



(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)



1-0049968

(51)^{2022.01} G10L 21/04

(13) B

(21) 1-2023-01997

(22) 18/06/2014

(62) 1-2016-00192

(86) PCT/EP2014/062833 18/06/2014

(87) WO 2014/202672 A3 24/12/2014

(30) 13173159.8 21/06/2013 EP; 14167055.4 05/05/2014 EP

(45) 25/08/2025 449

(43) 25/07/2023 424A

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

(72) REUSCHL, Stefan (DE); DOEHLA, Stefan (DE); LECOMTE, Jérémie (FR);
JANDER, Manuel (DE); FAERBER, Nikolaus (DE).

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

(54) BỘ ĐỊNH TỈ LỆ THỜI GIAN, BỘ GIẢI MÃ ÂM THANH VÀ PHƯƠNG PHÁP
CUNG CẤP PHIÊN BẢN ĐƯỢC ĐỊNH TỈ LỆ THỜI GIAN CỦA TÍN HIỆU ÂM
THANH ĐẦU VÀO

(21) 1-2023-01997

(57) Sáng chế đề cập đến bộ định tỉ lệ thời gian, bộ giải mã âm thanh và phương pháp cung cấp phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào. Bộ định tỉ lệ thời gian để cung cấp phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào được tạo cấu hình để tính toán hoặc ước lượng chất lượng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bởi sự định tỉ lệ thời gian của tín hiệu âm thanh đầu vào. Bộ định tỉ lệ thời gian được tạo cấu hình để thực hiện sự định tỉ lệ thời gian của tín hiệu âm thanh đầu vào phụ thuộc vào sự tính toán hoặc sự ước lượng chất lượng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bởi sự định tỉ lệ thời gian. Bộ giải mã âm thanh bao gồm bộ định tỉ lệ thời gian như vậy

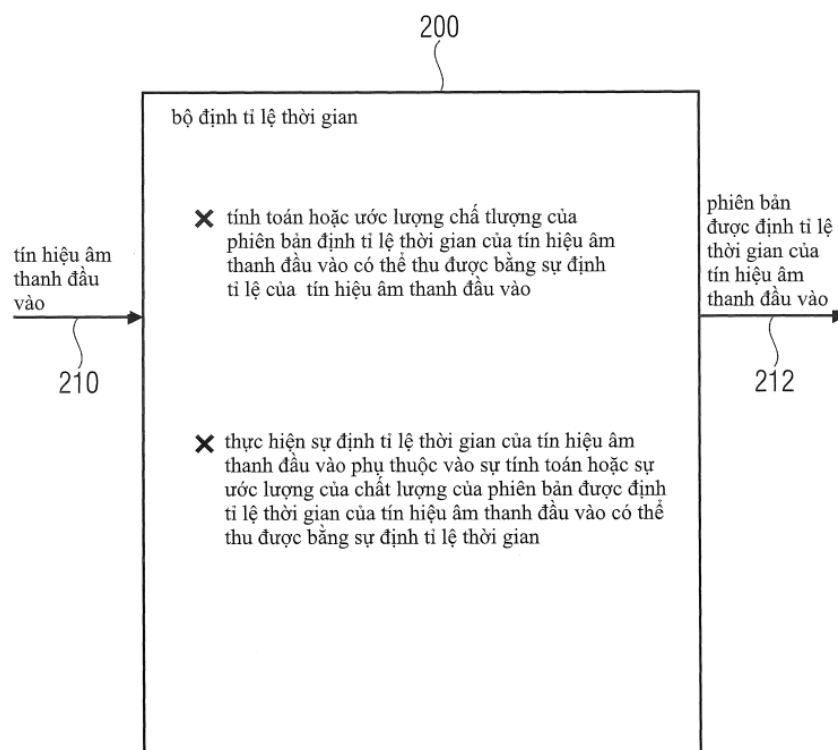


FIG 2

Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến bộ định tỉ lệ thời gian để cung cấp phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào.

Sáng chế còn đề cập đến bộ giải mã âm thanh để cung cấp nội dung âm thanh được giải mã trên cơ sở nội dung âm thanh đầu vào.

Sáng chế còn đề cập đến phương pháp cung cấp phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào.

Sáng chế còn đề cập đến chương trình máy tính để thực hiện các phương pháp đã nêu.

Tình trạng kỹ thuật của sáng chế

Sự lưu trữ và truyền dẫn nội dung âm thanh (bao gồm nội dung âm thanh chung, như nội dung âm nhạc, nội dung tiếng nói và nội dung âm thanh/tiếng nói chung được trộn) là một lĩnh vực kỹ thuật quan trọng. Thách thức cụ thể gây ra bởi thực tế rằng người nghe mong muốn sự phát lại liên tục các nội dung âm thanh, mà không có sự gián đoạn và cũng không có các thành phần lạ có thể nghe được gây ra bởi sự lưu trữ và sự truyền dẫn nội dung âm thanh. Cùng lúc đó, mong muốn để giữ các yêu cầu đối với các phương tiện lưu trữ và các phương tiện truyền dẫn dữ liệu càng thấp càng tốt, để giữ chi phí trong giới hạn có thể chấp nhận.

Các vấn đề nảy sinh, ví dụ, nếu việc đọc ra từ vật ghi lưu trữ bị gián đoạn hoặc bị trễ tạm thời, hoặc nếu sự truyền dẫn giữa nguồn dữ liệu và bộ chứa dữ liệu bị gián đoạn hoặc bị trễ tạm thời. Ví dụ, sự truyền dẫn qua liên mạng không có độ tin cậy cao, vì các gói TCP/IP có thể bị mất, và vì độ trễ truyền dẫn qua liên mạng có thể biến thiên, ví dụ, phụ thuộc vào tình trạng tải biến thiên của các nút liên mạng. Tuy nhiên, nó được yêu cầu, để có được trải nghiệm hài lòng người sử dụng, rằng có sự phát lại liên tục nội dung âm thanh, mà không có các "khoảng trống" có thể nghe được hoặc các thành phần lạ có thể nghe được. Hơn nữa, nó được mong muốn tránh khỏi các độ

trễ đáng kể mà sẽ bị gây ra bởi việc đặt vào vùng đệm một lượng lớn thông tin âm thanh.

Xem xét sự thảo luận như trên, có thể nhận ra rằng nhu cầu khái niệm mà cung cấp chất lượng âm thanh tốt, thậm chí trong trường hợp sự cung cấp gián đoạn thông tin âm thanh.

Tài liệu tham khảo

[Lia01] Y. J. Liang, N. Faerber, B. Girod: “Adaptive playout scheduling using time-scale modification in packet voice communications”, 2001

[Lef03] P. Gournay, F. Rousseau, R. Lefebvre: “Improved packet loss recovery using late frames for prediction-based speech coders”, 2003

Bản chất kỹ thuật của sáng chế

Phương án theo sáng chế tạo ra bộ định tỉ lệ thời gian để cung cấp phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào. Bộ định tỉ lệ thời gian được tạo cấu hình để tính toán hoặc ước lượng chất lượng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bởi sự định tỉ lệ thời gian của tín hiệu âm thanh đầu vào. Hơn nữa, bộ định tỉ lệ thời gian được tạo cấu hình để thực hiện sự định tỉ lệ thời gian của tín hiệu âm thanh đầu vào phụ thuộc vào sự tính toán hoặc sự ước lượng chất lượng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bởi sự định tỉ lệ thời gian. Phương án này theo sáng chế là dựa trên ý tưởng rằng các tình huống trong sự định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể dẫn đến độ méo có thể nghe được đáng kể. Hơn nữa, phương án theo sáng chế dựa trên việc tìm ra rằng cơ cấu điều khiển chất lượng giúp tránh được độ méo có thể nghe được như vậy bằng cách ước lượng liệu sự định tỉ lệ thời gian mong muốn có thực sự cung cấp chất lượng đủ của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào không. Theo đó, sự định tỉ lệ thời gian không chỉ được điều khiển bởi sự kéo dài thời gian hoặc sự rút ngắn thời gian mong muốn, nhưng cũng là sự đánh giá chất lượng có thể thu được. Theo đó, có thể, ví dụ, hoãn sự định tỉ lệ thời gian nếu sự định tỉ lệ thời gian dẫn đến chất lượng thấp không thể chấp nhận được của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào. Tuy nhiên, sự ước lượng tính

toán của chất lượng (mong muốn) của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể cũng được sử dụng để điều chỉnh bất kỳ tham số nào khác của sự định tỉ lệ thời gian. Kết luận, cơ cấu điều khiển chất lượng được sử dụng trong phương án được đề cập ở trên giúp giảm hoặc tránh các thành phần lạ có thể nghe được trong hệ thống mà sự định tỉ lệ thời gian được áp dụng.

Trong phương án được ưu tiên, bộ định tỉ lệ thời gian được tạo cấu hình để thực hiện thuật toán chồng lấp và cộng sử dụng khối các mẫu thứ nhất của tín hiệu âm thanh đầu vào và khối các mẫu thứ hai của tín hiệu âm thanh đầu vào (trong đó khối các mẫu thứ nhất của tín hiệu âm thanh đầu vào và khối các mẫu thứ hai của tín hiệu âm thanh đầu vào có thể là các khối các mẫu chồng lấp hoặc không chồng lấp, mà thuộc về khung đơn hoặc thuộc về các khung khác). Bộ định tỉ lệ thời gian được tạo cấu hình để dịch chuyển thời gian khối các mẫu thứ hai so với khối các mẫu thứ nhất (ví dụ, khi được so sánh với đường thời gian ban đầu được kết hợp với khối các mẫu thứ nhất và khối các mẫu thứ hai), và để chồng lấp và cộng khối các mẫu thứ nhất và khối các mẫu thứ hai được dịch chuyển thời gian, nhờ đó thu được phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào. Phương án này theo sáng chế được dựa trên việc tìm ra rằng thuật toán chồng lấp và cộng sử dụng khối các mẫu thứ nhất và khối các mẫu thứ hai thường dẫn đến sự định tỉ lệ thời gian tốt, trong đó sự điều chỉnh của sự dịch chuyển thời gian của khối các mẫu thứ hai so với khối các mẫu thứ nhất cho phép giữ độ méo nhỏ có thể chấp nhận được trong nhiều trường hợp. Tuy nhiên, các tác giả sáng chế đã tìm ra rằng việc đưa cơ cấu điều khiển chất lượng bổ sung, mà kiểm tra liệu sự chồng lấp và cộng được hình dung của khối các mẫu thứ nhất và khối các mẫu thứ hai được dịch chuyển thời gian thực sự dẫn đến chất lượng thích đáng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào, giúp tránh được các thành phần lạ có thể nghe được với sự tin cậy thậm chí tốt hơn. Nói cách khác, các tác giả sáng chế đã tìm ra rằng sẽ có lợi khi thực hiện sự kiểm tra chất lượng (dựa trên sự ước lượng chất lượng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bằng sự định tỉ lệ thời gian) sau sự dịch chuyển thời gian mong muốn (hoặc có lợi) của khối các mẫu thứ hai so với khối các mẫu thứ nhất đã được nhận ra, vì quy trình này giúp giảm hoặc tránh các thành phần lạ có thể nghe được.

