độ của các sóng hài trong tín hiệu âm thanh và để xuất ra phiên bản được lọc của tín hiệu âm thanh. Bộ lọc 2 có thể là bên trong so với bộ ước lượng thứ nhất 12 như thể hiện trên Fig.1 hoặc có thể là bên ngoài so với bộ ước lượng thứ nhất 12. Bộ ước lượng thứ nhất 12 sử dụng phiên bản được lọc của tín hiệu âm thanh trong phép ước lượng số đo chất lượng thứ nhất. Nói cách khác, bộ ước lượng thứ nhất 12 ước lượng số đo chất lượng thứ nhất mà phần tín hiệu âm thanh sẽ có nếu được mã hóa và giải mã sử dụng thuật toán mã hóa thứ nhất, mà không thực sự mã hóa và giải mã phần tín hiệu âm thanh sử dụng thuật toán mã hóa thứ nhất. Thiết bị 10 bao gồm bộ ước lượng thứ hai 14 để ước lượng số đo chất lượng thứ hai cho phần tín hiệu. Số đo chất lượng thứ hai được kết hợp với thuật toán mã hóa thứ hai. Nói cách khác, bộ ước lượng thứ hai 14 ước lượng số đo chất lượng thứ hai mà phần tín hiệu âm thanh sẽ có nếu được mã hóa và giải mã sử dụng thuật toán mã hóa thứ hai, mà không thực sự mã hóa và giải mã phần tín hiệu âm thanh sử dụng thuật toán mã hóa thứ hai. Hơn nữa, thiết bị 10 bao gồm bộ điều khiển 16 để lựa chọn thuật toán mã hóa thứ nhất hoặc thuật toán mã hóa thứ hai dựa trên phép so sánh giữa số đo chất lượng thứ nhất và số đo chất lượng thứ hai. Bộ điều khiển có thể bao gồm đầu ra 18 biểu thị thuật toán mã hóa được lựa chọn.

Trong sự mô tả dưới đây, bộ ước lượng thứ nhất sử dụng phiên bản được lọc của tín hiệu âm thanh, tức là, phiên bản được lọc của phần tín hiệu âm thanh trong phép ước lượng số đo chất lượng nếu bộ lọc 2 được tạo cấu hình để giảm biên độ của các sóng hài được cung cấp và không được vô hiệu hóa, thậm chí không được biểu thị một cách rõ ràng.

Trong phương án, đặc điểm thứ nhất được kết hợp với thuật toán mã hóa thứ nhất được làm thích hợp hơn cho các tín hiệu tương tự âm nhạc và tín hiệu tương tự nhiễu, và đặc điểm mã hóa thứ hai được kết hợp với thuật toán mã hóa thứ hai được làm thích hợp hơn cho các tín hiệu tương tự tiếng nói và tín hiệu tương tự chuyển tiếp. Trong các phương án của sáng chế, thuật toán mã hóa thứ nhất là thuật toán mã hóa âm thanh, như thuật toán mã hóa biến đổi, tức là, thuật toán mã hóa phép biến đổi cosin rời rạc cải biên (modified discrete cosine transform - MDCT), như thuật toán mã hóa kích thích mã hóa biến đổi (transform coding excitation - TCX). Các thuật toán mã hóa biến đổi khác có thể dựa trên phép biến đổi FFT hoặc bất kỳ phép biến đổi hoặc giàn lọc khác. Trong các phương án của sáng chế, thuật toán mã hóa thứ hai là thuật toán mã hóa tiếng nói như thuật toán mã hóa dự báo tuyến tính kích thích mã (code excited linear prediction - CELP), như thuật toán mã hóa dự báo tuyến tính xuất phát từ mã đại số (algebraic code excited linear prediction - ACELP).

Trong các phương án, số đo chất lượng biểu diễn số đo chất lượng cảm giác. Giá trị đơn mà là phép ước lượng của chất lượng chủ quan của thuật toán mã hóa thứ nhất và giá trị đơn mà là phép ước lượng của chất lượng chủ quan của thuật toán mã hóa thứ hai có thể được tính toán. Thuật toán mã hóa mà cho chất lượng chủ quan được ước lượng tốt nhất có thể được chọn chỉ dựa trên phép so sánh giữa hai giá trị này. Điều này khác với những gì được làm trong tiêu chuẩn AMR-WB+ trong đó nhiều đặc tính biểu diễn các đặc điểm khác nhau của tín hiệu được tính toán và, sau đó, bộ phân loại được ứng dụng để quyết định thuật toán nào để lựa chọn.

Trong các phương án, số đo chất lượng tương ứng được ước lượng dựa trên một phần của tín hiệu âm thanh được gán trọng số, tức là, phiên bản được gán trọng số của tín hiệu âm thanh. Trong các phương án, tín hiệu âm thanh được gán trọng số có thể được định nghĩa như tín hiệu âm thanh được lọc bởi hàm gán trọng số, trong đó hàm gán trọng số là bộ lọc được gán trọng số LPC $A(z)/g$ với $A(z)$ là bộ lọc LPC và g là trọng số giữa 0 và 1 như 0,68. Hóa ra, các số đo chất lượng cảm giác tốt có thể thu được bằng cách này. Lưu ý rằng bộ lọc LPC $A(z)$ và bộ lọc được gán trọng số LPC $A(z)/g$ được xác định trong giai đoạn xử lý trước và rằng chúng cũng được sử dụng trong cả hai thuật toán mã hóa. Trong các phương án khác, hàm gán trọng số có thể là bộ lọc tuyến tính, bộ lọc FIR hoặc bộ lọc dự báo tuyến tính.

Trong các phương án, số đo chất lượng là tỉ số tín hiệu trên nhiễu (signal to noise ratio - SNR) phân đoạn trong miền tín hiệu được gán trọng số. Hóa ra là SNR phân đoạn trong miền tín hiệu được gán trọng số biểu diễn số đo chất lượng cảm giác tốt và, do đó, có thể được sử dụng như số đo chất lượng theo cách có lợi. Điều này cũng là số đo chất lượng được sử dụng trong cả hai thuật toán mã hóa ACELP và TCX để ước lượng các tham số mã hóa.

Số đo chất lượng khác có thể là SNR trong miền tín hiệu được gán trọng số. Các số đo chất lượng khác có thể là SNR phân đoạn, SNR của phần tương ứng của tín hiệu âm thanh trong miền tín hiệu âm thanh không được gán trọng số, tức là, không được lọc bởi các hệ số LPC (được gán trọng số).

Thông thường, SNR so sánh các tín hiệu âm thanh ban đầu và các tín hiệu âm thanh được xử lý (như các tín hiệu âm thanh tiếng nói) mẫu cạnh mẫu. Mục đích của điều này là để đo độ méo của các bộ mã hóa dạng sóng mà tái tạo dạng sóng đầu vào. SNR có thể được tính toán như thể hiện trên Fig.5a, trong đó $x(i)$ và $y(i)$ là các mẫu ban đầu và các mẫu được xử lý được biểu thị bởi i và N là tổng số các mẫu. SNR phân đoạn, thay vì làm việc trên toàn bộ tín hiệu, tính toán giá trị trung bình của các giá trị SNR của các phân đoạn ngắn, như nằm trong khoảng từ 1 đến 10 miligiây (millisecond - ms), chẳng hạn 5ms. SNR có thể được tính toán như được thể hiện trên Fig.5b, trong đó N và M lần lượt là độ dài phân đoạn và số lượng phân đoạn.

Trong các phương án của sáng chế, phần của tín hiệu âm thanh biểu thị khung của tín hiệu âm thanh mà thu được bằng cách tạo cửa sổ tín hiệu âm thanh và sự lựa chọn thuật toán mã hóa thích hợp được thực hiện cho nhiều khung liên tiếp thu được bằng cách tạo cửa sổ tín hiệu âm thanh. Trong phần mô tả dưới đây, đề cập đến tín hiệu âm thanh, các thuật ngữ "phần" và "khung" được sử dụng theo cách có thể trao đổi. Trong các phương án, mỗi khung được chia thành các khung con và SNR phân đoạn được ước lượng cho mỗi khung bằng cách tính toán SNR cho mỗi khung, được chuyển đổi theo dB và tính toán giá trị trung bình của các SNR khung con theo dB.

Do đó, theo các phương án, không phải SNR (phân đoạn) giữa tín hiệu âm thanh đầu vào và tín hiệu âm thanh được giải mã mà được ước lượng, nhưng SNR (phân đoạn) giữa tín hiệu âm thanh đầu vào được gán trọng số và tín hiệu âm thanh được giải

mã được gán trọng số được ước lượng. Miền là SNR (phân đoạn) này được đề cập, sự viện dẫn có thể được thực hiện đến chương 5.2.3 của tiêu chuẩn AMR-WB+ (Tiêu chuẩn quốc tế 3GPP TS 26.290 V6.1.0 2004-12).

Trong các phương án của sáng chế, số đo chất lượng tương ứng được ước lượng dựa trên năng lượng của một phần của tín hiệu âm thanh được gán trọng số và dựa trên độ méo được ước lượng được đưa vào khi mã hóa phần tín hiệu bằng thuật toán tương ứng, trong đó các bộ ước lượng thứ nhất và thứ hai được tạo cấu hình để xác định các độ méo được ước lượng phụ thuộc vào năng lượng của tín hiệu âm thanh được gán trọng số.

Trong các phương án của sáng chế, độ méo bộ lượng tử hóa được ước lượng được đưa vào bởi bộ lượng tử hóa được sử dụng trong thuật toán mã hóa thứ nhất khi lượng tử hóa phần của tín hiệu âm thanh được xác định và số đo chất lượng thứ nhất được xác định dựa trên năng lượng của phần của tín hiệu âm thanh được gán trọng số và độ méo bộ lượng tử hóa được ước lượng. Trong các phương án như vậy, độ khuếch đại toàn cục cho phần của tín hiệu âm thanh có thể được ước lượng sao cho phần của tín hiệu âm thanh sẽ sản xuất tốc độ truyền dữ liệu đích đã cho khi được mã hóa với bộ lượng tử hóa và bộ mã hóa entropi được sử dụng trong thuật toán mã hóa thứ nhất, trong đó độ méo bộ lượng tử hóa được ước lượng được xác định dựa trên độ khuếch đại toàn cục được ước lượng. Trong các phương án như vậy, độ méo bộ lượng tử hóa được ước lượng có thể được xác định dựa trên công suất khuếch đại được ước lượng. Khi bộ lượng tử hóa được sử dụng trong thuật toán mã hóa thứ nhất là bộ lượng tử hóa vô hướng đồng đều, bộ ước lượng thứ nhất có thể được tạo cấu hình để xác định độ méo bộ lượng tử hóa được ước lượng sử dụng công thức $D = G \cdot G / 12$, trong đó D là độ méo bộ lượng tử hóa được ước lượng và G là độ khuếch đại toàn cục được ước lượng. Trong trường hợp thuật toán mã hóa thứ nhất sử dụng bộ lượng tử hóa khác, độ méo bộ lượng tử hóa có thể được xác định từ độ khuếch đại toàn cục theo cách khác.

Các tác giả đã nhận ra rằng số đo chất lượng, như SNR phân đoạn, mà sẽ thu được khi mã hóa và giải mã phần của tín hiệu âm thanh sử dụng thuật toán mã hóa thứ nhất, như thuật toán TCX, có thể được ước lượng theo cách thích hợp bằng cách sử dụng các đặc tính trên trong bất kỳ tổ hợp nào của chúng.

Trong các phương án của sáng chế, số đo chất lượng thứ nhất là SNR phân đoạn và SNR phân đoạn được ước lượng bằng cách tính toán SNR được ước lượng được kết hợp với mỗi phần con trong nhiều phần con của phần của tín hiệu âm thanh dựa trên năng lượng của phần con tương ứng của tín hiệu âm thanh được gán trọng số và độ méo bộ lọc tử hóa được ước lượng và bằng cách tính toán giá trị trung bình của các SNR được kết hợp với các phần con của phần của tín hiệu âm thanh được gán trọng số để thu được SNR phân đoạn được ước lượng cho phần của tín hiệu âm thanh được gán trọng số.

Trong các phương án của sáng chế, độ méo băng mã thích ứng được ước lượng được đưa vào bởi băng mã thích ứng được sử dụng trong thuật toán mã hóa thứ hai khi sử dụng băng mã thích ứng để mã hóa phần của tín hiệu âm thanh được xác định, và số đo chất lượng thứ hai được ước lượng dựa trên năng lượng của phần của tín hiệu âm thanh được gán trọng số và độ méo băng mã thích ứng được ước lượng.

Trong các phương án như vậy, đối với mỗi phần con trong số nhiều phần con của phần của tín hiệu âm thanh, băng mã thích ứng có thể được làm gần đúng dựa trên phiên bản của phần con của tín hiệu âm thanh được gán trọng số được chuyển dịch về trước một độ trễ bước được xác định trong giai đoạn trước xử lý, độ khuếch đại băng mã thích ứng có thể được ước lượng sao cho sai số giữa phần con của phần của tín hiệu âm thanh được gán trọng số và băng mã thích ứng được làm gần đúng được tối thiểu hóa, và độ méo băng mã thích ứng được ước lượng có thể được xác định dựa trên năng lượng của sai số giữa phần con của phần tín hiệu âm thanh được gán trọng số và băng mã thích ứng được làm gần đúng được định tỉ lệ bởi độ khuếch đại băng mã thích ứng.