Trong phương án được ưu tiên, bộ định tỉ lệ thời gian được tạo cấu hình để tính toán hoặc ước lượng chất lượng (ví dụ, chất lượng mong muốn) của thuật toán chồng lấp và cộng giữa khối các mẫu thứ nhất và khối các mẫu thứ hai được dịch chuyển thời gian, để tính toán hoặc ước lượng chất lượng (mong muốn) của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bởi sự định tỉ lệ thời gian. Các tác giả sáng chế đã tìm ra rằng chất lượng của thuật toán chồng lấp và cộng thực sự có ảnh hưởng mạnh đến chất lượng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh có thể thu được bởi sự định tỉ lệ thời gian.

Trong phương án được ưu tiên, bộ định tỉ lệ thời gian được tạo cấu hình để xác định sự dịch chuyển thời gian của khối các mẫu thứ hai so với khối các mẫu thứ nhất phụ thuộc vào sự xác định mức tương tự giữa khối các mẫu thứ nhất, hoặc phần của khối các mẫu thứ nhất (ví dụ, phần phía bên phải, tức là, các mẫu tại cuối khối các mẫu thứ nhất), và khối các mẫu thứ hai, hoặc phần của khối các mẫu thứ hai (ví dụ, phần phía bên trái, tức là, các mẫu tại đầu khối các mẫu thứ hai). Khái niệm này dựa trên việc tìm ra rằng sự xác định sự tương tự giữa khối các mẫu thứ nhất và khối các mẫu thứ hai được dịch chuyển thời gian cung cấp cho sự ước lượng chất lượng của thuật toán chồng lấp và cộng, và tiếp theo cũng cung cấp cho sự ước lượng có ý nghĩa chất lượng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bằng cách định tỉ lệ thời gian. Hơn nữa, các tác giả sáng chế đã tìm ra rằng mức độ tương tự giữa khối các mẫu thứ nhất (hoặc phần phía bên phải của khối các mẫu thứ nhất) và khối các mẫu thứ hai được dịch chuyển thời gian (hoặc phần phía bên trái của khối các mẫu thứ hai được dịch chuyển thời gian) có thể được xác định với sự chính xác tốt sử dụng sự phức tạp tính toán thích hợp.

Trong phương án được ưu tiên, bộ định tỉ lệ thời gian được tạo cấu hình để xác định thông tin về mức tương tự giữa khối các mẫu thứ nhất, hoặc phần (ví dụ, phần phía bên phải) của khối các mẫu thứ nhất, và khối các mẫu thứ hai, hoặc phần (ví dụ, phần phía bên trái) của khối các mẫu thứ hai, đối với nhiều sự dịch chuyển thời gian khác nhau giữa khối các mẫu thứ nhất và khối các mẫu thứ hai, và để xác định sự dịch chuyển thời gian (ứng cử), để được sử dụng cho thuật toán chồng lấp và cộng, trên cơ sở thông tin về mức tương tự đối với nhiều sự dịch chuyển thời gian khác nhau. Theo đó, sự dịch chuyển thời gian của khối các mẫu thứ hai hoặc đối với khối các mẫu thứ

nhất có thể được chọn để thích ứng với nội dung âm thanh. Tuy nhiên, việc điều khiển chất lượng, mà bao gồm sự tính toán hoặc sự ước lượng của chất lượng (mong muốn) của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bằng sự định tỉ lệ thời gian của tín hiệu âm thanh đầu vào, có thể được thực hiện tiếp theo sự xác định sự dịch chuyển thời gian (ứng cử) để được sử dụng cho thuật toán chồng lấp và cộng. Nói cách khác, bằng cách sử dụng cơ cấu điều khiển chất lượng âm thanh, có thể đảm bảo rằng sự dịch chuyển thời gian được xác định trên cơ sở thông tin về mức tương tự giữa khối các mẫu thứ nhất (hoặc phần của khối các mẫu thứ nhất) và khối các mẫu thứ hai (hoặc phần của khối các mẫu thứ hai) cho nhiều sự dịch chuyển thời gian khác nhau thực sự dẫn đến chất lượng âm thanh đủ tốt. Do đó, các thành phần lạ có thể được giảm hoặc tránh một cách hiệu quả.

Trong phương án được ưu tiên, bộ định tỉ lệ thời gian được tạo cấu hình để xác định sự dịch chuyển thời gian của khối các mẫu thứ hai so với khối các mẫu thứ nhất, mà sự dịch chuyển thời gian được sử dụng cho thuật toán chồng lấp và cộng (trừ khi thuật toán chồng lấp và cộng bị hoãn phản hồi sự ước lượng chất lượng không đủ), phụ thuộc vào thông tin dịch chuyển thời gian đích. Nói cách khác, thông tin dịch chuyển thời gian đích được xem xét, và nỗ lực được thực hiện để xác định sự dịch chuyển thời gian của khối các mẫu thứ hai so với khối các mẫu thứ nhất sao cho sự dịch chuyển thời gian đã nêu của khối các mẫu thứ hai so với khối các mẫu thứ nhất gần với sự dịch chuyển thời gian đích được mô tả bởi thông tin dịch chuyển thời gian đích. Tiếp theo, có thể đạt được rằng sự dịch chuyển thời gian (ứng cử), mà thu được bằng sự chồng lấp và cộng của khối các mẫu thứ nhất và khối các mẫu thứ hai được dịch chuyển thời gian, là thỏa thuận với yêu cầu (được định nghĩa bởi thông tin dịch chuyển thời gian đích), trong đó sự thực hiện thực sự thuật toán chồng lấp và cộng có thể được ngăn chặn nếu sự tính toán hoặc sự ước lượng chất lượng (mong muốn) của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bằng sự định tỉ lệ thời gian biểu thị chất lượng không đủ.

Trong phương án được ưu tiên, bộ định tỉ lệ thời gian được tạo cấu hình để tính toán hoặc ước lượng chất lượng (ví dụ, chất lượng mong muốn) của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bởi sự định tỉ lệ thời gian của tín hiệu âm thanh đầu vào trên cơ sở thông tin về mức tương tự giữa khối các

mẫu thứ nhất, hoặc phần (ví dụ, phần phía bên phải) của khối các mẫu thứ nhất, và khối các mẫu thứ hai, được dịch chuyển thời gian bởi sự dịch chuyển thời gian được xác định, hoặc phần (ví dụ, phần phía bên trái) của khối mẫu thứ hai, được dịch chuyển thời gian bởi sự dịch chuyển thời gian được xác định. Các tác giả sáng chế đã tìm ra rằng mức tương tự giữa khối các mẫu thứ nhất, hoặc phần của khối các mẫu thứ nhất, và khối các mẫu thứ hai, được dịch chuyển thời gian bởi sự dịch chuyển thời gian được xác định, hoặc phần của khối các mẫu thứ hai, được dịch chuyển thời gian bởi sự dịch chuyển thời gian được xác định, tạo thành tiêu chuẩn tốt để quyết định liệu phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bằng sự định tỉ lệ thời gian sẽ có chất lượng đủ hay không.

Trong phương án được ưu tiên, bộ định tỉ lệ thời gian được tạo cấu hình để quyết định, trên cơ sở thông tin về mức tương tự giữa khối các mẫu thứ nhất, hoặc phần (ví dụ, phần phía bên phải) của khối các mẫu thứ nhất, và khối các mẫu thứ hai, được dịch chuyển thời gian bởi sự dịch chuyển thời gian được xác định, hoặc phần (ví dụ, phần phía bên trái) của khối mẫu thứ hai, được dịch chuyển thời gian bởi sự dịch chuyển thời gian được xác định, liệu sự định tỉ lệ thời gian có thực sự được thực hiện hay không. Theo đó, sự xác định sự dịch chuyển thời gian, mà được nhận ra như sự dịch chuyển thời gian ứng cử sử dụng thuật toán (thường đơn giản hơn về mặt tính toán và không tin cậy cao) thứ nhất được theo sau bởi sự kiểm tra chất lượng, mà dựa trên thông tin về mức tương tự giữa khối các mẫu thứ nhất (hoặc phần của khối các mẫu thứ nhất) và khối các mẫu thứ hai, được dịch chuyển thời gian bởi sự dịch chuyển thời gian được xác định (hoặc phần của khối các mẫu thứ hai, được dịch chuyển thời gian bằng sự dịch chuyển thời gian được xác định). "Sự kiểm tra chất lượng" trên cơ sở thông tin đã nêu thường tin cậy hơn sự xác định đơn thuần của sự dịch chuyển thời gian ứng cử, và do đó được sử dụng để cuối cùng quyết định liệu sự định tỉ lệ thời gian có thực sự được thực hiện. Do đó, sự định tỉ lệ thời gian có thể được tránh nếu sự định tỉ lệ thời gian dẫn đến các thành phần lạ có thể nghe được (hoặc độ méo) quá mức.

Trong phương án được ưu tiên, bộ định tỉ lệ thời gian được tạo cấu hình để dịch chuyển thời gian khối các mẫu thứ hai so với khối các mẫu thứ nhất, và để chồng lấp và cộng khối các mẫu thứ nhất và khối các mẫu thứ hai được dịch chuyển thời gian, để nhờ đó thu được phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào, nếu

sự tính toán hoặc ước lượng chất lượng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bằng sự định tỉ lệ thời gian biểu thị chất lượng mà lớn hơn hoặc bằng giá trị ngưỡng chất lượng. Bộ định tỉ lệ thời gian được tạo cấu hình để xác định sự dịch chuyển thời gian của khối các mẫu thứ hai so với khối các mẫu thứ nhất phụ thuộc vào sự xác định mức tương tự, được đánh giá sử dụng số đo độ tương tự thứ nhất, giữa khối các mẫu thứ nhất, hoặc phần (ví dụ, phần phía bên phải) của khối các mẫu thứ nhất, và khối các mẫu thứ hai, hoặc phần (ví dụ, phần phía bên trái) của khối các mẫu thứ hai. Bộ định tỉ lệ thời gian còn được tạo cấu hình để tính toán hoặc ước lượng chất lượng (ví dụ, chất lượng mong muốn) của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bởi sự định tỉ lệ thời gian của tín hiệu âm thanh đầu vào trên cơ sở thông tin về mức tương tự, được đánh giá sử dụng số đo độ tương tự thứ hai, giữa khối các mẫu thứ nhất, hoặc phần (ví dụ, phần phía bên phải) của khối các mẫu thứ nhất, và khối các mẫu thứ hai, được dịch chuyển thời gian bởi sự dịch chuyển thời gian được xác định, hoặc phần (ví dụ, phần phía bên trái) của khối mẫu thứ hai, được dịch chuyển thời gian bởi sự dịch chuyển thời gian được xác định. Sự sử dụng số đo độ tương tự thứ nhất và số đo độ tương tự thứ hai cho phép nhanh chóng xác định sự dịch chuyển thời gian của khối các mẫu thứ hai so với khối các mẫu thứ nhất với độ phức tạp tính toán vừa phải, và điều này cũng cho phép tính toán hoặc ước lượng chất lượng phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được nhờ sự định tỉ lệ thời gian của tín hiệu âm thanh đầu vào với độ chính xác cao. Do đó, quy trình hai bước, sử dụng hai số đo độ tương tự khác nhau, cho phép sự tổ hợp độ phức tạp tính toán tương đối nhỏ trong bước thứ nhất với độ chính xác cao trong bước (điều khiển chất lượng) thứ hai và cho phép giảm hoặc tránh các thành phần lạ có thể nghe được mặc dù số đo độ tương tự thứ nhất, mà thường đơn giản về tính toán, được sử dụng cho sự xác định sự dịch chuyển thời gian (ứng cử) của khối các mẫu thứ hai so với khối các mẫu thứ nhất (trong đó sẽ là thường quá đòi hỏi để sử dụng số đo độ tương tự có độ phức tạp tính toán cao, như số đo độ tương tự thứ hai, khi xác định sự dịch chuyển thời gian ứng cử của khối các mẫu thứ hai so với khối các mẫu thứ nhất).

Trong phương án được ưu tiên, số đo độ tương tự thứ hai phức tạp về tính toán hơn số đo độ tương tự thứ nhất. Theo đó, sự kiểm tra chất lượng "cuối cùng" có thể

được thực hiện với độ chính xác cao, trong khi sự xác định dễ dàng của sự dịch chuyển thời gian của khối các mẫu thứ hai so với khối các mẫu thứ nhất có thể được thực hiện theo cách hiệu quả.

Trong phương án được ưu tiên, số đo độ tương tự thứ nhất là sự tương quan chéo hoặc sự tương quan chéo được tiêu chuẩn hóa hoặc hàm chênh lệch độ lớn trung bình hoặc tổng các sai số được bình phương. Hơn nữa, số đo độ tương tự thứ hai là tổ hợp các sự tương quan chéo hoặc của các sự tương quan được tiêu chuẩn hóa cho nhiều sự dịch chuyển thời gian khác nhau. Các tác giả sáng chế đã tìm ra rằng, sự tương quan chéo, sự tương quan chéo được tiêu chuẩn hóa, hàm chênh lệch độ cao trung bình hoặc tổng các sai số được bình phương cho phép sự xác định tốt và hiệu quả của sự dịch chuyển thời gian (ứng cử) của khối các mẫu thứ hai so với khối các mẫu thứ nhất. Hơn nữa, các tác giả sáng chế đã tìm ra rằng số đo độ tương tự mà là tổ hợp của sự tương quan chéo hoặc của nhiều sự dịch chuyển thời gian khác nhau là chất lượng có thể tin cậy cao để đánh giá (tính toán hoặc ước lượng) chất lượng của phiên bản được định tỉ lệ theo thời gian của tín hiệu âm thanh đầu vào có thể thu được bằng sự định tỉ lệ thời gian.

Trong phương án được ưu tiên, số đo độ tương tự thứ hai là tổ hợp của các sự tương quan chéo cho ít nhất bốn sự dịch chuyển thời gian khác nhau. Các tác giả sáng chế đã tìm ra rằng sự tổ hợp của các sự tương quan chéo cho ít nhất bốn sự dịch chuyển khác nhau cho phép sự đánh giá chất lượng chính xác vì các thay đổi của tín hiệu theo thời gian có thể cũng được xem xét bằng cách xác định sự tương quan cho ít nhất bốn sự dịch chuyển thời gian khác nhau. Tương tự, sóng hài có thể được xem đến mức độ nào đó bằng cách sử dụng sự tương quan chéo cho ít nhất bốn sự dịch chuyển thời gian khác nhau. Tiếp theo, sự đánh giá tốt đặc biệt của chất lượng âm thanh có thể thu được có thể đạt được.

Trong phương án được ưu tiên, số đo độ tương tự thứ hai là tổ hợp của giá trị tương quan chéo thứ nhất và giá trị tương quan chéo thứ hai, mà thu được cho sự dịch chuyển thời gian mà được tạo khoảng cách bằng bội số nguyên của thời khoảng chu kỳ của tần số cơ bản của nội dung âm thanh của khối các mẫu thứ nhất hoặc khối các mẫu thứ hai, và của giá trị tương quan chéo thứ ba và giá trị tương quan chéo thứ tư, mà thu

được cho các sự dịch chuyển thời gian mà được chèn khoảng trống bởi bội số nguyên của thời khoảng chu kỳ của tần số cơ bản của nội dung âm thanh, trong đó sự dịch chuyển thời gian mà giá trị tương quan chéo thứ nhất thu được được chèn khoảng cách từ sự dịch chuyển thời gian mà giá trị tương quan chéo thứ ba thu được bằng bội số lẻ của một nửa thời khoảng chu kỳ của tần số cơ bản của nội dung âm thanh. Theo đó, giá trị tương quan chéo thứ nhất và giá trị tương quan chéo thứ hai có thể cung cấp thông tin liệu nội dung âm thanh ít nhất gần như ổn định theo thời gian. Tương tự, giá trị tương quan chéo thứ ba và giá trị tương quan chéo thứ tư cũng cung cấp thông tin liệu nội dung âm thanh ít nhất gần như ổn định theo thời gian. Hơn nữa, thực tế rằng giá trị tương quan chéo thứ ba và giá trị tương quan chéo thứ tư là "chênh lệch theo thời gian" so với giá trị tương quan chéo thứ nhất và giá trị tương quan chéo thứ hai cho phép sự xem xét các sóng hài. Kết luận, sự tính toán số đo độ tương tự thứ hai trên cơ sở sự tổ hợp giá trị tương quan chéo thứ nhất, giá trị tương quan chéo thứ hai, giá trị tương quan chéo thứ ba, và giá trị tương quan chéo thứ tư mang lại độ chính xác cao, và do đó kết quả có thể tin tưởng được cho sự tính toán (hoặc sự ước lượng) của chất lượng (mong đợi) của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bởi sự định tỉ lệ thời gian.

Trong phương án được ưu tiên, số đo độ tương tự thứ hai q thu được theo $q = c(p) * c(2*p) + c(3/2*p) * c(1/2*p)$ hoặc theo $q = c(p) * c(-p) + c(-1/2*p) * c(1/2*p)$. Trong các phương trình trên, $c(p)$ là giá trị tương quan chéo giữa khối các mẫu thứ nhất và khối các mẫu thứ hai, mà được dịch chuyển trong thời gian (so với nhau, và so với trục thời gian gốc) bằng thời khoảng chu kỳ p của tần số cơ bản của nội dung âm thanh của khối các mẫu thứ nhất hoặc của khối các mẫu thứ hai. $c(2*p)$ là giá trị tương quan chéo giữa khối các mẫu thứ nhất và khối các mẫu thứ hai, mà được dịch chuyển trong thời gian bởi $2*p$. $c(3/2*p)$ là giá trị tương quan chéo giữa khối các mẫu thứ nhất và khối các mẫu thứ hai, mà được dịch chuyển trong thời gian bằng $3/2*p$. $c(1/2*p)$ là giá trị tương quan chéo giữa khối các mẫu thứ nhất và khối các mẫu thứ hai, mà được dịch chuyển trong thời gian bằng $1/2*p$. $c(-p)$ là giá trị tương quan chéo giữa khối các mẫu thứ nhất và khối các mẫu thứ hai, mà được dịch chuyển trong thời gian bằng $-p$ và $c(-1/2*p)$ là giá trị tương quan giữa khối các mẫu thứ nhất và khối các mẫu thứ hai, mà được dịch chuyển trong thời gian bằng $-1/2*p$. Các tác giả

sáng chế đã tìm ra rằng việc sử dụng các phương trình trên dẫn đến sự tính toán (hoặc ước lượng) đặc biệt tốt và có thể tin tưởng chất lượng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bằng sự định tỉ lệ thời gian.

Trong phương án được ưu tiên, bộ định tỉ lệ thời gian được tạo cấu hình để so sánh giá trị chất lượng, mà dựa trên sự tính toán hoặc sự ước lượng chất lượng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bằng sự định tỉ lệ thời gian, với giá trị ngưỡng có thể biến thiên, để quyết định liệu việc định tỉ lệ thời gian nên được thực hiện hay không. Việc sử dụng giá trị ngưỡng có thể thay đổi cho phép thích ứng ngưỡng để quyết định liệu sự định tỉ lệ thời gian nên được thực hiện hay không cho tình huống. Theo đó, yêu cầu chất lượng để thực hiện sự định tỉ lệ thời gian có thể tăng trong một số tình huống, và có thể giảm trong các tình huống khác, ví dụ, phụ thuộc vào các thuật toán định tỉ lệ thời gian trước, hoặc bất kỳ đặc điểm nào khác của tín hiệu. Do đó, sự quan trọng của sự quyết định liệu thực hiện sự định tỉ lệ thời gian hay không có thể còn tăng.

Trong phương án được ưu tiên, bộ định tỉ lệ thời gian được tạo cấu hình để giảm giá trị ngưỡng có thể thay đổi, để nhờ đó tăng yêu cầu chất lượng, phản hồi việc tìm ra rằng chất lượng của sự định tỉ lệ thời gian sẽ không đủ cho một hoặc nhiều khối các mẫu trước đó. Bằng cách giảm giá trị ngưỡng biến thiên, điều này có thể tránh việc định tỉ lệ thời gian được loại bỏ trên chu kỳ thời gian được mở rộng, vì điều này có thể dẫn đến sự chạy non vùng đệm hoặc sự chạy tràn vùng đệm và do đó có thể bất lợi hơn so với sự tạo ra một số thành phần lạ được gây ra bởi định tỉ lệ thời gian. Do đó, các vấn đề mà sẽ được gây ra bởi sự trì hoãn quá mức của sự định tỉ lệ thời gian có thể được ngăn ngừa.