Trong các phương án của sáng chế, độ méo băng mã thích ứng được ước lượng được xác định cho mỗi phần con của phần của tín hiệu âm thanh có thể được giảm đi thừa số không đổi để xem xét sự giảm độ méo mà thu được bởi băng mã đổi mới trong thuật toán mã hóa thứ hai.

Trong các phương án của sáng chế, số đo chất lượng thứ hai là SNR phân đoạn và SNR phân đoạn được ước lượng bằng cách tính toán SNR được ước lượng được kết hợp với mỗi phần con dựa trên năng lượng phần con tương ứng của tín hiệu âm thanh

được gán trọng số và độ méo bằng mã thích ứng được ước lượng và bằng cách tính toán giá trị trung bình của các SNR được kết hợp với các phân con để thu được SNR phân đoạn được ước lượng.

Trong các phương án của sáng chế, bảng mã thích ứng được làm gần đúng dựa trên phiên bản của phần tín hiệu âm thanh được gán trọng số được chuyển dịch về trước một độ trễ bước được xác định trong giai đoạn xử lý trước, độ khuếch đại bằng mã thích ứng được ước lượng sao cho sai số giữa phần tín hiệu âm thanh được gán trọng số và bảng mã thích ứng được làm gần đúng được tối thiểu hóa, và độ méo bằng mã thích ứng được ước lượng được xác định dựa trên năng lượng giữa phần của tín hiệu âm thanh được gán trọng số và bảng mã thích ứng được làm gần đúng được định tỉ lệ bởi độ khuếch đại bằng mã thích ứng. Do đó, độ méo bằng mã được ước lượng có thể được xác định với độ phức tạp thấp.

Các tác giả sáng chế đã nhận ra rằng số đo chất lượng, như SNR phân đoạn, mà sẽ thu được khi mã hóa và giải mã phần của tín hiệu âm thanh sử dụng thuật toán mã hóa thứ hai, như thuật toán ACELP, có thể được ước lượng theo cách thích hợp bằng cách sử dụng các đặc tính trên trong bất kỳ tổ hợp nào của chúng.

Trong các phương án của sáng chế, cơ cấu độ trễ được sử dụng trong việc so sánh các số đo chất lượng được ước lượng. Điều này có thể quyết định thuật toán nào được sử dụng ổn định hơn. Cơ cấu độ trễ có thể phụ thuộc vào các số đo chất lượng được ước lượng (như sự chênh lệch giữa chúng) và các tham số khác, như số liệu thống kê về các quyết định trước đó, số lượng các khung tính theo thời gian, các chuyển tiếp trong khung. Miễn là các cơ cấu độ trễ được đề cập, sự viển dẫn có thể được thực hiện đến, ví dụ, công bố đơn sáng chế số WO 2012/110448 A1.

Trong các phương án của sáng chế, bộ mã hóa để mã hóa tín hiệu âm thanh bao gồm thiết bị 10, giai đoạn để thực hiện thuật toán mã hóa thứ nhất và giai đoạn thực hiện thuật toán mã hóa thứ hai, trong đó bộ mã hóa được tạo cấu hình để mã hóa phần của tín hiệu âm thanh sử dụng thuật toán mã hóa thứ nhất hoặc thuật toán mã hóa thứ hai phụ thuộc vào sự lựa chọn bởi bộ điều khiển 16. Trong các phương án của sáng chế, hệ thống để mã hóa và giải mã bao gồm bộ mã hóa và bộ giải mã được tạo cấu hình để nhận phiên bản được mã hóa của phần tín hiệu âm thanh và sự chỉ thị của thuật

toán được sử dụng để mã hóa phần tín hiệu âm thanh và để giải mã phiên bản được mã hóa của phần tín hiệu âm thanh sử dụng thuật toán được chỉ thị.

Thuật toán lựa chọn chế độ chu trình mở như thể hiện trên Fig.1 và được mô tả ở trên (ngoại trừ bộ lọc 2) được mô tả trong đơn đăng ký sáng chế PCT/EP2014/051557 trước đó. Thuật toán này được sử dụng để tạo sự lựa chọn giữa hai chế độ, như ACELP và TCX, trên cơ sở khung cạnh khung. Sự lựa chọn có thể dựa trên phép ước lượng của SNR phân đoạn của cả ACELP và TCX. Chế độ với SNR được phân đoạn được ước lượng cao nhất được lựa chọn. Một cách tùy ý, cơ cấu độ trễ có thể được sử dụng để cung cấp sự lựa chọn chống nhiễu hơn. SNR phân đoạn của ACELP có thể được ước lượng sử dụng phép làm gần đúng của độ méo bảng mã thích ứng và phép làm gần đúng của độ méo bảng mã đổi mới. Bảng mã thích ứng có thể được làm gần đúng trong miền tín hiệu được gán trọng số sử dụng độ trễ bước được ước lượng bởi thuật toán phân tích bước. Độ méo có thể được tính toán trong miền tín hiệu được gán trọng số giả định độ khuếch đại tối ưu. Độ méo có thể sau đó được giảm đi thừa số không đổi, làm gần đúng độ méo bảng mã đổi mới. SNR phân đoạn hoặc TCX có thể được ước lượng sử dụng phiên bản đơn giản hóa của bộ mã hóa TCX thực. Tín hiệu đầu vào có thể đầu tiên được biến đổi với MDCT, và sau đó được tạo hình sử dụng bộ lọc LPC được gán trọng số. Cuối cùng, độ méo có thể được ước lượng trong miền MDCT được gán trọng số, sử dụng độ khuếch đại toàn cục và bộ ước lượng độ khuếch đại toàn cục.

Hóa ra là thuật toán lựa chọn chế độ chu trình mở như được mô tả trong đơn đăng ký sáng chế trước đây cung cấp quyết định mong đợi hầu hết thời gian, lựa chọn ACELP trên các tín hiệu tương tự tiếng nói và tương tự chuyển tiếp và TCX trên các tín hiệu tương tự âm nhạc và tương tự nhiễu. Tuy nhiên, các tác giả sáng chế đã nhận ra rằng nó có thể xảy ra rằng ACELP đôi khi được lựa chọn trên một số tín hiệu âm nhạc sóng hài. Trên các tín hiệu như vậy, bảng mã thích ứng thường có độ khuếch đại dự báo cao, do tính có thể dự đoán được cao của các tín hiệu sóng hài, sản xuất độ méo thấp và sau đó SNR phân đoạn cao hơn TCX. Tuy nhiên, TCX nghe tốt hơn trên hầu hết các tín hiệu âm nhạc sóng hài, do đó TCX nên được ưu tiên trong các trường hợp này.

Do đó, sáng chế này đề xuất thực hiện phép ước lượng SNR hoặc SNR phân đoạn như số đo chất lượng thứ nhất sử dụng phiên bản tín hiệu đầu vào, mà được lọc để giảm các sóng hài của nó. Do đó, có thể thu được sự lựa chọn chế độ cải thiện trên các tín hiệu âm nhạc sóng hài.

Thông thường, bất kỳ bộ lọc thích hợp nào để giảm các sóng hài có thể được sử dụng. Trong các phương án của sáng chế, bộ lọc là bộ lọc dự báo dài hạn. Một ví dụ đơn giản của bộ lọc dự báo dài hạn là

$$F(z) = 1 - g \cdot z^{-T}$$

trong đó các tham số bộ lọc là độ khuếch đại "g" và độ trễ bước "T", mà được xác định từ tín hiệu âm thanh.

Các phương án của sáng chế dựa trên bộ lọc dự báo dài hạn mà được áp dụng cho tín hiệu âm thanh trước sự phân tích MDCT trong phép ước lượng SNR phân đoạn TCX. Bộ lọc dự báo dài hạn giảm biên độ các sóng hài trong tín hiệu đầu vào trước sự phân tích MDCT. Hậu quả là độ méo trong miền MDCT được gán trọng số được giảm, SNR phân đoạn được ước lượng của TCX tăng, và cuối cùng TCX được lựa chọn thường xuyên hơn trên các tín hiệu âm nhạc sóng hài.

Trong các phương án của sáng chế, hàm chuyển của bộ lọc dự báo dài hạn bao gồm một phần nguyên của độ trễ bước và bộ lọc đa nhánh phụ thuộc vào phần phân số của độ trễ bước. Điều này cho phép sự thực hiện hiệu quả vì phần nguyên được sử dụng chỉ trong khung tỉ lệ mẫu chuẩn ($z^{-T_{int}}$). Cùng thời gian, có thể đạt được độ chính xác cao do sự sử dụng phần nguyên trong bộ lọc đa nhánh. Bằng cách xem xét phần phân số trong sự loại bỏ bộ lọc đa nhánh của năng lượng của các sóng hài có thể đạt được khi tránh khỏi sự loại bỏ năng lượng của các phần gần các sóng hài.

Trong các phương án theo sáng chế, bộ lọc dự báo dài hạn được mô tả như sau:

$$P(z) = 1 - \beta g B(z, T_{fr}) z^{-T_{int}}$$

trong đó T_{int} và T_{fr} là số nguyên và phần phân số của độ trễ bước, g là độ khuếch đại, β là trọng số, và $B(z, T_{fr})$ là bộ lọc thông thấp FIR mà các hệ số phụ thuộc vào phần phân số của độ trễ bước. Các chi tiết khác về các phương án của bộ lọc dự báo dài hạn như vậy sẽ được nêu dưới đây.

Độ trễ bước và độ khuếch đại có thể được ước lượng trên cơ sở khung cạnh khung.

Bộ lọc dự báo có thể được vô hiệu hóa (độ khuếch đại = 0) dựa trên sự tổ hợp của một hoặc nhiều số đo độ hài hòa (ví dụ, sự tương quan được chuẩn hóa hoặc độ khuếch đại dự báo) và/hoặc một hoặc nhiều số đo kết cấu theo thời gian (ví dụ, số đo độ phẳng theo thời gian hoặc sự thay đổi năng lượng).

Bộ lọc có thể được áp dụng cho tín hiệu âm thanh đầu vào trên cơ sở khung cạnh khung. Nếu các tham số bộ lọc thay đổi từ một khung đến khung kế tiếp, sự gián đoạn có thể được đưa vào tại đường viền giữa hai khung. Trong các phương án, thiết bị còn bao gồm bộ phận để loại bỏ sự gián đoạn trong tín hiệu âm thanh được gây ra bởi bộ lọc. Để loại bỏ sự gián đoạn có khả năng, bất kỳ công nghệ nào có thể được sử dụng, như các công nghệ có thể so sánh với các công nghệ được mô tả trong các tài liệu US5012517, EP0732687A2, US5999899A, hoặc US7353168B2. Công nghệ khác để loại bỏ các sự gián đoạn có khả năng được mô tả dưới đây.

Trước khi mô tả phương án của bộ ước lượng thứ nhất 12 và bộ ước lượng thứ hai 14 một cách cụ thể viện dẫn đến Fig.3, phương án của bộ mã hóa 20 được mô tả viện dẫn đến Fig.2.

Bộ mã hóa 20 bao gồm bộ ước lượng thứ nhất 12, bộ ước lượng thứ hai 14, bộ điều khiển 16, bộ phận xử lý trước 22, công tắc 24, giai đoạn mã hóa thứ nhất 26 được tạo cấu hình để thực hiện thuật toán TCX, giai đoạn mã hóa thứ hai 28 được tạo cấu hình để thực hiện thuật toán ACELP, và giao diện đầu ra 30. Bộ phận xử lý trước 22 có thể là phần của bộ mã hóa USAC thông thường và có thể được tạo cấu hình để xuất ra các hệ số LPC, các hệ số LPC được gán trọng số, tín hiệu âm thanh được gán trọng số, và tập hợp độ trễ bước. Lưu ý rằng tất cả các tham số này được sử dụng trong cả thuật toán mã hóa, tức là, thuật toán TCX và thuật toán ACELP. Do đó, các tham số như vậy không phải được tính toán cho sự lựa chọn chế độ chu trình mở một cách bổ sung. Lợi ích của việc sử dụng các tham số đã được tính toán trong sự lựa chọn chế độ chu trình mở là tiết kiệm độ phức tạp.

Như được thể hiện trên Fig.2, thiết bị bao gồm bộ lọc giảm các sóng hài 2. Thiết bị còn bao gồm bộ phận vô hiệu tùy chọn 4 để vô hiệu bộ lọc giảm các sóng hài 2 dựa

trên tổ hợp của một hoặc nhiều số đo độ hài hòa (ví dụ, sự tương quan được chuẩn hóa hoặc độ khuếch đại dự báo) và/hoặc một hoặc nhiều số đo kết cấu theo thời gian (ví dụ, số đo độ phẳng theo thời gian hoặc sự thay đổi năng lượng). Thiết bị bao gồm bộ phận loại bỏ sự gián đoạn tùy chọn 6 để loại bỏ sự gián đoạn từ phiên bản được lọc của tín hiệu âm thanh. Ngoài ra, thiết bị tùy chọn bao gồm bộ phận 8 để ước lượng các tham số bộ lọc của bộ lọc giảm các sóng hài 2. Trên Fig.2, các thành phần này (2, 4, 6, và 8) được thể hiện như phần của bộ ước lượng thứ nhất 12. Điều đó rõ ràng mà không phải nói rằng các thành phần này có thể được thực hiện bên ngoài hoặc tách khỏi bộ ước lượng thứ nhất và có thể được tạo cấu hình để cung cấp phiên bản được lọc của tín hiệu âm thanh đến bộ ước lượng thứ nhất.