Trong phương án được ưu tiên, bộ định tỉ lệ thời gian được tạo cấu hình để tăng giá trị ngưỡng có thể thay đổi, để nhờ đó tăng yêu cầu chất lượng, phản hồi sự thật rằng sự định tỉ lệ đã được áp dụng cho một hoặc nhiều khối các mẫu trước. Theo đó, có thể đảm bảo rằng khối các mẫu tiếp theo chỉ được định tỉ lệ thời gian nếu mức chất lượng tương đối cao (cao hơn mức chất lượng "thông thường") có thể đạt được. Ngược lại, sự định tỉ lệ chuỗi các khối các mẫu tiếp theo được ngăn ngừa nếu sự định tỉ lệ thời gian sẽ không đáp ứng các yêu cầu chất lượng tương đối cao. Điều này phù hợp, vì sự

ứng dụng sự định tỉ lệ thời gian cho nhiều khối các mẫu tiếp theo sẽ thường dẫn đến các thành phần lạ trừ khi sự định tỉ lệ thời gian đáp ứng các yêu cầu chất lượng tương đối cao (mà thường cao hơn các yêu cầu chất lượng "thường" có thể áp dụng nếu chỉ khối các mẫu đơn hơn là chuỗi các khối các mẫu kề nhau, là để được định tỉ lệ thời gian).

Trong phương án được ưu tiên, bộ định tỉ lệ thời gian bao gồm bộ đếm thứ nhất được giới hạn phạm vi để đếm số lượng khối các mẫu hoặc số lượng khung mà đã được định tỉ lệ thời gian vì sự yêu cầu chất lượng tương ứng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bằng sự định tỉ lệ thời gian đã đạt được. Hơn nữa, bộ định tỉ lệ thời gian bao gồm bộ đếm thứ hai được giới hạn phạm vi để đếm số lượng khối các mẫu hoặc số lượng khung mà chưa được định tỉ lệ thời gian vì sự yêu cầu chất lượng tương ứng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bằng sự định tỉ lệ thời gian chưa đạt được. Bộ định tỉ lệ thời gian được tạo cấu hình để tính toán giá trị ngưỡng biến thiên phụ thuộc vào giá trị bộ đếm thứ nhất và phụ thuộc vào giá trị của bộ đếm thứ hai. Bằng các sử dụng bộ đếm thứ nhất được giới hạn phạm vi và bộ đếm thứ hai được giới hạn phạm vi, cơ cấu đơn giản cho sự điều chỉnh giá trị ngưỡng biến thiên thu được, mà cho phép thích ứng giá trị ngưỡng biến thiên vào tình huống tương ứng trong khi tránh các giá trị quá nhỏ hoặc quá lớn của giá trị ngưỡng.

Trong phương án được ưu tiên, bộ định tỉ lệ thời gian được tạo cấu hình để cộng giá trị mà tỉ lệ với giá trị của bộ đếm thứ nhất vào giá trị ngưỡng ban đầu, và để trừ giá trị mà tỉ lệ với giá trị của bộ đếm thứ hai từ đó, để thu được giá trị ngưỡng biến thiên. Bằng cách sử dụng khái niệm như vậy, giá trị ngưỡng biến thiên có thể thu được bằng cách rất đơn giản.

Trong phương án được ưu tiên, bộ định tỉ lệ thời gian được tạo cấu hình để thực hiện sự định tỉ lệ thời gian của tín hiệu âm thanh đầu vào phụ thuộc vào sự tính toán hoặc ước lượng chất lượng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bằng sự định tỉ lệ thời gian, trong đó sự tính toán hoặc sự ước lượng chất lượng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào bao gồm sự tính toán hoặc ước lượng của các thành phần lạ trong phiên bản

được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào mà sẽ được gây ra bởi sự định tỉ lệ thời gian. Bằng cách tính toán hoặc ước lượng các thành phần lạ trong phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào mà sẽ được gây ra bởi sự định tỉ lệ thời gian, tiêu chí ý nghĩa cho sự tính toán hoặc ước lượng chất lượng có thể được sử dụng, vì các thành phần lạ thường sẽ giảm ấn tượng nghe của người nghe.

Trong phương án được ưu tiên, sự ước lượng tính toán của chất lượng (mong muốn) của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào bao gồm sự tính toán hoặc sự ước lượng của các thành phần lạ trong phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào mà sẽ được gây ra bởi thuật toán chồng lấp và cộng của các khối các mẫu tiếp theo của tín hiệu âm thanh đầu vào. Các tác giả sáng chế đã nhận ra rằng thuật toán chồng lấp và cộng có thể là nguồn đầu tiên của các thành phần lạ khi thực hiện sự định tỉ lệ thời gian. Theo đó, các tác giả sáng chế đã tìm ra rằng phương pháp hiệu quả là tính toán hoặc ước lượng các thành phần lạ của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào mà sẽ được gây ra bởi thuật toán chồng lấp và cộng của các khối các mẫu tiếp theo của tín hiệu âm thanh đầu vào.

Trong phương án được ưu tiên, bộ định tỉ lệ thời gian được tạo cấu hình để tính toán hoặc ước lượng chất lượng (mong muốn) của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bởi sự định tỉ lệ thời gian của tín hiệu âm thanh đầu vào phụ thuộc vào mức tương tự của các khối các mẫu tiếp theo của tín hiệu âm thanh đầu vào. Các tác giả sáng chế đã tìm ra rằng sự định tỉ lệ thời gian có thể được thực hiện với chất lượng tốt nếu các khối các mẫu tiếp theo của tín hiệu âm thanh đầu vào bao gồm sự tương tự tương đối cao, và rằng sự làm méo thường được tạo ra bởi sự định tỉ lệ thời gian nếu các khối các mẫu tiếp theo của tín hiệu âm thanh đầu vào bao gồm các sự chênh lệch đáng kể.

Trong phương án được ưu tiên, bộ định tỉ lệ thời gian được tạo cấu hình để tính toán hoặc ước lượng liệu có các thành phần lạ có thể nghe được trong phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bởi sự định tỉ lệ thời gian của tín hiệu âm thanh đầu vào. Các tác giả sáng chế đã tìm ra rằng sự tính toán

hoặc ước lượng của các thành phần lạ có thể nghe được cung cấp thông tin chất lượng mà được thích ứng tốt với ấn tượng nghe của con người.

Trong phương án được ưu tiên, bộ định tỉ lệ thời gian được tạo cấu hình để hoãn sự định tỉ lệ thời gian đến khung tiếp theo hoặc khối các mẫu tiếp theo nếu sự tính toán hoặc ước lượng chất lượng (mong muốn) của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bởi sự định tỉ lệ thời gian biểu thị chất lượng không đủ. Theo đó, có thể thực hiện sự định tỉ lệ thời gian tại thời gian mà phù hợp hơn cho sự định tỉ lệ thời gian trong đó ít các thành phần lạ được tạo ra. Nói cách khác, bằng cách lựa chọn một cách linh hoạt thời gian tại đó sự định tỉ lệ thời gian được thực hiện phụ thuộc vào chất lượng có thể đạt được bằng sự định tỉ lệ thời gian, ấn tượng nghe của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể được cải thiện. Hơn nữa, ý tưởng này dựa trên sự tìm ra rằng độ trễ nhẹ của thuật toán định tỉ lệ thời gian thường không cung cấp bất kỳ đáng kể nào.

Trong phương án được ưu tiên, bộ định tỉ lệ thời gian được tạo cấu hình để hoãn sự định tỉ lệ thời gian đến thời gian khi sự định tỉ lệ thời gian có thể nghe được ít hơn nếu sự tính toán hoặc ước lượng chất lượng (mong muốn) của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bởi sự định tỉ lệ thời gian biểu thị chất lượng không đủ. Theo đó, ấn tượng nghe có thể được cải thiện bằng cách tránh độ méo có thể nghe được.

Phương án theo sáng chế tạo ra bộ giải mã âm thanh để cung cấp nội dung âm thanh được giải mã trên cơ sở nội dung âm thanh đầu vào. Bộ giải mã âm thanh bao gồm vùng đệm chấp chèn được tạo cấu hình để tạo vùng đệm nhiều khung âm thanh biểu diễn các khối của các mẫu âm thanh. Bộ giải mã âm thanh cũng bao gồm lõi bộ giải mã được tạo cấu hình để cung cấp các khối mẫu âm thanh trên cơ sở các khung âm thanh nhận được từ vùng đệm chấp chèn. Hơn nữa, bộ giải mã âm thanh bao gồm bộ định tỉ lệ thời gian trên cơ sở mẫu như được liệt kê ở trên. Bộ định tỉ lệ thời gian trên cơ sở mẫu được tạo cấu hình để cung cấp các khối các mẫu âm thanh được định tỉ lệ thời gian trên cơ sở các khối các mẫu âm thanh được cung cấp bởi lõi bộ giải mã. Bộ giải mã âm thanh được dựa trên ý tưởng rằng bộ định tỉ lệ thời gian, mà được tạo cấu hình để thực hiện sự định tỉ lệ thời gian của tín hiệu âm thanh đầu vào phụ thuộc

vào sự tính toán hoặc ước lượng của chất lượng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bằng sự định tỉ lệ thời gian được thích ứng tốt cho sự sử dụng trong bộ giải mã âm thanh bao gồm vùng đệm chấp chừa và lỗi bộ giải mã. Sự có mặt của vùng đệm chấp chừa cho phép, ví dụ, hoãn thuật toán định tỉ lệ thời gian nếu sự tính toán hoặc ước lượng của chất lượng (mong muốn) của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bằng sự định tỉ lệ thời gian biểu thị rằng chất lượng xấu có thể thu được. Do đó, bộ định tỉ lệ thời gian trên cơ sở mẫu, mà bao gồm cơ cấu điều khiển chất lượng, cho phép tránh, hoặc ít nhất giảm, các thành phần lạ có thể nghe được trong bộ giải mã âm thanh bao gồm vùng đệm chấp chừa và lỗi bộ giải mã.