Tín hiệu âm thanh đầu vào 40 được cung cấp trên đường đầu vào. Tín hiệu âm thanh đầu vào 40 được áp dụng cho bộ ước lượng thứ nhất 12, bộ phận xử lý trước 22 và cả hai giai đoạn bộ mã hóa 26, 28. Trong bộ ước lượng thứ nhất 12, tín hiệu âm thanh đầu vào 40 được áp dụng cho bộ lọc 2 và phiên bản được lọc của tín hiệu âm thanh đầu vào được sử dụng trong phép ước lượng số đo chất lượng thứ nhất. Trong trường hợp bộ lọc được vô hiệu hóa bởi bộ phận vô hiệu 4, tín hiệu âm thanh đầu vào 40 được sử dụng trong phép ước lượng số đo chất lượng thứ nhất, hơn là phiên bản được lọc của tín hiệu âm thanh đầu vào. Bộ phận xử lý trước xử lý tín hiệu âm thanh đầu vào theo cách thông thường để suy ra các hệ số LPC và các hệ số LPC được gán trọng số 42 và để lọc tín hiệu âm thanh 40 với các hệ số LPC được gán trọng số 42 để thu được tín hiệu âm thanh được gán trọng số 44. Bộ phận xử lý trước 22 xuất ra các hệ số LPC được gán trọng số 42, tín hiệu âm thanh được gán trọng số 44 và tập hợp các độ trễ bước 48. Như được hiểu bởi những người có trình độ trong lĩnh vực kỹ thuật, các hệ số LPC được gán trọng số 42 và tín hiệu âm thanh được gán trọng số 44 có thể được phân đoạn thành các khung hoặc các khung con. Sự phân đoạn có thể thu được bằng cách tạo cửa sổ tín hiệu âm thanh theo cách phù hợp.

Trong các phương án thay thế, bộ xử lý trước có thể được cung cấp, mà được tạo cấu hình để tạo ra các hệ số LPC được gán trọng số và tín hiệu âm thanh được gán trọng số dựa trên phiên bản được lọc của tín hiệu âm thanh. Các hệ số LPC được gán trọng số và tín hiệu âm thanh được gán trọng số, mà dựa trên phiên bản được lọc của tín hiệu âm thanh mà sau đó được áp dụng đến bộ ước lượng thứ nhất để ước lượng số

đo chất lượng thứ nhất, hơn là các hệ số LPC được gán trọng số 42 và tín hiệu âm thanh được gán trọng số 44.

Trong các phương án của sáng chế, các hệ số LPC được lượng tử hóa hoặc các hệ số LPC được gán trọng số được lượng tử hóa có thể được sử dụng. Do đó, nên được hiểu rằng thuật ngữ "các hệ số LPC" được dự định để cũng bao gồm "các hệ số LPC được gán trọng số", và thuật ngữ "các hệ số LPC được gán trọng số" cũng được dự định để bao gồm "các hệ số LPC được lượng tử hóa được gán trọng số". Về mặt này, đáng lưu ý rằng thuật toán TCX của USAC sử dụng các hệ số LPC được gán trọng số được lượng tử hóa để tạo hình phổ MCDT.

Bộ ước lượng thứ nhất 12 nhận tín hiệu âm thanh 40, các hệ số LPC được gán trọng số 42 và tín hiệu âm thanh được gán trọng số 44 ước lượng số đo chất lượng thứ nhất 46 dựa trên đó và xuất số đo chất lượng thứ nhất đến bộ điều khiển 16. Bộ ước lượng thứ hai 16 nhận tín hiệu âm thanh được gán trọng số 44 và tập hợp các độ trễ bước 48, ước lượng số đo chất lượng thứ hai 50 dựa trên đó và xuất số đo chất lượng thứ hai 50 đến bộ điều khiển 16. Như được biết đến đối với những người có hiểu biết trung bình trong lĩnh vực kỹ thuật, các hệ số LPC được gán trọng số 42, tín hiệu âm thanh được gán trọng số 44 và tập hợp các độ trễ bước 48 đã được tính toán trong môđun trước đó (tức là, bộ phận xử lý trước 22) và, do đó, là có sẵn mà không có chi phí.

Bộ điều khiển đưa ra quyết định để lựa chọn hoặc thuật toán TCX hoặc thuật toán ACELP dựa trên phép so sánh của các số đo chất lượng nhận được. Như được biểu thị ở trên, bộ điều khiển có thể sử dụng cơ cấu độ trễ trong việc quyết định thuật toán nào được sử dụng. Sự lựa chọn giai đoạn bộ mã hóa thứ nhất 26 hoặc giai đoạn bộ mã hóa thứ hai 28 được thể hiện dưới dạng giản đồ trên Fig.2 bằng công tắc 24 mà được điều khiển bởi tín hiệu điều khiển được xuất ra bởi bộ điều khiển 16. Tín hiệu điều khiển 52 biểu thị liệu giai đoạn bộ mã hóa thứ nhất 26 hoặc giai đoạn bộ mã hóa thứ hai 28 là để được sử dụng. Dựa trên tín hiệu điều khiển 52, các tín hiệu yêu cầu được biểu thị dưới dạng giản đồ bởi mũi tên 54 trên Fig.2 và ít nhất bao gồm các hệ số LPC, các hệ số LPC được gán trọng số, tín hiệu âm thanh, tín hiệu âm thanh được gán trọng số, tập hợp các độ trễ bước được áp dụng cho hoặc giai đoạn bộ mã hóa thứ nhất

26 hoặc giai đoạn bộ mã hóa thứ hai 28. Giai đoạn bộ mã hóa được lựa chọn áp dụng thuật toán mã hóa được kết hợp và xuất ra sự biểu diễn được mã hóa 56 hoặc 58 đến giao diện đầu ra 30. Giao diện đầu ra 30 có thể được tạo cấu hình để xuất ra tín hiệu âm thanh được mã hóa 60 mà có thể bao gồm giữa các dữ liệu khác sự biểu diễn được mã hóa 56 hoặc 58, các hệ số LPC hoặc các hệ số LPC được gán trọng số, các tham số cho thuật toán mã hóa được lựa chọn và thông tin về thuật toán mã hóa được lựa chọn.

Các phương án cụ thể để ước lượng các số đo chất lượng thứ nhất và thứ hai, trong đó các số đo chất lượng thứ nhất và thứ hai là các SNR phân đoạn trong miền tín hiệu được gán trọng số và giờ được mô tả viện dẫn đến Fig.3. Fig.3 thể hiện bộ ước lượng thứ nhất 12 và bộ ước lượng thứ hai 14 và các chức năng của chúng dưới dạng lưu đồ thể hiện phép ước lượng tương ứng từng bước một.

Phép ước lượng SNR phân đoạn TCX

Bộ ước lượng (TCX) thứ nhất nhận tín hiệu âm thanh 40 (tín hiệu đầu vào), các hệ số LPC được gán trọng số 42 và tín hiệu âm thanh được gán trọng số 44 như đầu vào. Phiên bản được lọc của tín hiệu âm thanh 40 được tạo ra, bước 98. Trong phiên bản được lọc của tín hiệu âm thanh 40 các sóng hài được giảm hoặc được triệt.

Tín hiệu âm thanh 40 có thể được phân tích để xác định một hoặc nhiều số đo độ hài hòa (ví dụ, sự tương quan được chuẩn hóa hoặc độ khuếch đại dự báo) và/hoặc một hoặc nhiều số đo kết cấu theo thời gian (ví dụ, số đo độ phẳng theo thời gian hoặc sự thay đổi năng lượng). Dựa trên một trong số các số đo này hoặc tổ hợp của các số đo này, bộ lọc 2 và, do đó, sự lọc 98 có thể được vô hiệu hóa. Nếu sự lọc 98 được vô hiệu hóa, phép ước lượng số đo chất lượng thứ nhất được thực hiện sử dụng tín hiệu âm thanh 40 hơn là phiên bản được lọc của nó.

Trong các phương án của sáng chế, bước loại bỏ sự gián đoạn (không được thể hiện trên Fig.3) có thể tiếp theo sự lọc 98 để loại bỏ sự gián đoạn trong tín hiệu âm thanh, mà có thể là kết quả từ sự lọc 98.

Trong bước 100, phiên bản được lọc của tín hiệu âm thanh 40 được tạo cửa sổ. Việc tạo cửa sổ có thể diễn ra với cửa sổ dạng sin chồng lấp thấp 10ms. Khi khung trước là ACELP, kích thước khối có thể được tăng 5ms, bên trái của cửa sổ có thể là

hình chữ nhật và đáp ứng xung bằng 0 được tạo cửa sổ của bộ lọc tổng hợp ACELP có thể được loại bỏ khỏi tín hiệu đầu vào được tạo cửa sổ. Điều này tương tự với những gì được làm trong thuật toán TCX. Khung của phiên bản được lọc của tín hiệu âm thanh 40, mà biểu diễn phần của tín hiệu âm thanh, được xuất ra từ bước 100.

Trong bước 102, tín hiệu âm thanh được tạo cửa sổ, tức là, khung thu được, được biến đổi với phép biến đổi cosin rời rạc cải biên (modified discrete cosine transform - MDCT). Trong bước 104, sự tạo hình phổ được thực hiện bằng cách tạo hình phổ MDCT với các hệ số LPC được gán trọng số.

Trong bước 106 độ khuếch đại toàn cục G được ước lượng sao cho phổ được gán trọng số được lượng tử hóa với độ khuếch đại G sẽ tạo ra đích đã cho R, khi được mã hóa với bộ mã hóa entropi, ví dụ, bộ mã hóa số học. Thuật ngữ "độ khuếch đại toàn cục" được sử dụng vì một độ khuếch đại được xác định cho toàn bộ khung.

Ví dụ sự thực hiện của phép ước lượng độ khuếch đại toàn cục bây giờ sẽ được giải thích. Lưu ý rằng phép ước lượng độ khuếch đại toàn cục là thích hợp cho các phương án trong đó thuật toán mã hóa TCX sử dụng bộ lượng tử hóa với bộ mã hóa số học. Bộ lượng tử hóa vô hướng như vậy với bộ mã hóa số học được giả định theo tiêu chuẩn MPEG USAC.

Sự khởi tạo

Thứ nhất, các biến được sử dụng trong phép ước lượng độ khuếch đại được khởi tạo bởi:

1. Thiết lập $en[i] = 9,0 + 10,0 \cdot \log_{10}(c[4*i+0] + c[4*i+1] + c[4*i+2] + c[4*i+3])$, trong đó $0 \leq i < L/4$, $c[]$ là vectơ các hệ số để lượng tử hóa, và L là độ dài của $c[]$.
2. Thiết lập $fac = 128$, độ lệch = fac và đích = bất kỳ giá trị nào (ví dụ, 1000)

Sự lặp lại

Sau đó, khối thao tác sau đây được thực hiện NITER lần (ví dụ ở đây, NITER = 10).

1. $fac = fac/2$
2. $\text{độ lệch} = \text{độ lệch} - fac$

3. $ener = 0$

4. đối với mỗi i trong đó $0 \leq i < L/4$ thực hiện việc sau đây:

nếu $en[i] - \text{độ lệch} > 3,0$, sau đó $ener = ener + en[i] - \text{độ lệch}$

5. nếu $ener > \text{đích}$, vậy $\text{độ lệch} = \text{độ lệch} + fac$

Kết quả của sự lặp lại là giá trị độ lệch. Sau sự lặp lại, độ khuếch đại toàn cục được ước lượng như $G = 10^{(\text{độ lệch}/20)}$.

Cách cụ thể trong đó độ khuếch đại toàn cục được ước lượng có thể biến thiên phụ thuộc vào bộ lượng tử hóa và bộ mã hóa entropi được sử dụng. Trong tiêu chuẩn USAC MPEG bộ lượng tử hóa vô hướng với bộ mã hóa số học được giả định. Các phương pháp TCX khác có thể sử dụng bộ lượng tử hóa và được hiểu bởi những người có trình độ trung bình trong lĩnh vực kỹ thuật cách để ước lượng độ khuếch đại toàn cục cho các bộ lượng tử hóa khác nhau. Ví dụ, tiêu chuẩn AMR-WB+ giả định rằng bộ lượng tử hóa dàn RE8 được sử dụng. Đối với bộ lượng tử hóa như vậy, phép ước lượng độ khuếch đại toàn cục có thể được ước lượng như được mô tả trong chương 5.3.5.7 trên trang 34 của 3GPP TS 26.290 V6.1.0 2004-12, trong đó tốc độ truyền dữ liệu đích cố định được giả định.

Sau khi đã ước lượng độ khuếch đại toàn cục trong bước 106, phép ước lượng độ méo diễn ra trong bước 108. Cụ thể hơn, độ méo bộ lượng tử hóa được làm gần đúng dựa trên độ khuếch đại toàn cục được ước lượng. Trong phương án này, được giả định rằng bộ lượng tử hóa vô hướng đồng đều được sử dụng. Do đó độ méo bộ lượng tử hóa được xác định với công thức đơn giản $D = G * G/12$, trong đó D biểu diễn độ méo bộ lượng tử hóa được xác định và G biểu diễn độ khuếch đại toàn cục được ước lượng. Điều này tương ứng với phép làm gần đúng tốc độ cao của độ méo bộ lượng tử hóa vô hướng đồng đều.

Dựa trên độ méo bộ lượng tử hóa được xác định, sự tính toán SNR phân đoạn được thực hiện trong bước 110. SNR trong mỗi khung con của khung được tính toán như tỉ lệ năng lượng tín hiệu âm thanh được gán trọng số và độ méo D mà được giả định là hằng số trong các khung con. Ví dụ, khung được chia thành bốn khung con liên

tiếp (xem Fig.4). SNR phân đoạn sau đó là giá trị trung bình của các SNR của bốn khung con và có thể được chỉ thị trong dB.

Phương pháp này cho phép phép ước lượng SNR phân đoạn thứ nhất mà có thể thu được khi thực sự mã hóa và giải mã khung chủ thể sử dụng thuật toán TCX, tuy nhiên không phải thực sự mã hóa và giải mã tín hiệu âm thanh và, do đó, với độ phức tạp được giảm và thời gian tính toán được giảm mạnh.

Phép ước lượng SNR phân đoạn ACELP

Bộ ước lượng thứ hai 14 nhận tín hiệu âm thanh được gán trọng số 44 và tập hợp các độ trễ bước 48 mà đã được tính toán trong bộ phận xử lý trước 22.

Như được thể hiện trong bước 112, trong mỗi khung con, bảng mã thích ứng được làm gần đúng bằng cách đơn giản sử dụng tín hiệu âm thanh được gán trọng số và độ trễ bước T. Bảng mã thích ứng được làm gần đúng bởi

$$xw(n-T), n = 0, \dots, N$$

trong đó xw là tín hiệu âm thanh được gán trọng số, T là độ trễ bước của khung con tương ứng và N là độ dài khung con. Theo đó, bảng mã thích ứng được làm gần đúng bằng cách sử dụng phiên bản khung phụ được dịch chuyển về trước bởi T . Theo đó, trong các phương án của sáng chế, bảng mã thích ứng được làm gần đúng theo cách rất đơn giản.

Trong bước 114, độ khuếch đại bảng mã thích ứng cho mỗi khung phụ được xác định. Cụ thể hơn, trong mỗi khung phụ, độ khuếch đại bảng mã G được ước lượng sao cho độ khuếch đại này tối thiểu hóa sai số giữa tín hiệu âm thanh được gán trọng số và bảng mã thích ứng được làm gần đúng. Điều này có thể được thực hiện bằng cách so sánh đơn giản sự chênh lệch giữa cả hai tín hiệu cho mỗi mẫu và tìm độ khuếch đại sao cho tổng các độ chênh lệch này được tối thiểu hóa.

Trong bước 116, độ khuếch đại bảng mã thích ứng cho mỗi khung phụ được xác định. Trong mỗi khung phụ, độ méo D được đưa vào bởi bảng mã thích ứng đơn giản là năng lượng của sai số giữa tín hiệu âm thanh được gán trọng số và bảng mã thích ứng được làm gần đúng được định tỉ lệ bởi độ khuếch đại G .

Độ méo được xác định trong bước 116 có thể được điều chỉnh trong bước tùy chọn 118 để xem xét bảng mã đổi mới. Độ méo của bảng mã đổi mới được sử dụng trong các thuật toán ACELP có thể được ước lượng như các giá trị hằng số. Trong phương án được mô tả của sáng chế, đơn giản giả định rằng bảng mã đổi mới giảm độ méo D bởi một thừa số không đổi. Do đó, các độ méo thu được trong bước 116 cho mỗi khung con có thể được nhân trong bước 118 với một thừa số không đổi, như thừa số không đổi, thứ tự từ 0 đến 1, như 0,055.

Trong bước 120 sự tính toán SNR phân đoạn diễn ra. Trong mỗi khung con, SNR được tính toán như tỉ lệ năng lượng tín hiệu âm thanh được gán trọng số và độ méo D. SNR phân đoạn sau đó là giá trị trung bình của SNR của bốn khung con và có thể được biểu thị theo dB.

Phương pháp này cho phép phép ước lượng SNR thứ hai mà có thể thu được khi thực sự mã hóa và giải mã khung chủ thể sử dụng thuật toán ACELP, tuy nhiên không phải thực sự mã hóa và giải mã tín hiệu âm thanh và, do đó, với độ phức tạp được giảm và thời gian tính toán được giảm mạnh.

Các bộ ước lượng thứ nhất và thứ hai 12 và 14 xuất ra các SNR phân đoạn được ước lượng 46, 50 đến bộ điều khiển 16 và bộ điều khiển 16 quyết định thuật toán nào được sử dụng cho phần được kết hợp của tín hiệu âm thanh dựa trên các SNR phân đoạn được ước lượng 46, 50. Bộ điều khiển có thể tùy chọn sử dụng cơ cấu độ trễ để quyết định ổn định hơn. Ví dụ, cơ cấu độ trễ tương tự như trong quyết định chu trình đóng có thể được sử dụng với các tham số tinh chỉnh khác một chút. Cơ cấu độ trễ như vậy có thể tính toán giá trị "dsnr" mà có thể phụ thuộc vào các SNR phân đoạn được ước lượng (như sự chênh lệch giữa chúng) và các tham số khác, như số liệu thống kê về các quyết định trước đó, số lượng các khung tính theo thời gian, và các chuyển tiếp trong khung.

Không có cơ cấu độ trễ, bộ điều khiển có thể lựa chọn thuật toán mã hóa có SNR được ước lượng cao hơn, tức là, ACELP được lựa chọn nếu SNR được ước lượng thứ hai lớn hơn SNR được ước lượng thứ nhất và TCX được lựa chọn nếu SNR được ước lượng thứ nhất lớn hơn SNR được ước lượng thứ hai. Với cơ cấu độ trễ, bộ điều khiển

có thể lựa chọn thuật toán mã hóa theo quy tắc quyết định sau, trong đó $acelp_snr$ là SNR được ước lượng thứ hai và tcx_snr là SNR được ước lượng thứ nhất:

nếu $acelp_snr + dsnr > tcx_snr$ khi đó lựa chọn ACELP, nếu không lựa chọn TCX.

Sự xác định các tham số của bộ lọc để giảm biên độ của các sóng hài

Phương án xác định các tham số của bộ lọc để giảm biên độ của các sóng hài giờ sẽ được mô tả. Các tham số bộ lọc có thể được ước lượng tại phía bộ mã hóa, như trong bộ phận 8.

Phép ước lượng bước

Một độ trễ bước (phần nguyên + phần phân số) ôi khung được ước lượng (kích thước khung, ví dụ, 20ms). Điều này được thực hiện trong ba bước để giảm độ phức tạp và để cải thiện độ chính xác phép ước lượng.

a) Phép ước lượng thứ nhất của phần nguyên của độ trễ bước

Thuật toán phân tích bước mà tạo ra đường viền phát triển bước nhẵn được sử dụng (ví dụ, sự phân tích bước chu trình mở được mô tả trong Rec. ITU-T G.718, sec. 6.6). Sự phân tích này thường được thực hiện trên cơ sở khung con (kích thước khung con, ví dụ, 10ms), và sản xuất một số ước lượng độ trễ bước mỗi khung con. Lưu ý rằng các số ước lượng độ trễ bước này không có bất kỳ phần phân số nào và thường được ước lượng trên tín hiệu được xuống mẫu (tốc độ lấy mẫu, ví dụ, 6400Hz). Tín hiệu được sử dụng có thể là bất kỳ tín hiệu âm thanh nào, ví dụ, tín hiệu âm thanh được gán trọng số LPC như được mô tả trong Rec. ITU-T G.718, sec. 6.5.

b) Sự làm mịn phần nguyên T_{int} của độ trễ bước

Phần nguyên cuối cùng của độ trễ bước được ước lượng trên tín hiệu âm thanh $x[n]$ chạy tại tốc độ lấy mẫu bộ mã hóa lỗi, mà thường cao hơn tốc độ lấy mẫu của tín hiệu được xuống mẫu được sử dụng trong a) (ví dụ, 12,8kHz, 16kHz, 32kHz, v.v). Tín hiệu $x[n]$ có thể là tín hiệu âm thanh bất kỳ, ví dụ, tín hiệu âm thanh được gán trọng số LPC.

Phần nguyên T_{int} của độ trễ bước sau đó là độ trễ mà tối đa hóa hàm tự tương quan

$$C(d) = \sum_{n=0}^N x[n]x[n-d]$$

với d là khoảng độ trễ bước T được ước lượng trong a).

$$T - \delta_1 \leq d \leq T + \delta_2$$

c) Phép ước lượng phần phân số T_{fr} của độ trễ bước

Phần phân số T_{fr} được tìm ra bằng cách nội suy hàm tự tương quan $C(d)$ được tính toán trong bước b) và lựa chọn độ trễ bước phân số mà tối đa hóa hàm tự tương quan được nội suy. Phép nội suy có thể được thực hiện sử dụng bộ lọc FIR thông thấp như được mô tả trong ví dụ, Rec. ITU-T G.718, sec. 6.6.7.

Phép ước lượng độ khuếch đại và sự lượng tử hóa

Độ khuếch đại thường được ước lượng trên tín hiệu âm thanh đầu vào tại tốc độ lấy mẫu bộ mã hóa lõi, nhưng có thể cũng là bất kỳ tín hiệu âm thanh như tín hiệu âm thanh được gán trọng số LPC. Tín hiệu này được ghi chú là $y[n]$ và có thể là giống hoặc khác với $x[n]$.

Sự dự báo $y_P[n]$ của $y[n]$ được tìm thấy đầu tiên bằng cách lọc $y[n]$ với bộ lọc sau

$$P(z) = B(z, T_{fr})z^{-T_{int}}$$

với T_{int} phần nguyên của độ trễ bước (được ước lượng trong b)) và $B(z, T_{fr})$ bộ lọc FIR thông thấp mà các hệ số của chúng phụ thuộc vào phần phân số của độ trễ bước T_{fr} (được ước lượng trong c)).

Một ví dụ của $B(z)$ khi độ phân giải độ trễ bước là $\frac{1}{4}$:

$$\begin{array}{ll} T_{fr} = \frac{0}{4} & B(z) = 0,0000z^{-2} + 0,2325z^{-1} + 0,5349z^0 + 0,2325z^1 \\ T_{fr} = \frac{1}{4} & B(z) = 0,0152z^{-2} + 0,3400z^{-1} + 0,5094z^0 + 0,1353z^1 \\ T_{fr} = \frac{2}{4} & B(z) = 0,0609z^{-2} + 0,4391z^{-1} + 0,4391z^0 + 0,0609z^1 \\ T_{fr} = \frac{3}{4} & B(z) = 0,1353z^{-2} + 0,5094z^{-1} + 0,3400z^0 + 0,0152z^1 \end{array}$$

Độ khuếch đại g sau đó được tính toán như sau:

$$g = \frac{\sum_{n=0}^{N-1} y[n]y_P[n]}{\sum_{n=0}^{N-1} y_P[n]y_P[n]}$$

và được giới hạn giữa 0 và 1.

Cuối cùng, độ khuếch đại g được lượng tử hóa, ví dụ, trên 2 bit, sử dụng ví dụ, sự lượng tử hóa đồng đều.

β được sử dụng để điều khiển độ mạnh của bộ lọc. β bằng 1 tạo ra toàn bộ hiệu quả. β bằng 0 vô hiệu hóa bộ lọc. Do đó, trong các phương án của sáng chế, bộ lọc có thể được vô hiệu hóa bằng cách thiết lập β đến giá trị 0. Trong các phương án của sáng chế, nếu bộ lọc được kích hoạt, β có thể được thiết lập đến giá trị nằm trong khoảng từ 0,5 đến 0,75. Trong các phương án của sáng chế, nếu bộ lọc được kích hoạt, β có thể được thiết lập đến giá trị 0,625. Ví dụ của $B(z, T_{fr})$ được cho như trên. Thứ tự và các hệ số của $B(z, T_{fr})$ có thể cũng phụ thuộc vào tốc độ truyền dữ liệu và tốc độ lấy mẫu đầu ra. Đáp ứng tần số khác nhau có thể được thiết kế và tinh chỉnh cho mỗi tổ hợp của tốc độ truyền dữ liệu và tốc độ lấy mẫu đầu ra.

Vô hiệu hóa bộ lọc

Bộ lọc có thể được vô hiệu hóa dựa trên sự tổ hợp của một hoặc nhiều số đo độ hài hòa và/hoặc một hoặc nhiều số đo kết cấu theo thời gian. Các ví dụ của các số đo như vậy được mô tả dưới đây:

i) Số đo độ hài hòa như sự tương quan được chuẩn hóa tại độ trễ bước nguyên được ước lượng trong bước b).

$$norm. corr. = \frac{\sum_{n=0}^N x[n]x[n - T_{int}]}{\sqrt{\sum_{n=0}^N x[n]x[n]} \sqrt{\sum_{n=0}^N x[n - T_{int}]x[n - T_{int}]}}$$

Sự tương quan được chuẩn hóa là 1 nếu tín hiệu đầu vào được dự báo một cách hoàn hảo bởi độ trễ bước nguyên, và là 0 nếu nó không thể được dự báo. Giá trị cao (gần đến 1) sau đó sẽ biểu thị tín hiệu sóng hài. Đối với quyết định mạnh hơn, sự tương quan được chuẩn hóa của khung trước có thể cũng được sử dụng trong quyết định, ví dụ:

Nếu $(\text{norm.corr}(\text{curr.}) * \text{norm.corr}(\text{prev.})) > 0,25$, khi đó bộ lọc không được vô hiệu hóa

ii) Các số đo kết cấu theo thời gian được tính toán, ví dụ, trên cơ sở các mẫu năng lượng cũng được sử dụng bởi bộ phát hiện chuyển tiếp cho sự phát hiện chuyển tiếp (ví dụ, số đo độ phẳng theo thời gian, sự thay đổi năng lượng), ví dụ

nếu (số đo độ phẳng theo thời gian $> 3,5$ hoặc sự thay đổi năng lượng $> 3,5$) khi đó bộ lọc được vô hiệu hóa.