Trong phương án được ưu tiên, bộ giải mã âm thanh còn bao gồm bộ điều khiển vùng đệm chấp chừa. Bộ điều khiển vùng đệm chấp chừa được tạo cấu hình để cung cấp thông tin điều khiển cho bộ định tỉ lệ thời gian trên cơ sở mẫu, trong đó thông tin điều khiển biểu thị liệu sự định tỉ lệ thời gian trên cơ sở mẫu nên được thực hiện hay không. Cách khác, hoặc ngoài ra, thông tin điều khiển có thể biểu thị lượng mong muốn của sự định tỉ lệ thời gian. Theo đó, bộ định tỉ lệ thời gian trên cơ sở mẫu có thể được điều khiển phụ thuộc vào yêu cầu của bộ giải mã âm thanh. Ví dụ, bộ điều khiển vùng đệm chấp chừa có thể thực hiện sự điều khiển thích ứng tín hiệu, và có thể lựa chọn liệu sự định tỉ lệ thời gian trên cơ sở khung hoặc sự định tỉ lệ thời gian trên cơ sở mẫu nên được thực hiện bằng cách thích ứng tín hiệu. Theo đó, có độ linh hoạt bổ sung. Tuy nhiên, cơ cấu điều khiển chất lượng của bộ định tỉ lệ thời gian trên cơ sở mẫu có thể, ví dụ, gạt bỏ thông tin điều khiển được cung cấp bởi bộ điều khiển vùng đệm chấp chừa, sao cho sự định tỉ lệ thời gian trên cơ sở mẫu được tránh (hoặc được vô hiệu hóa) thậm chí trong trường hợp mà thông tin điều khiển được cung cấp bằng bộ điều khiển vùng đệm chấp chừa biểu thị rằng sự định tỉ lệ thời gian trên cơ sở mẫu nên được thực hiện. Do đó, bộ định tỉ lệ thời gian trên cơ sở mẫu "thông minh" có thể gạt bỏ bộ điều khiển vùng đệm chấp chừa, vì bộ định tỉ lệ thời gian trên cơ sở mẫu có thể thu được nhiều thông tin chi tiết về chất lượng có thể thu được bằng sự định tỉ lệ thời gian. Kết luận, bộ định tỉ lệ thời gian trên cơ sở mẫu có thể được dẫn bởi thông tin điều khiển được cung cấp bởi bộ điều khiển vùng đệm chấp chừa, nhưng tuy nhiên có thể "từ chối" sự định tỉ lệ thời gian nếu chất lượng sẽ cơ bản bị làm lộ bằng cách tiếp

theo thông tin điều khiển được cung cấp bằng bộ điều khiển vùng đệm chấp chờn, mà giúp đảm bảo chất lượng âm thanh hài lòng.

Phương án khác theo sáng chế tạo ra phương pháp để cung cấp phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào. Phương pháp bao gồm tính toán hoặc ước lượng chất lượng (ví dụ, chất lượng mong muốn) của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bởi sự định tỉ lệ thời gian của tín hiệu âm thanh đầu vào. Phương pháp còn bao gồm thực hiện sự định tỉ lệ thời gian của tín hiệu âm thanh đầu vào phụ thuộc vào sự tính toán hoặc ước lượng của chất lượng (mong muốn) của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bằng sự định tỉ lệ thời gian. Phương pháp này dựa trên sự cân nhắc tương tự như bộ định tỉ lệ thời gian được đề cập ở trên.

Phương án khác nữa theo sáng chế tạo ra chương trình máy tính để thực hiện phương pháp đã nêu khi chương trình máy tính chạy trên máy tính. Chương trình máy tính đã nêu được dựa trên sự cân nhắc tương tự như phương pháp và cũng như vùng đệm chấp chờn được mô tả ở trên.

Mô tả vắn tắt các hình vẽ

Các phương án theo sáng chế tiếp theo sẽ được mô tả dựa trên sự viện dẫn đến các hình vẽ đính kèm, trong đó:

Fig.1 thể hiện sơ đồ khối dạng giản lược của bộ điều khiển vùng đệm chấp chờn, theo phương án của sáng chế;

Fig.2 thể hiện sơ đồ khối dạng giản lược của bộ định tỉ lệ thời gian, theo phương án của sáng chế;

Fig.3 thể hiện sơ đồ khối dạng giản lược của bộ giải mã âm thanh, theo phương án của sáng chế;

Fig.4 thể hiện sơ đồ khối dạng giản lược của bộ giải mã âm thanh theo phương án khác của sáng chế, trong đó sự khái quát suốt sự quản lý vùng đệm chấp chờn (jitter buffer management - JBM) được thể hiện;

Fig.5 thể hiện mã chương trình giả của thuật toán để điều khiển mức vùng đệm PCM;

Fig.6 thể hiện mã chương trình giả của thuật toán để tính toán giá trị độ trễ và giá trị chênh lệch từ thời gian nhận và tem thời gian RTP của gói RTP;

Fig.7 thể hiện mã chương trình giả của thuật toán để tính toán các giá trị độ trễ đích;

Fig.8 thể hiện lưu đồ của logic điều khiển sự quản lý vùng đệm chấp chờn;

Fig.9 thể hiện sự biểu diễn sơ đồ khối dạng giản lược của WSOLA được thay đổi với sự điều khiển chất lượng;

Fig.10a và Fig.10b thể hiện lưu đồ của phương pháp điều khiển bộ định tỉ lệ thời gian;

Fig.11 thể hiện mã chương trình giả của thuật toán để điều khiển chất lượng cho sự định tỉ lệ thời gian;

Fig.12 thể hiện sự biểu diễn đồ họa của độ trễ đích và độ trễ sự phát ra, mà thu được bằng phương án theo sáng chế;

Fig.13 thể hiện sự biểu diễn đồ họa của sự định tỉ lệ thời gian, mà được thực hiện trong phương án theo sáng chế;

Fig.14 thể hiện lưu đồ của phương pháp để điều khiển sự cung cấp nội dung âm thanh được giải mã trên cơ sở nội dung âm thanh đầu vào; và

Fig.15 thể hiện lưu đồ của phương pháp để cung cấp phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào, theo phương án của sáng chế.

Mô tả chi tiết sáng chế

Bộ điều khiển vùng đệm chấp chờn theo Fig.1

Fig.1 thể hiện sơ đồ khối dạng giản lược của bộ điều khiển vùng đệm chấp chờn, theo phương án của sáng chế. Bộ điều khiển vùng đệm chấp chờn 100 để điều khiển sự cung cấp nội dung âm thanh được giải mã trên cơ sở nội dung âm thanh đầu vào nhận tín hiệu âm thanh 110 hoặc thông tin về tín hiệu âm thanh (mà thông tin có thể mô tả một hoặc nhiều đặc điểm của tín hiệu âm thanh, hoặc của các khung hoặc các phần tín hiệu khác của tín hiệu âm thanh).

Hơn nữa, bộ điều khiển vùng đệm chấp chờn 100 cung cấp thông tin điều khiển (ví dụ, tín hiệu điều khiển) 112 để định tỉ lệ trên cơ sở khung. Ví dụ, thông tin điều khiển 112 có thể bao gồm tín hiệu kích hoạt (cho sự định tỉ lệ thời gian trên cơ sở khung) và/hoặc thông tin điều khiển định lượng (cho sự định tỉ lệ thời gian trên cơ sở khung).

Hơn nữa, bộ điều khiển vùng đệm chấp chờn 100 cung cấp thông tin điều khiển (ví dụ, tín hiệu điều khiển) 114 để định tỉ lệ thời gian trên cơ sở mẫu. Thông tin điều khiển 114 có thể, ví dụ, bao gồm tín hiệu kích hoạt và/hoặc thông tin điều khiển định lượng cho sự định tỉ lệ thời gian trên cơ sở mẫu.

Bộ điều khiển vùng đệm chấp chờn 110 được tạo cấu hình để lựa chọn sự định tỉ lệ thời gian trên cơ sở khung hoặc sự định tỉ lệ thời gian trên cơ sở mẫu theo cách thích ứng tín hiệu. Theo đó, bộ điều khiển vùng đệm chấp chờn có thể được tạo cấu hình để đánh giá tín hiệu âm thanh hoặc thông tin về tín hiệu âm thanh 110 và để cung cấp, trên cơ sở của nó, thông tin điều khiển 112 và/hoặc thông tin điều khiển 114. Theo đó, quyết định liệu sự định tỉ lệ thời gian trên cơ sở khung hoặc sự định tỉ lệ thời gian trên cơ sở mẫu được sử dụng có thể thích ứng với các đặc điểm của tín hiệu âm thanh, ví dụ, theo cách mà sự định tỉ lệ trên cơ sở khung đơn giản về tính toán được sử dụng nếu nó mong muốn (hoặc được ước lượng) trên cơ sở tín hiệu âm thanh và/hoặc trên cơ sở thông tin về một hoặc nhiều đặc điểm của tín hiệu âm thanh mà sự định tỉ lệ thời gian trên cơ sở khung không dẫn đến sự giảm đáng kể nội dung âm thanh. Ngược lại, bộ điều khiển vùng đệm chấp chờn thường quyết định sử dụng sự định tỉ lệ thời gian trên cơ sở mẫu nếu nó được mong đợi hoặc được ước lượng (bởi bộ điều khiển vùng đệm chấp chờn), trên cơ sở sự đánh giá các đặc điểm của tín hiệu âm thanh 110, rằng sự định tỉ lệ thời gian trên cơ sở mẫu được yêu cầu để ngăn ngừa các thành phần lạ có thể nghe được khi thực hiện sự định tỉ lệ thời gian.

Hơn nữa, nên lưu ý rằng sự điều khiển vùng đệm chấp chờn 110 tự nhiên có thể cũng nhận thông tin điều khiển bổ sung, ví dụ, thông tin điều khiển biểu thị liệu sự định tỉ lệ thời gian nên được thực hiện hay không.

Dưới đây, một số chi tiết tùy chọn của bộ điều khiển vùng đệm chấp chờn 100 sẽ được mô tả. Ví dụ, bộ điều khiển vùng đệm chấp chờn 100 có thể cung cấp thông tin

điều khiển 112, 114 sao cho các khung âm thanh mà được thả hoặc được chèn để điều khiển độ sâu của vùng đệm chập chờn khi sự định tỉ lệ thời gian trên cơ sở khung được sử dụng, và sao cho sự chồng lấp và cộng được dịch chuyển thời gian của các phần tín hiệu âm thanh được thực hiện khi sự định tỉ lệ thời gian trên cơ sở mẫu được sử dụng. Nói cách khác, bộ điều khiển vùng đệm chập chờn 100 có thể kết hợp, ví dụ, với vùng đệm chập chờn (cũng được chỉ rõ như vùng đệm khử chập chờn trong một số trường hợp) và điều khiển vùng đệm chập chờn để thực hiện sự định tỉ lệ trên cơ sở khung. Trong trường hợp này, độ sâu của vùng đệm chập chờn có thể được điều khiển bằng cách thả khung từ vùng đệm chập chờn, hoặc bằng cách chèn khung (ví dụ, các khung đơn giản bao gồm sự tạo tín hiệu mà khung là "không hoạt động" và sự tạo ra nhiễu âm nền nhân tạo nên được sử dụng) vào trong vùng đệm chập chờn. Hơn nữa, bộ điều khiển vùng đệm chập chờn 100 có thể điều khiển bộ định tỉ lệ thời gian (ví dụ, bộ định tỉ lệ thời gian trên cơ sở mẫu) để thực hiện sự chồng lấp và cộng được dịch chuyển thời gian của các phần tín hiệu âm thanh.