Nhiều chi tiết liên quan đến sự xác định của một hoặc nhiều số đo sự hài hòa được nêu dưới đây.

Số đo của sự hài hòa là, ví dụ, được tính toán bởi sự tương quan được chuẩn hóa của tín hiệu âm thanh hoặc phiên bản được thay đổi trước của nó tại hoặc quanh độ trễ bước. Độ trễ bước thậm chí có thể được xác định trong các bước bao gồm giai đoạn thứ nhất và giai đoạn thứ hai, trong đó, trong giai đoạn thứ nhất, phép ước lượng sơ bộ độ trễ bước được xác định tại miền được xuống mẫu của tốc độ lấy mẫu thứ nhất và, nằm trong giai đoạn thứ hai, phép ước lượng sơ bộ của độ trễ bước được tinh chỉnh tại tốc độ lấy mẫu thứ hai, cao hơn tốc độ lấy mẫu thứ nhất. Độ trễ bước, ví dụ, được xác định sử dụng sự tự tương quan. Ít nhất một số đo kết cấu theo thời gian, ví dụ, được xác định nằm trong vùng theo thời gian được đặt theo thời gian phụ thuộc vào thông tin bước. Đầu hướng về trước theo thời gian của vùng theo thời gian, ví dụ, được đặt phụ thuộc vào thông tin bước. Đầu hướng về trước theo thời gian của vùng theo thời gian có thể được đặt sao cho đầu hướng về trước theo thời gian của vùng theo thời gian được đổi thành hướng về trước bởi lượng thời gian tăng một cách đơn điệu với sự tăng của thông tin bước. Đầu hướng về phía trước theo thời gian của vùng theo thời gian có thể được định vị phụ thuộc vào kết cấu theo thời gian của tín hiệu âm thanh nằm trong vùng ứng cử theo thời gian mở rộng từ đầu hướng về trước theo thời gian của vùng theo thời gian hoặc, của vùng của sự ảnh hưởng cao hơn lên trên sự xác định của số đo kết cấu theo thời gian, đến đầu hướng về tương lai theo thời gian của khung hiện thời. Biên độ hoặc tỉ lệ giữa các mẫu năng lượng tối đa và tối thiểu trong vùng ứng cử theo thời gian có thể được sử dụng cho đầu này. Ví dụ, ít nhất một số đo cấu trúc theo thời gian có thể đo giá trị trung bình hoặc sự biến thiên năng lượng tối đa của

tín hiệu âm thanh nằm trong vùng theo thời gian và điều kiện của sự vô hiệu hóa có thể được đáp ứng nếu cả hai ít nhất một số đo kết cấu theo thời gian là nhỏ hơn ngưỡng thứ nhất được định trước và số đo độ hài hòa là, đối với khung hiện thời và/hoặc khung trước, trên ngưỡng thứ hai. Điều kiện cũng được đáp ứng nếu số đo độ hài hòa là, đối với khung hiện thời, trên ngưỡng thứ ba và số đo độ hài hòa là, đối với khung hiện thời và/hoặc khung trước, trên ngưỡng thứ tư mà giảm với sự tăng của độ trễ bước.

Sự mô tả từng bước của phương án cụ thể để xác định các số đo sẽ được trình bày sau đây.

Bước 1: Sự phát hiện chuyển tiếp và các số đo theo thời gian

Tín hiệu đầu vào $s_{HP}(n)$ được nhập vào trong bộ phát hiện chuyển tiếp miền thời gian. Tín hiệu đầu vào $s_{HP}(n)$ được lọc thông cao. Hàm chuyển của bộ lọc HP của sự phát hiện chuyển tiếp được cho bởi

$$H_{TD}(z) = 0,375 - 0,5z^{-1} + 0,125z^{-2} \quad (1)$$

Tín hiệu, được lọc bởi bộ lọc HP của sự phát hiện chuyển tiếp, được biểu thị là $s_{TD}(n)$. Tín hiệu được lọc HP $s_{TD}(n)$ được phân đoạn thành 8 phân đoạn liên tiếp có cùng độ dài. Năng lượng của tín hiệu được lọc HP $s_{TD}(n)$ cho mỗi phân đoạn được tính như:

$$E_{TD}(i) = \sum_{n=0}^{L_{segment}-1} (s_{TD}(iL_{segment} + n))^2, \quad i = 0, \dots, 7 \quad (2)$$

trong đó $L_{segment} = \frac{L}{8}$ là số lượng các mẫu trong phân đoạn 2,5ms tại tần số lấy mẫu đầu vào.

Năng lượng được tích lũy được tính sử dụng:

$$E_{Acc} = \max(E_{TD}(i-1), 0,8125E_{Acc}) \quad (3)$$

Sự tấn công được phát hiện nếu năng lượng của phân đoạn $E_{TD}(i)$ vượt quá năng lượng được tích tụ một thừa số không đổi $attackRatio = 8,5$ và $attackIndex$ được thiết lập là i :

$$E_{TD}(i) > attackRatio \cdot E_{Acc} \quad (4)$$

Nếu không có sự tấn công nào được phát hiện dựa trên tiêu chí i , chỉ số tấn công được thiết lập thành i mà không biểu thị sự có mặt của sự tấn công. Chỉ số tấn công về cơ bản được thiết lập đến vị trí của sự tấn công cuối cùng trong khung với sự hạn chế bổ sung.

Sự thay đổi năng lượng cho mỗi phân đoạn được tính là:

$$E_{chg}(i) = \begin{cases} \frac{E_{TD}(i)}{E_{TD}(i-1)}, & E_{TD}(i) > E_{TD}(i-1) \\ \frac{E_{TD}(i-1)}{E_{TD}(i)}, & E_{TD}(i-1) > E_{TD}(i) \end{cases} \quad (5)$$

Số đo độ phẳng theo thời gian được tính toán là:

$$TFM(N_{past}) = \frac{1}{8 + N_{past}} \sum_{i=-N_{past}}^7 E_{chg}(i) \quad (6)$$

Sự thay đổi năng lượng tối đa được tính là:

$$MEC(N_{past}, N_{new}) = \max(E_{chg}(-N_{past}), E_{chg}(-N_{past} + 1), \dots, E_{chg}(N_{new} - 1)) \quad (7)$$

Nếu chỉ số $E_{chg}(i)$ hoặc $E_{TD}(i)$ là âm khi đó nó biểu thị giá trị từ phân đoạn trước, với sự biểu thị phân đoạn liên quan đến khung hiện thời.

N_{past} là số lượng các phân đoạn từ khung trước. Nó bằng với 0 nếu số đo độ phẳng theo thời gian được tính toán cho sự sử dụng trong quyết định ACELP/TCX. Nếu số đo độ phẳng theo thời gian được tính toán cho quyết định LTP TCX khi đó nó sẽ bằng:

$$N_{past} = 1 + \min\left(8, \left\lceil 8 \frac{pitch}{L} + 0.5 \right\rceil\right) \quad (8)$$

N_{new} là số lượng các phân đoạn từ khung hiện thời. Nó bằng 8 đối với các khung không chuyển tiếp. Đối với các khung chuyển tiếp, các vị trí của các phân đoạn với năng lượng tối đa và tối thiểu được tìm ra:

$$i_{max} = \arg \max_{i \in \{-N_{past}, \dots, 7\}} E_{TD}(i) \quad (9)$$

$$i_{\min} = \arg \min_{i \in \{-N_{past}, \dots, 7\}} E_{TD}(i) \quad (10)$$

Nếu $E_{TD}(i_{\min}) > 0.375 E_{TD}(i_{\max})$ khi đó N_{new} được thiết lập là $i_{\max} - 3$, nếu không N_{new} được thiết lập là 8.

Bước 2. Sự chuyển đổi độ dài khối biến đổi

Độ dài chồng lấp và độ dài khối biến đổi của TCX phụ thuộc vào sự có mặt của sự chuyển tiếp và vị trí của nó.

Bảng 1: Mã hóa sự chồng lấp và độ dài biến đổi dựa trên vị trí sự chuyển tiếp

Chỉ số tấn công	Sự chồng lấp với cửa sổ thứ nhất của khung sau	Quyết định biến đổi dài/ngắn (được mã hóa nhị phân) 0 – Dài, 1 - Ngắn	Mã hóa nhị phân cho độ rộng chồng lấp	Mã hóa chồng lấp
Không	ALDO	0	0	00
-2	ĐẦY	1	0	10
-1	ĐẦY	1	0	10
0	ĐẦY	1	0	10
1	ĐẦY	1	0	10
2	TỐI THIỂU	1	10	110
3	MỘT NỬA	1	11	111
4	MỘT NỬA	1	11	111
5	TỐI THIỂU	1	10	110
6	TỐI THIỂU	0	10	010
7	MỘT NỬA	0	11	011

Bộ phát hiện sự chuyển tiếp được mô tả như trên về cơ bản trả lại chỉ số cho sự tấn công cuối cùng với sự giới hạn rằng nếu có nhiều sự chuyển tiếp khi đó sự chồng lấp TỐI THIỂU được ưu tiên hơn sự chồng lấp MỘT NỬA mà được ưu tiên hơn sự chồng lấp ĐẦY. Nếu sự tấn công tại vị trí 2 hoặc 6 không đủ mạnh thì khi đó sự chồng lấp MỘT NỬA được lựa chọn thay vì sự chồng lấp TỐI THIỂU.

Bước 3. Phép ước lượng bước

Một độ trễ bước (phần nguyên + phần phân số) mỗi khung được ước lượng (kích thước khung, ví dụ, 20ms) như đã nêu trên trong 3 bước a) đến c) để giảm độ phức tạp và cải thiện độ chính xác phép ước lượng.

Bước 4. Bit quyết định

Nếu tín hiệu âm thanh đầu vào không chứa bất kỳ nội dung sóng hài nào hoặc nếu công nghệ dựa trên sự dự báo sẽ đưa vào sự méo trong kết cấu thời gian (ví dụ, sự lặp lại các sự chuyển tiếp ngắn), khi đó quyết định rằng bộ lọc được vô hiệu hóa được thực hiện.

Quyết định được thực hiện dựa trên một số tham số như sự tương quan được chuẩn hóa tại độ trễ bước nguyên và các số đo kết cấu theo thời gian.

Sự tương quan được chuẩn hóa tại độ trễ bước nguyên norm_corr được ước lượng như đã nêu trên. Sự tương quan được chuẩn hóa là 1 nếu tín hiệu đầu vào được dự báo một cách hoàn hảo bởi độ trễ bước nguyên, và là 0 nếu nó không thể được dự báo. Giá trị cao (gần đến 1) sau đó sẽ biểu thị tín hiệu sóng hài. Đối với quyết định mạnh hơn, ngoài sự tương quan được chuẩn hóa cho khung hiện thời ($\text{norm_corr}(\text{curr})$) sự tương quan được chuẩn hóa của khung trước ($\text{norm_corr}(\text{prev})$) có thể cũng được sử dụng trong quyết định, ví dụ:

$$\text{Nếu } (\text{norm_corr}(\text{curr}) * \text{norm_corr}(\text{prev})) > 0,25$$

hoặc

$$\text{Nếu } \max(\text{norm_corr}(\text{curr}) * \text{norm_corr}(\text{prev})) > 0,5$$

khi đó khung hiện thời chứa một số nội dung sóng hài.

Các số đo kết cấu theo thời gian có thể được tính toán bởi bộ phát hiện chuyển tiếp (ví dụ, số đo độ phẳng theo thời gian (phương trình (6)) và phương trình thay đổi năng lượng tối đa (7)), để tránh sự kích hoạt bộ lọc trên tín hiệu chứa sự chuyển tiếp mạnh hoặc sự thay đổi theo thời gian lớn. Các đặc tính theo thời gian được tính toán trên tín hiệu chứa khung hiện thời (các phân đoạn N_{new}) và khung trước lên đến độ trễ bước (các phân đoạn N_{past}). Đối với bước như các sự chuyển tiếp mà dần dần phân rã, tất cả hoặc một số các đặc tính được tính toán chỉ lên đến vị trí của sự chuyển tiếp (

$i_{\max} - 3$) vì độ méo trong phần không hài hòa của phổ được đưa vào bởi sự lọc LTP sẽ được triệt bởi sự tạo mạng chắn của sự chuyển tiếp kéo dài mạnh (ví dụ, tiếng chũm chọe trung).

Chuỗi xung đối với các tín hiệu bước thấp có thể được phát hiện như sự chuyển tiếp bởi bộ phát hiện chuyển tiếp. Đối với các tín hiệu với bước thấp, các đặc tính từ bộ phát hiện chuyển tiếp sau đó được bỏ qua và thay vào đó là ngưỡng bổ sung đối với sự tương quan được chuẩn hóa mà phụ thuộc vào độ trễ bước, ví dụ:

Nếu $\text{norm_corr} \leq 1,2 - T_{\text{int}}/L$, khi đó vô hiệu hóa bộ lọc.