Bộ điều khiển vùng đệm chập chờn 100 có thể được tạo cấu hình để chuyển đổi giữa sự định tỉ lệ thời gian trên cơ sở khung, sự định tỉ lệ thời gian trên cơ sở mẫu và sự vô hiệu hóa sự định tỉ lệ thời gian theo cách thích ứng tín hiệu. Nói cách khác, bộ điều khiển vùng đệm chập chờn thường không chỉ phân biệt giữa sự định tỉ lệ thời gian trên cơ sở khung và sự định tỉ lệ thời gian trên cơ sở mẫu, nhưng cũng lựa chọn trạng thái mà trong đó không có sự định tỉ lệ thời gian nào. Ví dụ, trạng thái sau có thể được chọn nếu không cần thiết để định tỉ lệ thời gian vì độ sâu của vùng đệm chập chờn nằm trong phạm vi có thể chấp nhận được. Nói cách khác, sự định tỉ lệ thời gian trên cơ sở khung và sự định tỉ lệ thời gian trên cơ sở mẫu thường không chỉ là hai chế độ của thuật toán mà có thể được lựa chọn bởi bộ điều khiển vùng đệm chập chờn.

Bộ điều khiển vùng đệm chập chờn 100 có thể cũng xem xét thông tin về độ sâu của vùng đệm chập chờn để quyết định chế độ nào của thuật toán (ví dụ, sự định tỉ lệ thời gian trên cơ sở khung, sự định tỉ lệ thời gian trên cơ sở mẫu hoặc không có sự định tỉ lệ thời gian) nên được sử dụng. Ví dụ, bộ điều khiển vùng đệm chập chờn có thể so sánh giá trị đích mô tả độ sâu mong muốn của vùng đệm chập chờn (cũng được chỉ rõ là vùng đệm khử chập chờn) và giá trị thực sự mô tả độ sâu thực sự của vùng đệm chập chờn và lựa chọn chế độ của thuật toán (sự định tỉ lệ thời gian trên cơ sở

khung, sự định tỉ lệ thời gian trên cơ sở mẫu, hoặc không có sự định tỉ lệ thời gian) phụ thuộc vào sự so sánh đã nêu, sao cho sự định tỉ lệ thời gian trên cơ sở khung hoặc sự định tỉ lệ thời gian trên cơ sở mẫu được chọn để điều khiển độ sâu của vùng đệm chập chờn.

Bộ điều khiển vùng đệm chập chờn 100 có thể, ví dụ, được tạo cấu hình để lựa chọn sự chèn nhiều âm nền nhân tạo hoặc sự xóa nhiều âm nền nhân tạo nếu khung trước đó là không hoạt động (mà có thể, ví dụ, được tự nhận ra trên cơ sở tín hiệu âm thanh 110, hoặc trên cơ sở thông tin về tín hiệu âm thanh, tương tự, ví dụ, cờ bộ nhận dạng sự im lặng SID trong trường hợp của chế độ truyền dẫn gián đoạn). Theo đó, bộ điều khiển vùng đệm chập chờn 100 có thể tạo tín hiệu đến vùng đệm chập chờn (cũng được chỉ ra như vùng đệm khử chập chờn) mà khung nhiều âm nền nhân tạo nên được chèn, nếu sự kéo dài thời gian là mong muốn và khung trước đó (hoặc khung hiện thời) không hoạt động. Hơn nữa, bộ điều khiển vùng đệm chập chờn 100 có thể hướng dẫn vùng đệm chập chờn (hoặc vùng đệm khử chập chờn) để loại bỏ khung nhiều âm nền nhân tạo (ví dụ, khung bao gồm thông tin tạo tín hiệu biểu thị rằng sự tạo ra nhiều âm nền nhân tạo nên được thực hiện) nếu nó mong muốn được thực hiện sự rút ngắn thời gian và khung trước đó không hoạt động (hoặc khung hiện thời không hoạt động). Nên lưu ý rằng khung tương ứng có thể được coi là không hoạt động khi khung tương ứng mang thông tin tạo tín hiệu biểu thị sự tạo ra nhiều âm nền nhân tạo (và thường không bao gồm nội dung âm thanh được mã hóa bổ sung nào). Thông tin tạo tín hiệu như vậy có thể, ví dụ, lấy dạng của cờ biểu thị sự im lặng (silence indication flag - SID) trong trường hợp chế độ truyền dẫn gián đoạn.

Ngược lại, bộ điều khiển vùng đệm chập chờn 100 được ưu tiên cấu hình để lựa chọn tại sự chồng lấp và cộng được dịch chuyển thời gian của các phần tín hiệu âm thanh nếu khung trước đó hoạt động (ví dụ, nếu khung trước đó không bao gồm thông tin tạo tín hiệu biểu thị rằng nhiều âm nền nhân tạo nên được tạo ra). Sự chồng lấp và cộng được dịch chuyển thời gian của các phần tín hiệu âm thanh thường cho phép sự điều chỉnh sự dịch chuyển thời gian giữa các khối các mẫu âm thanh thu được trên cơ sở các khung tiếp theo của thông tin âm thanh đầu vào với độ phân giải tương đối cao (ví dụ, với độ phân giải mà nhỏ hơn chiều dài của các khối các mẫu âm thanh, hoặc nhỏ hơn một phần tư chiều dài của các khối các mẫu âm thanh, hoặc thậm chí nhỏ hơn

hoặc bằng hai mẫu âm thanh, hoặc nhỏ như một mẫu âm thanh). Theo đó, sự lựa chọn sự định tỉ lệ thời gian trên cơ sở mẫu cho phép sự định tỉ lệ thời gian được tinh chỉnh, mà giúp tránh được các thành phần lạ có thể nghe được cho các khung hoạt động.

Trong trường hợp bộ điều khiển vùng đệm chấp chờn lựa chọn sự định tỉ lệ thời gian trên cơ sở mẫu, sự điều khiển vùng đệm chấp chờn có thể cũng cung cấp thông tin điều khiển bổ sung để điều chỉnh, hoặc tinh chỉnh, sự định tỉ lệ thời gian trên cơ sở mẫu. Ví dụ, bộ điều khiển vùng đệm chấp chờn 100 có thể được tạo cấu hình để xác định liệu khối các mẫu âm thanh biểu diễn phần tín hiệu âm thanh hoạt động nhưng "im lặng", ví dụ, phần tín hiệu âm thanh mà bao gồm năng lượng tương đối nhỏ. Trong trường hợp này, tức là, nếu phần tín hiệu âm thanh "hoạt động" (ví dụ, không phải phần tín hiệu âm thanh mà với nó sự tạo ra nhiều âm nền nhân tạo được sử dụng trong bộ giải mã âm thanh, hơn là sự giải mã chi tiết nội dung âm thanh) nhưng "im lặng" (ví dụ, trong đó, năng lượng tín hiệu dưới giá trị ngưỡng năng lượng nhất định, hoặc thậm chí bằng 0), bộ điều khiển vùng đệm chấp chờn có thể cung cấp thông tin điều khiển 114 để lựa chọn chế độ chồng lấp và cộng, trong đó sự dịch chuyển thời gian giữa khối các mẫu âm thanh biểu diễn phần tín hiệu âm thanh "im lặng" (nhưng hoạt động) và khối các mẫu âm thanh tiếp theo được thiết lập đến giá trị tối đa được định trước. Theo đó, bộ định tỉ lệ thời gian trên cơ sở mẫu không cần phải xác định lượng định tỉ lệ thời gian nhất định trên cơ sở sự so sánh chi tiết của các khối các mẫu âm thanh tiếp theo, nhưng có thể còn sử dụng một cách đơn giản giá trị tối đa được định trước cho sự dịch chuyển thời gian. Có thể hiểu rằng phần tín hiệu âm thanh "im lặng" sẽ thường không gây ra các thành phần lạ đáng kể trong thuật toán chồng lấp và cộng, không kể lựa chọn thực sự là sự dịch chuyển thời gian. Do đó, thông tin điều khiển 114 được cung cấp bởi bộ điều khiển vùng đệm chấp chờn có thể đơn giản hóa sự xử lý để được thực hiện bằng bộ định tỉ lệ thời gian trên cơ sở mẫu.

Ngược lại, bộ điều khiển vùng đệm chấp chờn 110 tìm ra rằng khối các mẫu âm thanh biểu diễn phần tín hiệu âm thanh "hoạt động" và không im lặng (ví dụ, phần tín hiệu âm thanh mà không có sự tạo ra nhiều âm nền nhân tạo cho phần tín hiệu âm thanh này, và cũng bao gồm năng lượng tín hiệu mà trên giá trị ngưỡng nhất định), bộ điều khiển vùng đệm chấp chờn cung cấp thông tin điều khiển 114 để nhờ đó lựa chọn chế độ chồng lấp và cộng trong đó sự dịch chuyển thời gian giữa các khối các mẫu âm

thanh được xác định theo cách thích ứng tín hiệu (ví dụ, bằng bộ định tỉ lệ thời gian trên cơ sở mẫu và sử dụng sự xác định sự tương tự giữa các khối các mẫu âm thanh tiếp theo).

Hơn nữa, bộ điều khiển vùng đệm chấp chờn 100 có thể cũng nhận thông tin trên sự làm đầy vùng đệm thực tế. Bộ điều khiển vùng đệm chấp chờn 100 có thể lựa chọn sự chèn của khung được che giấu (tức là, khung mà được tạo ra sử dụng cơ cấu phục hồi mất gói, ví dụ, sử dụng sự dự báo trên cơ sở các khung được giải mã trước đó) phản hồi sự xác định rằng sự kéo dài thời gian được yêu cầu và vùng đệm chấp chờn là trống. Nói cách khác, sự điều khiển vùng đệm chấp chờn có thể bắt đầu sự vận hành ngoại lệ cho trường hợp mà trong đó, về cơ bản, sự định tỉ lệ thời gian trên cơ sở mẫu sẽ được mong muốn (vì khung trước, hoặc khung hiện thời, là "hoạt động"), nhưng trong đó sự định tỉ lệ thời gian trên cơ sở mẫu (ví dụ, sử dụng sự chồng lấp và cộng) có thể không được thực hiện một cách hợp lý vì vùng đệm chấp chờn (hoặc vùng đệm khử chấp chờn) là trống. Do đó, bộ điều khiển vùng đệm chấp chờn 100 có thể được tạo cấu hình để cung cấp thông tin điều khiển thích hợp 112, 114 thậm chí cho các trường hợp ngoại lệ.

Để đơn giản hóa thuật toán của bộ điều khiển vùng đệm chấp chờn 100, bộ điều khiển vùng đệm chấp chờn 100 có thể được tạo cấu hình để lựa chọn sự định tỉ lệ thời gian trên cơ sở khung hoặc sự định tỉ lệ thời gian trên cơ sở mẫu phụ thuộc vào hoặc sự truyền dẫn gián đoạn (cũng được chỉ rõ là "DTX") cùng với sự tạo ra nhiễu âm nền nhân tạo (cũng được chỉ rõ là "CNG") hiện được sử dụng. Nói cách khác, bộ điều khiển vùng đệm chấp chờn 100 có thể, ví dụ, lựa chọn sự định tỉ lệ thời gian trên cơ sở khung nếu điều này được nhận ra, trên cơ sở tín hiệu âm thanh hoặc trên cơ sở thông tin về tín hiệu âm thanh, rằng khung trước (hoặc khung hiện thời) là khung "không hoạt động", ví dụ, sự tạo ra nhiễu âm nền nhân tạo nên được sử dụng). Điều này có thể được xác định, ví dụ, bằng cách đánh giá thông tin tạo tín hiệu (ví dụ, cờ hiệu, như được gọi là cờ hiệu "SID"), mà được bao gồm trong sự biểu diễn được mã hóa của tín hiệu âm thanh. Theo đó, sự điều khiển vùng đệm chấp chờn có thể quyết định rằng sự định tỉ lệ thời gian trên cơ sở khung nên được sử dụng nếu sự truyền dẫn gián đoạn cùng với sự tạo ra nhiễu âm nền nhân tạo được sử dụng, vì có thể mong đợi rằng chỉ các độ méo có thể nghe được nhỏ, hoặc không có độ méo có thể nghe được, được gây

ra bởi sự định tỉ lệ thời gian trong trường hợp này. Ngược lại, mặt khác, sự định tỉ lệ thời gian trên cơ sở mẫu có thể được sử dụng (ví dụ, nếu sự truyền dẫn gián đoạn cùng với sự tạo ra nhiều âm nền nhân tạo hiện không được sử dụng), trừ khi có các trường hợp ngoại lệ (như, ví dụ, vùng đệm chấp chờn trống).

Cụ thể hơn, bộ điều khiển vùng đệm chấp chờn có thể lựa chọn giữa một trong số (ít nhất) bốn chế độ trong trường hợp sự định tỉ lệ thời gian được yêu cầu. Ví dụ, bộ điều khiển vùng đệm chấp chờn có thể được tạo cấu hình để lựa chọn sự chen nhiều âm nền nhân tạo hoặc sự xóa nhiều âm nền nhân tạo cho sự định tỉ lệ thời gian nếu sự truyền dẫn gián đoạn cùng với sự tạo ra nhiều âm nền nhân tạo hiện được sử dụng. Ngoài ra, bộ điều khiển vùng đệm chấp chờn có thể được tạo cấu hình để lựa chọn thuật toán chồng lấp và cộng sử dụng sự dịch chuyển thời gian được định trước cho việc định tỉ lệ thời gian nếu phần tín hiệu âm thanh hiện thời hoạt động nhưng bao gồm năng lượng tín hiệu mà nhỏ hơn hoặc bằng giá trị ngưỡng năng lượng, và nếu vùng đệm chấp chờn không trống. Hơn nữa, bộ điều khiển vùng đệm chấp chờn có thể được tạo cấu hình để lựa chọn thuật toán chồng lấp - cộng sử dụng sự dịch chuyển thời gian thích ứng tín hiệu cho việc định tỉ lệ thời gian nếu phần tín hiệu âm thanh hiện thời hoạt động và bao gồm năng lượng tín hiệu mà lớn hơn hoặc bằng giá trị ngưỡng năng lượng, và nếu vùng đệm chấp chờn không trống. Cuối cùng, bộ điều khiển vùng đệm chấp chờn có thể được tạo cấu hình để lựa chọn sự chen khung được che giấu cho sự định tỉ lệ thời gian nếu phần tín hiệu âm thanh hiện thời hoạt động và nếu vùng đệm chấp chờn trống. Theo đó, có thể nhận thấy rằng bộ điều khiển vùng đệm chấp chờn 110 có thể được tạo cấu hình để lựa chọn sự định tỉ lệ thời gian trên cơ sở khung hoặc sự định tỉ lệ thời gian trên cơ sở mẫu theo cách thích ứng tín hiệu.

Hơn nữa, nên lưu ý rằng bộ điều khiển vùng đệm chấp chờn có thể được tạo cấu hình để lựa chọn thuật toán chồng lấp và cộng sử dụng sự dịch chuyển thời gian thích ứng tín hiệu và cơ cấu điều khiển chất lượng cho việc định tỉ lệ thời gian nếu phần tín hiệu âm thanh hiện thời hoạt động và bao gồm năng lượng tín hiệu mà lớn hơn hoặc bằng giá trị ngưỡng năng lượng và nếu vùng đệm chấp chờn không trống. Nói cách khác, có thể có cơ cấu điều khiển chất lượng bổ sung cho sự định tỉ lệ thời gian trên cơ sở mẫu, mà bổ sung sự lựa chọn thích ứng tín hiệu giữa sự định tỉ lệ thời gian trên cơ sở khung và sự định tỉ lệ thời gian trên cơ sở mẫu, mà được thực hiện bởi bộ điều

khiến vùng đệm chập chờn. Do đó, khái niệm phân cấp có thể được sử dụng, trong đó vùng đệm chập chờn thực hiện sự lựa chọn ban đầu giữa sự định tỉ lệ thời gian trên cơ sở khung và sự định tỉ lệ thời gian trên cơ sở mẫu, và trong đó cơ cấu điều khiển chất lượng bổ sung được thực hiện để đảm bảo rằng sự định tỉ lệ thời gian trên cơ sở mẫu không dẫn đến sự sụt giảm có thể chấp nhận của chất lượng âm thanh.

Kết luận, chức năng cơ bản của bộ điều khiển vùng đệm chập chờn 100 đã được giải thích, và các sự cải thiện tùy chọn của nó cũng đã được giải thích. Hơn nữa, nên lưu ý rằng bộ điều khiển vùng đệm chập chờn 100 có thể được bổ sung bởi bất kỳ đặc điểm và chức năng nào được mô tả ở đây.

Bộ định tỉ lệ thời gian theo Fig.2

Fig.2 thể hiện sơ đồ khối dạng giản lược của bộ định tỉ lệ thời gian 200, theo phương án của sáng chế. Bộ định tỉ lệ thời gian 200 được tạo cấu hình để nhận tín hiệu âm thanh đầu vào 210 (ví dụ, dưới dạng chuỗi các mẫu được cung cấp bởi lõi bộ giải mã) và cung cấp, trên cơ sở của nó, phiên bản được định tỉ lệ thời gian 212 của tín hiệu âm thanh đầu vào. Bộ định tỉ lệ thời gian 200 được tạo cấu hình để tính toán hoặc ước lượng chất lượng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bởi sự định tỉ lệ thời gian của tín hiệu âm thanh đầu vào. Chức năng này có thể được thực hiện, ví dụ, bằng bộ phận tính toán. Hơn nữa, bộ định tỉ lệ thời gian 200 được tạo cấu hình để thực hiện sự định tỉ lệ thời gian của tín hiệu âm thanh đầu vào 210 phụ thuộc vào sự tính toán hoặc sự ước lượng của chất lượng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bởi sự định tỉ lệ thời gian, để nhờ đó thu được phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào 212. Chức năng này có thể, ví dụ, được thực hiện bởi bộ phận định tỉ lệ thời gian.

Theo đó, bộ định tỉ lệ thời gian có thể thực hiện sự điều khiển chất lượng để đảm bảo rằng sự sụt giảm quá mức chất lượng âm thanh được ngăn ngừa khi thực hiện sự định tỉ lệ thời gian. Ví dụ, bộ định tỉ lệ thời gian có thể được tạo cấu hình để dự báo (hoặc ước lượng), trên cơ sở của tín hiệu âm thanh đầu vào, liệu thuật toán định tỉ lệ thời gian được dự tính (như, ví dụ, thuật toán chồng lấp và cộng được thực hiện trên cơ sở các khối các mẫu (âm thanh) được dịch chuyển thời gian được mong muốn có kết

quả là chất lượng âm thanh tốt. Nói cách khác, bộ định tỉ lệ thời gian có thể được tạo cấu hình để tính toán hoặc ước lượng chất lượng (mong muốn) của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bằng sự định tỉ lệ thời gian của tín hiệu âm thanh đầu vào trước sự định tỉ lệ thời gian của tín hiệu âm thanh đầu vào thực sự được thực hiện. Với mục đích này, bộ định tỉ lệ thời gian có thể, ví dụ, so sánh các phần của tín hiệu âm thanh đầu vào mà bao gồm thuật toán định tỉ lệ thời gian (ví dụ, trong các phần đã nêu của tín hiệu âm thanh đầu vào để được chồng lắp và được cộng để nhờ đó thực hiện sự định tỉ lệ thời gian). Kết luận, bộ định tỉ lệ thời gian 200 thường được tạo cấu hình để kiểm tra liệu nó có thể mong muốn rằng sự định tỉ lệ thời gian được dự tính sẽ có kết quả là chất lượng âm thanh đầy đủ của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào, và để quyết định liệu có thực hiện sự định tỉ lệ thời gian hoặc không trên cơ sở đó. Mặt khác, bộ định tỉ lệ thời gian có thể thích ứng bất kỳ tham số định tỉ lệ thời gian nào (ví dụ, sự định tỉ lệ thời gian giữa các khối các mẫu để được chồng lắp và cộng) phụ thuộc vào kết quả của sự ước lượng tính toán của chất lượng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bằng sự định tỉ lệ thời gian của tín hiệu âm thanh đầu vào.

Dưới đây, những cải thiện tùy chọn của bộ định tỉ lệ thời gian 200 sẽ được mô tả.