Một quyết định ví dụ được thể hiện dưới đây khi đó $b1$ là tốc độ truyền dữ liệu nào đó, ví dụ 48kbps, trong đó TCX_20 biểu thị rằng khung được mã hóa sử dụng khối dài đơn, trong đó TCX_10 biểu thị rằng khung được mã hóa sử dụng 2, 3, 4 hoặc nhiều các khối ngắn, trong đó sự quyết định $\text{TCX_20}/\text{TCX_10}$ được dựa trên đầu ra của bộ phát hiện chuyển tiếp được mô tả nêu trên. tempFlatness là số đo độ phẳng theo thời gian như được định nghĩa trong (6), maxEnergyChange là sự thay đổi năng lượng tối đa như được định nghĩa trong (7). Điều kiện $\text{norm_corr}(\text{curr}) > 1,2 - T_{\text{int}}/L$ có thể cũng được viết như $(1,2 - \text{norm_corr}(\text{curr})) * L < T_{\text{int}}$.

```
enableLTP =
(bitrate < b1 && tcxMode == TCX_20 && (norm_corr(curr) * norm_corr(prev)) > 0,25 && tempFlatness < 3,5) ||
(bitrate >= b1 && tcxMode == TCX_10 && max(norm_corr(curr), norm_corr(prev)) > 0,5 && maxEnergyChange < 3,5) ||
(bitrate >= b1 && norm_corr(curr) > 0,44 && norm_corr(curr) > 1,2 - T_int/L) ||
(bitrate >= b1 && tcxMode == TCX_20 && norm_corr(curr) > 0,44 &&
(tempFlatness < 6,0 || (tempFlatness < 7,0 && maxEnergyChange < 22,0)));
( bitrate >= b1 && tcxMode == TCX_20 && norm_corr > 0,44 &&
```

Rõ ràng từ các ví dụ nêu trên, sự phát hiện sự chuyển tiếp ảnh hưởng đến cơ cấu quyết định nào cho sự dự báo dài hạn sẽ được sử dụng và phần nào của tín hiệu sẽ được sử dụng cho sự đo lường được sử dụng trong quyết định, và không trực tiếp gây ra sự vô hiệu hóa bộ lọc dự báo dài hạn.

Các số đo theo thời gian được sử dụng cho quyết định độ dài biến đổi có thể hoàn toàn khác với các số đo theo thời gian được sử dụng cho quyết định bộ lọc LTP hoặc chúng có thể chồng lấp hoặc giống một cách chính xác nhưng được tính toán trong các

vùng khác nhau. Đối với các tín hiệu bước thấp, sự phát hiện các sự chuyển tiếp có thể được bỏ qua hoàn toàn nếu ngưỡng cho sự tương quan được chuẩn hóa mà phụ thuộc vào độ trễ bước đạt được.

Công nghệ loại bỏ sự gián đoạn có khả năng

Công nghệ khả thi để loại bỏ sự gián đoạn có thể được gây ra bằng cách áp dụng bộ lọc tuyến tính $H(z)$ khung cạnh khung bây giờ được mô tả. Bộ lọc tuyến tính có thể là bộ lọc LTP được mô tả. Bộ lọc tuyến tính có thể là bộ lọc đáp ứng xung hữu hạn (finite impulse response - FIR) hoặc bộ lọc đáp ứng xung vô hạn (infinite impulse response - IIR). Phương pháp đề xuất không lọc phần của khung hiện thời với các tham số bộ lọc của khung trước, và do đó tránh được các vấn đề có khả năng của các phương pháp đã biết. Phương pháp được đề xuất sử dụng bộ lọc LPC để loại bỏ sự gián đoạn. Bộ lọc LPC được ước lượng trên tín hiệu âm thanh (được lọc bởi bộ lọc bất biến thời gian tuyến tính $H(z)$ hoặc không) và do đó là mẫu tốt của hình dạng phổ của tín hiệu âm thanh (được lọc bởi $H(z)$ hoặc không). Bộ lọc LPC sau đó được sử dụng sao cho hình dạng phổ của tín hiệu âm thanh tạo mạng chắn sự gián đoạn.

Bộ lọc LPC có thể được ước lượng bằng các cách khác nhau. Nó có thể được ước lượng ví dụ, sử dụng tín hiệu âm thanh (khung hiện thời và/hoặc khung trước) và thuật toán Levinson-Durbin. Nó cũng có thể được tính toán trên tín hiệu khung được lọc trước, sử dụng thuật toán Levinson-Durbin.

Nếu $H(z)$ được sử dụng trong bộ mã hóa-giải mã âm thanh và bộ mã hóa-giải mã âm thanh đã sử dụng bộ lọc LPC (được lượng tử hóa hay không) để ví dụ, tạo hình nhiễu âm lượng tử hóa trong bộ mã hóa-giải mã âm thanh trên cơ sở biến đổi, khi đó bộ lọc LPC này có thể được sử dụng trực tiếp để làm nhẵn sự gián đoạn, mà không có sự phức tạp bổ sung nào cần thiết để ước lượng bộ lọc LPC mới.

Dưới đây mô tả sự xử lý khung hiện thời cho trường hợp bộ lọc FIR và trường hợp bộ lọc IIR. Khung trước được giả định là đã được xử lý.

Trường hợp bộ lọc FIR:

1. Lọc khung hiện thời với các tham số bộ lọc của khung hiện thời, tạo ra khung hiện thời được lọc.

2. Xem xét bộ lọc LPC (được lượng tử hóa hay không) với thứ tự M , được ước lượng trên tín hiệu âm thanh (được lọc hay không).
3. M mẫu cuối của khung trước được lọc với bộ lọc $H(z)$ và các hệ số của khung hiện thời, tạo ra phần thứ nhất của tín hiệu được lọc.
4. M mẫu cuối của khung trước được lọc sau đó được trừ từ phần thứ nhất của tín hiệu được lọc, tạo ra phần thứ hai của tín hiệu được lọc.
5. Đáp ứng xung bằng 0 (Zero Impulse Response - ZIR) của bộ lọc LPC sau đó được tạo ra bằng cách lọc khung của các mẫu 0 với bộ lọc LPC và các tình trạng ban đầu bằng với phần thứ hai của tín hiệu được lọc.
6. ZIR có thể được tạo ra một cách tùy chọn sao cho biên độ của nó tiến nhanh đến 0.
7. Phần mở đầu của ZIR được trừ từ phần bắt đầu tương ứng của khung hiện thời được lọc.

Trường hợp bộ lọc IIR:

1. Xem xét bộ lọc LPC (được lượng tử hóa hay không) với thứ tự M , được ước lượng trên tín hiệu âm thanh (được lọc hay không).
2. M mẫu cuối của khung trước được lọc với bộ lọc $H(z)$ và các hệ số của khung hiện thời, tạo ra phần thứ nhất của tín hiệu được lọc.
3. M mẫu cuối của khung trước được lọc sau đó được trừ từ phần thứ nhất của tín hiệu được lọc, tạo ra phần thứ hai của tín hiệu được lọc.
4. Đáp ứng xung bằng 0 (Zero Impulse Response - ZIR) của bộ lọc LPC sau đó được tạo ra bằng cách lọc khung của các mẫu 0 với bộ lọc LPC và các tình trạng ban đầu bằng với phần thứ hai của tín hiệu được lọc.
5. ZIR có thể được tạo ra một cách tùy chọn sao cho biên độ của nó tiến nhanh đến 0.
6. Phần bắt đầu của khung hiện thời sau đó được xử lý mẫu cạnh mẫu bắt đầu với mẫu thứ nhất của khung hiện thời.

7. Mẫu được lọc với bộ lọc $H(z)$ và các tham số khung hiện thời, tạo ra mẫu được lọc thứ nhất.
8. Mẫu tương ứng của ZIR sau đó được trừ từ mẫu được lọc thứ nhất, sản xuất mẫu tương ứng của khung hiện thời được lọc.
9. Chuyển đến mẫu tiếp theo.
10. Lặp lại từ 9 đến 12 cho đến khi mẫu cuối cùng của phần bắt đầu của khung hiện thời được xử lý.
11. Lọc các mẫu còn lại của khung hiện thời với các tham số bộ lọc của khung hiện thời.

Theo đó, các phương án của sáng chế cho phép ước lượng các SNR phân đoạn và sự lựa chọn của thuật toán mã hóa thích hợp theo cách đơn giản và chính xác. Cụ thể, các phương án của sáng chế cho phép sự lựa chọn chu trình mở của thuật toán mã hóa thích hợp, trong đó sự lựa chọn không thích hợp của thuật toán mã hóa trong trường hợp tín hiệu âm thanh có các sóng hài được tránh.

Trong các phương án nêu trên, các SNR phân đoạn được ước lượng bằng cách tính toán giá trị trung bình của các SNR được ước lượng cho các khung con tương ứng. Trong các phương án thay thế, SNR của toàn bộ khung có thể được ước lượng mà không chia khung thành các khung con.

Các phương án của sáng chế cho phép giảm mạnh về thời gian tính toán khi được so sánh với sự lựa chọn chu trình đóng vì số lượng các bước được yêu cầu trong sự lựa chọn chu trình đóng được loại bỏ.

Theo đó, số lượng lớn các bước và thời gian tính toán được kết hợp với nó có thể được lưu bởi phương pháp theo sáng chế trong khi vẫn cho phép sự lựa chọn của thuật toán mã hóa phù hợp với sự thực hiện tốt.

Mặc dù một số khía cạnh đã được mô tả trong bối cảnh của thiết bị, rõ ràng rằng các khía cạnh này cũng thể hiện sự mô tả của phương pháp tương ứng, trong đó khối hoặc thiết bị tương ứng với bước phương pháp hoặc đặc tính của bước phương pháp. Tương tự, các khía cạnh được mô tả trong ngữ cảnh của bước phương pháp cũng biểu diễn sự mô tả của khối hoặc mục hoặc dấu hiệu tương ứng của thiết bị tương ứng.

Các phương án của các thiết bị được mô tả trong đây và các đặc tính của nó có thể được thực hiện bởi máy tính, một hoặc nhiều bộ xử lý, một hoặc nhiều bộ vi xử lý, các mảng cổng có thể lập trình trường (field-programmable gate array - FPGA), các mạch tích hợp cụ thể ứng dụng (application specific integrated circuit - ASIC) và tương tự hoặc tổ hợp của chúng, mà được tạo cấu hình hoặc được lập trình để cung cấp các chức năng đã được mô tả.

Một vài hoặc tất cả các bước phương pháp có thể được thực hiện bởi (hoặc sử dụng) thiết bị phần cứng, tương tự ví dụ, bộ vi xử lý, máy tính có thể lập trình hoặc mạch điện tử. Theo một số phương án, một số một hoặc nhiều bước phương pháp quan trọng nhất có thể được thực hiện bởi thiết bị như vậy.

Phụ thuộc vào các yêu cầu thực hiện nhất định, các phương án của sáng chế có thể được thực hiện trong phần cứng hoặc trong phần mềm. Phương án có thể được thực hiện sử dụng vật ghi lưu trữ không tạm thời như vật ghi lưu trữ dạng số, ví dụ, đĩa mềm, DVD, Blu-Ray, CD, ROM, PROM, và EPROM, EEPROM hoặc bộ nhớ FLASH, có các tín hiệu điều khiển có thể đọc được bằng điện tử được lưu trữ trên đó, mà kết hợp (hoặc có khả năng kết hợp) với hệ thống máy tính có thể lập trình để phương pháp tương ứng được thực hiện. Do đó, vật ghi lưu trữ dạng số có thể là máy tính có thể đọc được.

Một số phương án theo sáng chế bao gồm vật mang dữ liệu có các tín hiệu điều khiển có thể đọc được bằng điện tử, mà có khả năng kết hợp với hệ thống máy tính có thể lập trình được, sao cho một trong số các phương pháp được mô tả ở đây được thực hiện.

Thông thường, các phương án theo sáng chế có thể được thực hiện như sản phẩm chương trình máy tính với mã chương trình, mã chương trình hoạt động để thực hiện một trong các phương pháp khi sản phẩm chương trình máy tính chạy trên máy tính. Ví dụ, mã chương trình có thể được lưu trữ trên vật mang có thể đọc được bằng máy.

Các phương án khác bao gồm chương trình máy tính để thực hiện một trong các phương pháp được mô tả ở đây, được lưu trữ trên vật mang có thể đọc được bằng máy.

Nói cách khác, một phương án của phương pháp theo sáng chế là, chương trình máy tính có mã chương trình để thực hiện một trong các phương pháp được mô tả ở đây, khi chương trình máy tính chạy trên máy tính.

Do đó, phương án nữa của phương pháp theo sáng chế là, vật mang dữ liệu (hoặc vật ghi lưu trữ số, hoặc vật ghi có thể đọc được bằng máy tính) bao gồm, đã được ghi lại trên đó, chương trình máy tính để thực hiện một trong các phương pháp được mô tả ở đây. Vật mang dữ liệu, vật ghi lưu trữ số hoặc vật ghi được ghi thường là hữu hình và/hoặc không tạm thời.

Do đó, phương án nữa của phương pháp theo sáng chế là, dòng dữ liệu hoặc chuỗi các tín hiệu thể hiện chương trình máy tính để thực hiện một trong các phương pháp được mô tả ở đây. Ví dụ, dòng dữ liệu hoặc chuỗi các tín hiệu có thể, ví dụ, được tạo cấu hình để được truyền thông qua sự kết nối truyền thông dữ liệu, ví dụ, thông qua liên mạng.

Phương án nữa bao gồm phương tiện xử lý, ví dụ, máy tính, hoặc thiết bị logic có thể lập trình được tạo cấu hình để, hoặc được lập trình để, thực hiện một trong các phương pháp được mô tả ở đây.

Phương án nữa bao gồm máy tính có chương trình máy tính được cài đặt trên máy tính để thực hiện một trong các phương pháp được mô tả ở đây.

Phương án khác theo sáng chế gồm có thiết bị hoặc hệ thống được tạo cấu hình để truyền (ví dụ, bằng điện tử hoặc quang học) chương trình máy tính để thực hiện một trong các phương pháp được mô tả ở đây đến bộ thu. Máy thu nhận có thể, ví dụ, là máy tính, thiết bị di động, thiết bị nhớ hoặc tương tự. Thiết bị hoặc hệ thống có thể, ví dụ, gồm có máy chủ tập tin để truyền chương trình máy tính đến bộ thu.

Trong một số phương án, thiết bị logic có thể lập trình được (ví dụ, mảng cổng có thể lập trình được dạng trường) có thể được sử dụng để thực hiện một số hoặc tất cả các chức năng của các phương pháp được mô tả ở đây. Theo một số phương án, mảng cổng lập trình được dạng trường có thể kết hợp với bộ vi xử lý để thực hiện một trong số các phương pháp được mô tả ở đây. Thông thường, các phương pháp ưu tiên được thực hiện bởi thiết bị phần cứng bất kỳ.

Các phương án được mô tả bên trên chỉ mang tính minh họa cho các nguyên lý của sáng chế. Cần hiểu rằng các sự biến thể và biến đổi của các phương án và các chi tiết được mô tả ở đây sẽ là rõ ràng đối với người có trình độ trong lĩnh vực kỹ thuật tương ứng. Do đó, mục đích chỉ bị giới hạn bởi phạm vi của các điểm yêu cầu bảo hộ sắp đưa ra dưới đây và không bị giới hạn bởi các chi tiết cụ thể được biểu diễn bằng cách mô tả và giải thích của các phương án ở đây.

YÊU CẦU BẢO HỘ

1. Thiết bị lựa chọn một trong số thuật toán mã hóa thứ nhất có đặc điểm thứ nhất và thuật toán mã hóa thứ hai có đặc điểm thứ hai (10) mã hóa một phần của tín hiệu âm thanh (40) để thu được phiên bản được mã hóa của phần của tín hiệu âm thanh (40), thiết bị bao gồm:

bộ lọc dự báo dài hạn được tạo cấu hình để nhận tín hiệu âm thanh, để giảm biên độ của các sóng hài trong tín hiệu âm thanh và để phát ra phiên bản được lọc của tín hiệu âm thanh;

bộ ước lượng thứ nhất (12) để sử dụng phiên bản được lọc của tín hiệu âm thanh trong phép ước lượng tỉ số tín hiệu trên nhiễu (signal to noise ratio - SNR) hoặc SNR phân đoạn của phần của tín hiệu âm thanh như số đo chất lượng thứ nhất cho phần tín hiệu âm thanh, số đo chất lượng thứ nhất được kết hợp với thuật toán mã hóa thứ nhất, trong đó phép ước lượng số đo chất lượng thứ nhất đã nêu bao gồm thực hiện phép làm gần đúng thuật toán mã hóa thứ nhất để thu được ước lượng độ méo của thuật toán mã hóa thứ nhất và để ước lượng số đo chất lượng thứ nhất dựa trên phần của tín hiệu âm thanh và phép ước lượng độ méo của thuật toán mã hóa thứ nhất mà không thực sự mã hóa và giải mã phần của tín hiệu âm thanh sử dụng thuật toán mã hóa thứ nhất;

bộ ước lượng thứ hai (14) để ước lượng SNR hoặc SNR phân đoạn như số đo chất lượng thứ hai cho phần của tín hiệu âm thanh, số đo chất lượng thứ hai được kết hợp với thuật toán mã hóa thứ hai, trong đó phép ước lượng số đo chất lượng thứ hai đã nêu bao gồm thực hiện phép làm gần đúng thuật toán mã hóa thứ hai để thu được phép ước lượng độ méo của thuật toán mã hóa thứ hai và để ước lượng số đo chất lượng thứ hai sử dụng phần của tín hiệu âm thanh và phép ước lượng độ méo của thuật toán mã hóa thứ hai mà không thực sự mã hóa và giải mã phần của tín hiệu âm thanh sử dụng thuật toán mã hóa thứ hai; và

bộ điều khiển (16) để lựa chọn thuật toán mã hóa thứ nhất hoặc thuật toán mã hóa thứ hai dựa trên phép so sánh giữa số đo chất lượng thứ nhất và số đo chất lượng thứ hai,

trong đó thuật toán mã hóa thứ nhất là thuật toán mã hóa biến đổi, thuật toán mã hóa trên cơ sở phép biến đổi cosin rời rạc cải biên (modified discrete cosine transform - MDCT) hoặc thuật toán mã hóa trên cơ sở kích thích mã hóa biến đổi (transform coding excitation - TCX) và trong đó thuật toán mã hóa thứ hai là thuật toán mã hóa dự báo tuyến tính được kích thích mã (code excited linear prediction - CELP) hoặc thuật toán mã hóa dự báo tuyến tính được kích thích mã đại số (algebraic code excited linear prediction - ACELP).

2. Thiết bị (10) theo điểm 1, trong đó hàm chuyển của bộ lọc dự báo dài hạn bao gồm một phần nguyên của độ trễ bước và bộ lọc đa nhánh phụ thuộc vào phần phân số của độ trễ bước.

3. Thiết bị (10) theo điểm 1, trong đó bộ lọc dự báo dài hạn có hàm chuyển:

$$P(z) = 1 - \beta g B(z, T_{fr}) z^{-T_{int}}$$

với T_{int} và T_{fr} là số nguyên và phần phân số của độ trễ bước, g là độ khuếch đại, β là trọng số và $B(z, T_{fr})$ là bộ lọc thông thấp FIR mà các hệ số của bộ lọc này phụ thuộc vào phần phân số của bước.

4. Thiết bị theo một trong số các điểm từ 1 đến 3, còn bao gồm bộ phận vô hiệu hóa để vô hiệu hóa bộ lọc dựa trên sự tổ hợp của một hoặc nhiều số đo sự hài hòa và/hoặc một hoặc nhiều số đo kết cấu theo thời gian.

5. Thiết bị theo điểm 4, trong đó một hoặc nhiều số đo độ hài hòa bao gồm ít nhất một trong số sự tương quan được chuẩn hóa hoặc độ khuếch đại dự báo và trong đó một hoặc nhiều số đo kết cấu theo thời gian bao gồm ít nhất một số đo độ phẳng theo thời gian và sự thay đổi năng lượng.

6. Thiết bị theo một trong số các điểm từ 1 đến 5, trong đó bộ lọc được áp dụng cho tín hiệu âm thanh trên cơ sở khung cạnh khung, thiết bị nêu trên còn bao gồm bộ phận để loại bỏ sự gián đoạn trong tín hiệu âm thanh được gây ra bởi bộ lọc.

7. Thiết bị (10) theo một trong số các điểm từ 1 đến 6, trong đó các bộ ước lượng thứ nhất và thứ hai được tạo cấu hình để ước lượng SNR hoặc SNR phân đoạn của phần của phiên bản được gán trọng số của tín hiệu âm thanh.

8. Thiết bị (10) theo một trong số các điểm từ 1 đến 7, trong đó bộ ước lượng thứ nhất (12) được tạo cấu hình để xác định độ méo bộ lượng tử hóa được ước lượng mà bộ lượng tử hóa được sử dụng trong thuật toán mã hóa thứ nhất sẽ đưa vào khi lượng tử hóa phần của tín hiệu âm thanh và để ước lượng số đo chất lượng thứ nhất dựa trên năng lượng của phần của phiên bản được gán trọng số của tín hiệu âm thanh và độ méo bộ lượng tử hóa được ước lượng, trong đó bộ ước lượng thứ nhất (12) được tạo cấu hình để ước lượng độ khuếch đại toàn cục cho phần của tín hiệu âm thanh sao cho phần của tín hiệu âm thanh sẽ tạo ra tốc độ truyền dữ liệu đích đã cho khi được mã hóa với bộ lượng tử hóa và bộ mã hóa entropi được sử dụng trong thuật toán mã hóa thứ nhất, trong đó bộ ước lượng thứ nhất (12) còn được tạo cấu hình để xác định độ méo bộ lượng tử hóa được ước lượng dựa trên độ khuếch đại toàn cục được ước lượng.

9. Thiết bị (10) theo một trong số các điểm từ 1 đến 8, trong đó bộ ước lượng thứ hai (14) được tạo cấu hình để xác định độ méo bảng mã thích ứng được ước lượng mà bảng mã thích ứng được sử dụng trong thuật toán mã hóa thứ hai sẽ đưa vào khi sử dụng bảng mã thích ứng để mã hóa phần của tín hiệu âm thanh, và trong đó bộ ước lượng thứ hai (14) được tạo cấu hình để ước lượng số đo chất lượng thứ hai dựa trên năng lượng của phần của phiên bản được gán trọng số của tín hiệu âm thanh và độ méo bảng mã thích ứng, trong đó, đối với mỗi phần con trong số nhiều phần con của phần của tín hiệu âm thanh, bộ ước lượng thứ hai (14) được tạo cấu hình để làm gần đúng bảng mã thích ứng dựa trên phiên bản của phần con của tín hiệu âm thanh được gán trọng số được chuyển dịch về trước bởi độ trễ bước được xác định trong giai đoạn xử lý trước, để ước lượng độ khuếch đại bảng mã thích ứng sao cho sai số giữa phần con của phần của tín hiệu âm thanh được gán trọng số và bảng mã thích ứng được làm gần đúng được tối thiểu hóa, và để xác định độ méo bảng mã thích ứng được ước lượng dựa trên năng lượng của sai số giữa phần con của phần của tín hiệu được gán trọng số và bảng mã thích ứng được làm gần đúng được ước lượng bởi độ khuếch đại bảng mã thích ứng.

10. Thiết bị (10) theo điểm 9, trong đó bộ ước lượng thứ hai (14) còn được tạo cấu hình để giảm độ méo bảng mã thích ứng được ước lượng được xác định cho mỗi phần con của phần của tín hiệu âm thanh bằng thừa số không đổi.

11. Thiết bị (10) theo một trong số các điểm từ 1 đến 8, trong đó bộ ước lượng thứ hai (14) được tạo cấu hình để xác định độ méo bảng mã thích ứng được ước lượng mà bảng mã thích ứng được sử dụng trong thuật toán mã hóa thứ hai sẽ đưa vào khi sử dụng bảng mã thích ứng để mã hóa phần của tín hiệu âm thanh, và trong đó bộ ước lượng thứ hai (14) được tạo cấu hình để ước lượng số đo chất lượng thứ hai dựa trên năng lượng của phần của phiên bản được gán trọng số của tín hiệu âm thanh và độ méo bảng mã thích ứng, trong đó bộ ước lượng thứ hai (14) được tạo cấu hình để làm gần đúng bảng mã thích ứng dựa trên phiên bản của phần của tín hiệu âm thanh được gán trọng số được chuyển dịch về trước bởi độ trễ bước được xác định trong giai đoạn xử lý trước, để ước lượng độ khuếch đại bảng mã thích ứng sao cho sai số giữa phần của tín hiệu âm thanh được gán trọng số và bảng mã thích ứng được làm gần đúng được tối thiểu hóa, và để xác định độ méo bảng mã thích ứng được ước lượng dựa trên năng lượng của sai số giữa phần của tín hiệu được gán trọng số và bảng mã thích ứng được làm gần đúng được ước lượng bởi độ khuếch đại bảng mã thích ứng.

12. Thiết bị mã hóa phần của tín hiệu âm thanh (20), bao gồm thiết bị (10) theo một điểm trong số các điểm từ 1 đến 11, giai đoạn bộ mã hóa thứ nhất (26) để thực hiện thuật toán mã hóa thứ nhất và giai đoạn bộ mã hóa thứ hai (28) để thực hiện thuật toán mã hóa thứ hai, trong đó thiết bị mã hóa (20) được tạo cấu hình để mã hóa phần của tín hiệu âm thanh sử dụng thuật toán mã hóa thứ nhất hoặc thuật toán mã hóa thứ hai phụ thuộc vào sự lựa chọn bởi bộ điều khiển (16).

13. Hệ thống mã hóa và giải mã bao gồm thiết bị (20) để mã hóa theo điểm 12 và bộ giải mã được tạo cấu hình để nhận phiên bản được mã hóa của phần của tín hiệu âm thanh và sự chỉ thị của thuật toán được sử dụng để mã hóa phần của tín hiệu âm thanh và để giải mã phiên bản được mã hóa của phần của tín hiệu âm thanh sử dụng thuật toán được chỉ thị.

14. Phương pháp lựa chọn một trong số thuật toán mã hóa thứ nhất có đặc điểm thứ nhất và thuật toán mã hóa thứ hai có đặc điểm thứ hai để mã hóa một phần của tín hiệu âm thanh để thu được phiên bản được mã hóa của phần của tín hiệu âm thanh, phương pháp bao gồm:

lọc tín hiệu âm thanh sử dụng bộ lọc dự báo dài hạn để giảm biên độ của các sóng hài trong tín hiệu âm thanh và để phát ra phiên bản được lọc của tín hiệu âm thanh;

sử dụng phiên bản được lọc của tín hiệu âm thanh trong phép ước lượng SNR hoặc SNR được phân đoạn của phần của tín hiệu âm thanh như số đo chất lượng thứ nhất cho phần của tín hiệu âm thanh, số đo chất lượng thứ nhất được kết hợp với thuật toán mã hóa thứ nhất, trong đó ước lượng số đo chất lượng thứ nhất đã nêu bao gồm thực hiện phép làm gần đúng thuật toán mã hóa thứ nhất để thu được phép ước lượng độ méo của thuật toán mã hóa thứ nhất và để ước lượng số đo chất lượng thứ nhất dựa trên phần của tín hiệu âm thanh thứ nhất và phép ước lượng độ méo của thuật toán mã hóa thứ nhất mà không thực sự mã hóa và giải mã phần của tín hiệu âm thanh sử dụng thuật toán mã hóa thứ nhất;

ước lượng SNR hoặc SNR được phân đoạn như số đo chất lượng thứ hai cho phần của tín hiệu âm thanh, số đo chất lượng thứ hai được kết hợp với thuật toán mã hóa thứ hai, trong đó ước lượng số đo chất lượng thứ hai đã nêu bao gồm thực hiện phép làm gần đúng thuật toán mã hóa thứ hai để thu được phép ước lượng độ méo của thuật toán mã hóa thứ hai và để ước lượng số đo chất lượng thứ hai sử dụng phần của tín hiệu âm thanh và phép ước lượng độ méo của thuật toán mã hóa thứ hai mà không thực sự mã hóa và giải mã phần của tín hiệu âm thanh sử dụng thuật toán mã hóa thứ hai; và

lựa chọn thuật toán mã hóa thứ nhất hoặc thuật toán mã hóa thứ hai dựa trên phép so sánh giữa số đo chất lượng thứ nhất và số đo chất lượng thứ hai,

trong đó thuật toán mã hóa thứ nhất là thuật toán mã hóa biến đổi, thuật toán mã hóa trên cơ sở phép biến đổi cosin rời rạc cải biên (modified discrete cosine transform - MDCT) hoặc thuật toán mã hóa trên cơ sở kích thích mã hóa biến đổi (transform coding excitation - TCX) và trong đó thuật toán mã hóa thứ hai là thuật toán mã hóa dự báo tuyến tính được kích thích (code excited linear prediction - CELP) hoặc thuật toán mã hóa dự báo tuyến tính được kích thích mã đại số (algebraic code excited linear prediction - ACELP).

15. Vật ghi có thể đọc được bằng máy tính bao gồm chương trình máy tính có mã chương trình để thực hiện, khi chạy trên máy tính, phương pháp theo điểm 14.

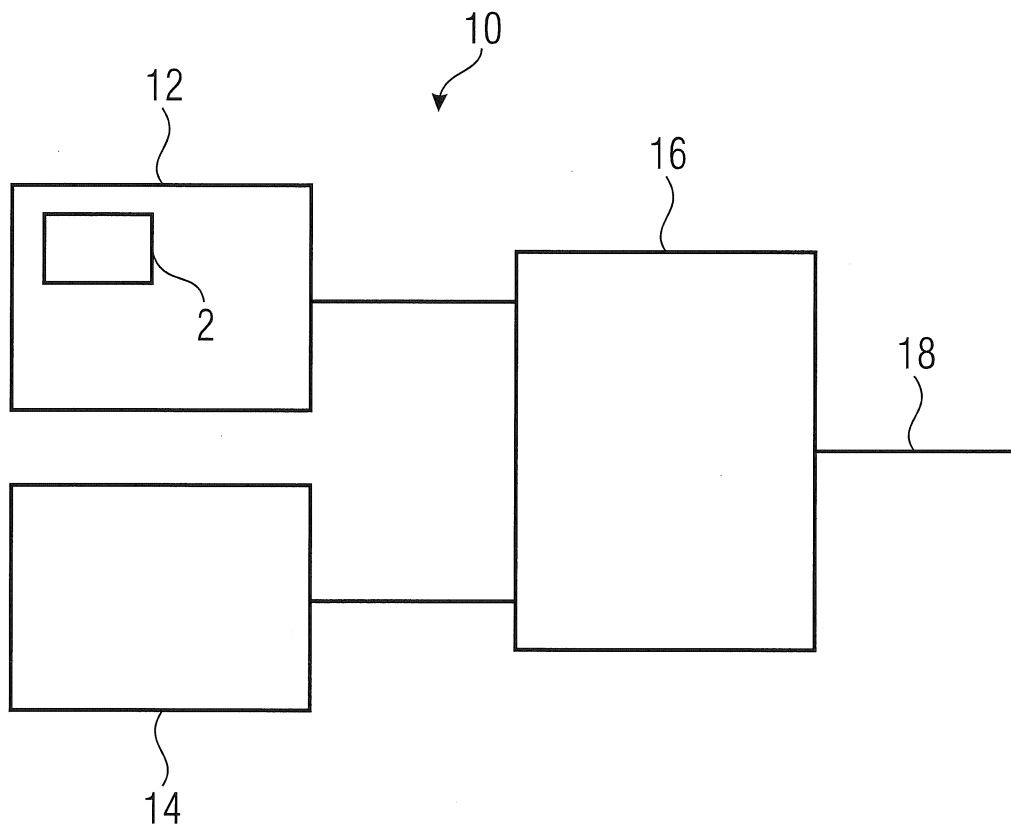


FIG 1

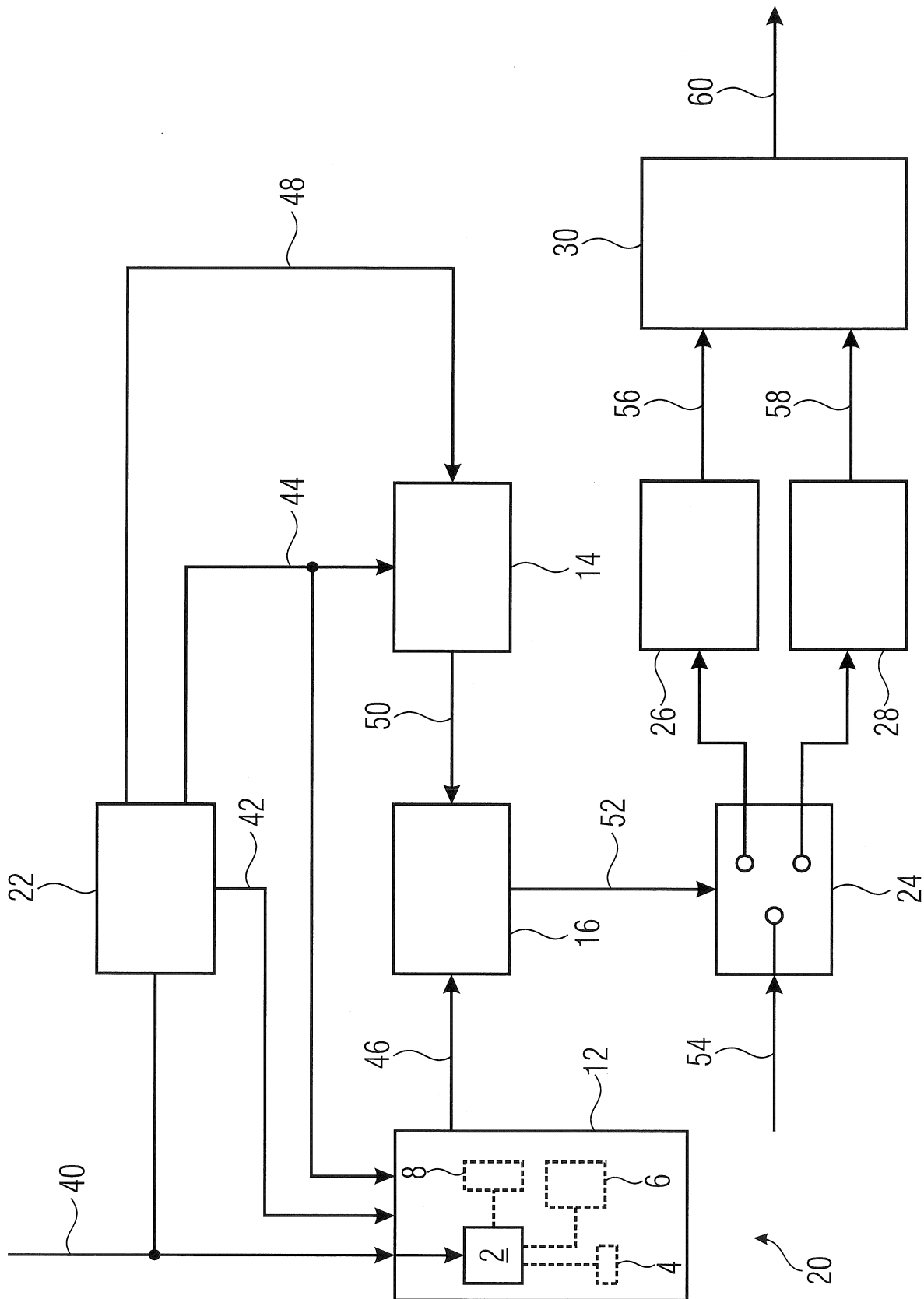


FIG 2

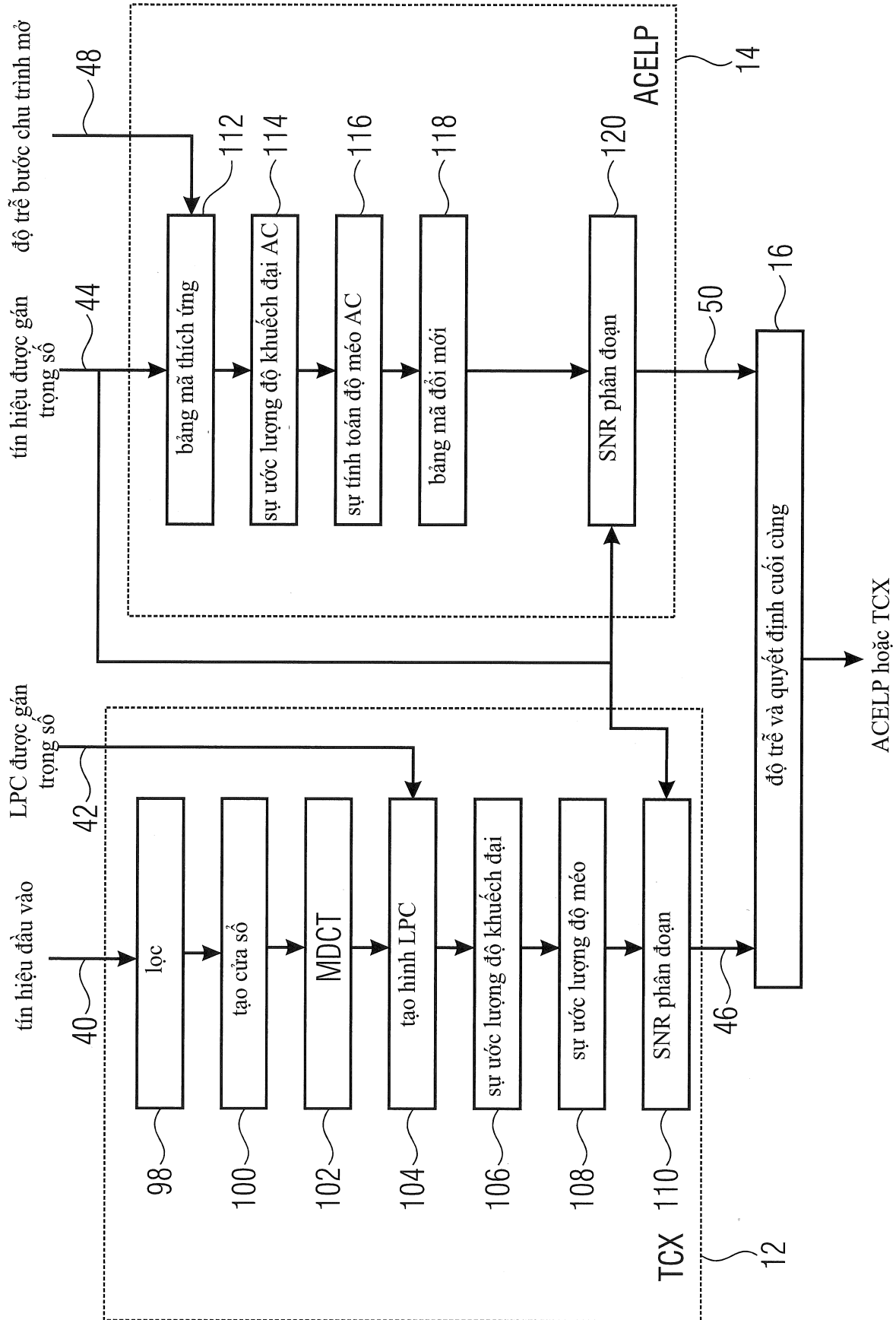


FIG 3

$$\text{SNR} = 10 \log_{10} \frac{\sum_{i=1}^N x^2(i)}{\sum_{i=1}^N (x(i) - y(i))^2}$$

FIG 4A

$$\text{SNRseg} = \frac{10}{M} \sum_{m=0}^{M-1} \log_{10} \sum_{i=Nm}^{Nm+N-1} \left(\frac{\sum_{i=1}^N x^2(i)}{\sum_{i=1}^N (x(i) - y(i))^2} \right)$$

FIG 4B