Trong phương án được ưu tiên, bộ định tỉ lệ thời gian được tạo cấu hình để thực hiện thuật toán chồng lắp và cộng sử dụng khối các mẫu thứ nhất của tín hiệu âm thanh đầu vào và khối các mẫu thứ hai của tín hiệu âm thanh đầu vào. Trong trường hợp này, bộ định tỉ lệ thời gian được tạo cấu hình để dịch chuyển thời gian khối các mẫu thứ hai so với khối các mẫu thứ nhất, và để chồng lắp và cộng khối các mẫu thứ nhất và khối các mẫu thứ hai được dịch chuyển thời gian, để nhờ đó thu được phiên bản định tỉ lệ thời gian của tín hiệu âm thanh đầu vào. Ví dụ, nếu sự rút ngắn thời gian được mong muốn, bộ định tỉ lệ thời gian có thể nhập vào số lượng các mẫu thứ nhất của tín hiệu âm thanh đầu vào và cung cấp, trên cơ sở của nó, số lượng các mẫu thứ hai của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào, trong đó số lượng các mẫu thứ hai là nhỏ hơn số lượng các mẫu thứ nhất. Để đạt được sự giảm số lượng các mẫu, số lượng các mẫu thứ nhất có thể được tách thành ít nhất khối các mẫu thứ nhất và khối các mẫu thứ hai (trong đó khối các mẫu thứ nhất và khối các mẫu thứ

hai có thể chồng lấp hoặc không chồng lấp), và khối các mẫu thứ nhất và khối các mẫu thứ hai có thể được dịch chuyển theo thời gian cùng nhau, sao cho các phiên bản được dịch chuyển theo thời gian của khối các mẫu thứ nhất và khối các mẫu thứ hai chồng lấp. Trong vùng chồng lấp giữa (các) phiên bản được dịch chuyển của khối các mẫu thứ nhất và khối các mẫu thứ hai, thuật toán chồng lấp và cộng được áp dụng. Thuật toán chồng lấp và cộng như vậy có thể được áp dụng mà không gây ra độ méo có thể nghe được đáng kể nếu khối các mẫu thứ nhất và khối các mẫu thứ hai là tương tự "đủ" trong miền chồng lấp (trong đó thuật toán chồng lấp và cộng được thực hiện) và ưu tiên cũng trong môi trường của vùng chồng lấp. Do đó, bằng các phần tín hiệu chồng lấp và cộng mà ban đầu không chồng lấp về thời gian, sự rút ngắn thời gian đạt được, vì tổng số lượng các mẫu được giảm một số lượng các mẫu mà không chồng lấp từ đầu (trong tín hiệu âm thanh đầu vào 210), nhưng được chồng lấp trong phiên bản được định tỉ lệ thời gian 212 của tín hiệu âm thanh đầu vào.

Ngược lại, sự kéo dài thời gian có thể cũng đạt được sử dụng thuật toán chồng lấp và cộng như vậy. Ví dụ, khối các mẫu thứ nhất và khối các mẫu thứ hai có thể được chọn để chồng lấp và có thể bao gồm toàn bộ sự mở rộng theo thời gian thứ nhất. Theo đó, khối các mẫu thứ hai có thể được dịch chuyển thời gian so với khối các mẫu thứ nhất, sao cho sự chồng lấp giữa khối các mẫu thứ nhất và khối các mẫu thứ hai được giảm. Nếu khối các mẫu thứ hai được dịch chuyển thời gian thích hợp với khối các mẫu thứ nhất, sự chồng lấp và cộng có thể được thực hiện, trong đó vùng chồng lấp giữa khối các mẫu thứ nhất và phiên bản được dịch chuyển thời gian của khối các mẫu thứ hai có thể ngắn hơn cả về số lượng các mẫu và về thời gian so với vùng chồng lấp ban đầu giữa khối các mẫu thứ nhất và khối các mẫu thứ hai. Theo đó, kết quả của thuật toán chồng lấp và cộng sử dụng khối các mẫu thứ nhất và phiên bản được dịch chuyển thời gian của khối các mẫu thứ hai có thể bao gồm sự mở rộng thời gian lớn hơn (cả về thời gian và cả về số lượng các mẫu) hơn là tổng sự mở rộng của khối các mẫu thứ nhất và khối các mẫu thứ hai trong dạng ban đầu của chúng.

Theo đó, rõ ràng rằng cả sự rút ngắn thời gian và sự kéo dài thời gian có thể thu được sử dụng thuật toán chồng lấp và cộng sử dụng khối các mẫu thứ nhất của tín hiệu âm thanh đầu vào và khối các mẫu thứ hai của các tín hiệu âm thanh đầu vào, trong đó khối các mẫu thứ hai được dịch chuyển thời gian so với khối các mẫu thứ nhất (hoặc

trong đó cả khối các mẫu thứ nhất và khối các mẫu thứ hai được dịch chuyển thời gian so với nhau).

Cụ thể hơn, bộ định tỉ lệ thời gian 200 được tạo cấu hình để tính toán và ước lượng chất lượng của thuật toán chồng lấp và cộng giữa khối các mẫu thứ nhất và phiên bản được dịch chuyển thời gian của khối các mẫu thứ hai, để tính toán và ước lượng chất lượng (mong muốn) của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bằng sự định tỉ lệ thời gian. Nên lưu ý rằng thường hiếm có thành phần lạ có thể nghe được nào nếu thuật toán chồng lấp và cộng được thực hiện cho các phần của các khối các mẫu mà tương tự đủ. Nói cách khác, chất lượng của thuật toán chồng lấp và cộng thường ảnh hưởng đáng kể đến chất lượng (mong muốn) của phiên bản được định tỉ lệ thời gian của các tín hiệu âm thanh đầu vào. Do đó, sự ước lượng (hoặc tính toán) chất lượng của thuật toán chồng lấp và cộng cung cấp số đo có thể tin cậy của chất lượng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào.

Cụ thể hơn, bộ định tỉ lệ thời gian 200 được tạo cấu hình để xác định sự dịch chuyển thời gian của khối các mẫu thứ hai so với khối các mẫu thứ nhất phụ thuộc vào sự xác định mức tương tự giữa khối các mẫu thứ nhất, hoặc phần (ví dụ, phần phía bên phải) của khối các mẫu thứ nhất, và khối các mẫu thứ hai được dịch chuyển thời gian, hoặc phần (ví dụ, phần phía bên trái) của khối các mẫu thứ hai được dịch chuyển thời gian. Nói cách khác, bộ định tỉ lệ thời gian có thể được tạo cấu hình để xác định, sự dịch chuyển thời gian nào giữa khối các mẫu thứ nhất và khối các mẫu thứ hai là thích hợp nhất để thu được kết quả chồng lấp và cộng tốt một cách thích đáng (hoặc ít nhất kết quả chồng lấp và bổ sung khả thi tốt nhất). Tuy nhiên, trong bước ("điều khiển chất lượng") bổ sung, nó có thể được xác nhận liệu sự dịch chuyển thời gian được xác định của khối các mẫu thứ hai so với khối các mẫu thứ nhất như vậy có thực sự mang lại kết quả chồng lấp và cộng đủ tốt (hoặc được mong muốn mang lại kết quả chồng lấp và cộng đủ tốt).

Cụ thể, bộ định tỉ lệ thời gian xác định thông tin về mức tương tự giữa khối các mẫu thứ nhất, hoặc phần (ví dụ, phần phía bên phải) của khối các mẫu thứ nhất, và khối các mẫu thứ hai, hoặc phần (ví dụ, phần phía bên trái) của khối các mẫu thứ hai,

đối với nhiều sự dịch chuyển thời gian khác nhau giữa khối các mẫu thứ nhất và khối các mẫu thứ hai, và xác định sự dịch chuyển thời gian (ứng cử) để được sử dụng cho thuật toán chồng lấp và cộng trên cơ sở thông tin về mức tương tự đối với nhiều sự dịch chuyển thời gian khác nhau. Nói cách khác, sự tìm kiếm cho sự thích hợp tốt nhất có thể được thực hiện, trong đó thông tin về mức tương tự cho các sự dịch chuyển thời gian khác nhau có thể được so sánh, để tìm ra sự dịch chuyển thời gian cho mức tương tự tốt nhất có thể đạt được.

Cụ thể hơn, bộ định tỉ lệ thời gian được tạo cấu hình để xác định sự dịch chuyển thời gian của khối các mẫu thứ hai so với khối các mẫu thứ nhất, mà sự dịch chuyển thời gian được sử dụng cho thuật toán chồng lấp và cộng, phụ thuộc vào thông tin dịch chuyển thời gian đích. Nói cách khác, thông tin dịch chuyển thời gian đích mà có thể, ví dụ, thu được trên cơ sở đánh giá sự đầy của vùng đệm, sự chấp chừa và các tiêu chuẩn bổ sung có khả năng khác, có thể được xem xét (cân nhắc) khi xác định liệu sự dịch chuyển thời gian nào được sử dụng (ví dụ, như sự dịch chuyển thời gian ứng cử) cho thuật toán chồng lấp và cộng. Do đó, sự chồng lấp và cộng được thích ứng cho các yêu cầu của hệ thống.

Trong một số phương án, bộ định tỉ lệ thời gian có thể được tạo cấu hình để tính toán hoặc ước lượng chất lượng phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bằng sự định tỉ lệ thời gian của tín hiệu âm thanh đầu vào trên cơ sở thông tin về mức tương tự giữa khối các mẫu thứ nhất, hoặc phần (ví dụ, phần phía bên phải) của khối các mẫu thứ nhất, và khối các mẫu thứ hai, được dịch chuyển thời gian bằng sự dịch chuyển thời gian (ứng cử) được xác định, hoặc phần (ví dụ, phần phía bên trái) của khối các mẫu thứ hai, được dịch chuyển thời gian bằng sự dịch chuyển thời gian (ứng cử) được xác định. Thông tin đã nêu về mức tương tự cung cấp thông tin về chất lượng (mong muốn) của thuật toán chồng lấp và cộng, và do đó cũng cung cấp thông tin (ít nhất là số đo) về chất lượng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bởi sự định tỉ lệ thời gian. Trong một số trường hợp, thông tin được tính toán hoặc được ước lượng về chất lượng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bởi sự định tỉ lệ thời gian có thể được sử dụng để quyết định liệu sự định tỉ lệ thời gian có thực sự được thực hiện hay không (trong đó sự định tỉ lệ thời gian có thể được hoãn

trong trường hợp sau). Nói cách khác, bộ định tỉ lệ thời gian có thể được tạo cấu hình để quyết định, trên cơ sở thông tin về mức tương tự giữa khối các mẫu thứ nhất, hoặc phần (ví dụ, phần phía bên phải) của khối các mẫu thứ nhất, và khối các mẫu thứ hai, được dịch chuyển thời gian bởi sự dịch chuyển thời gian (ứng cử) được xác định, hoặc phần (ví dụ, phần phía bên trái) của khối mẫu thứ hai, được dịch chuyển thời gian bởi sự dịch chuyển thời gian (ứng cử) được xác định, liệu sự định tỉ lệ thời gian có thực sự được thực hiện (hay không). Ví dụ, cơ cấu điều khiển chất lượng, mà đánh giá thông tin được tính toán hoặc được ước lượng về chất lượng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bằng sự định tỉ lệ thời gian, nhưng thực sự dẫn đến sự loại bỏ sự định tỉ lệ thời gian (ít nhất đối với khối khung các mẫu âm thanh hiện thời) nếu mong muốn rằng sự sụt giảm quá mức của nội dung âm thanh có thể bị gây ra bởi sự định tỉ lệ thời gian.

Trong một số phương án, các số đo độ tương tự khác nhau có thể được sử dụng cho sự xác định ban đầu của sự dịch chuyển thời gian (ứng cử) giữa khối các mẫu thứ nhất và khối các mẫu thứ hai và đối với cơ cấu điều khiển chất lượng cuối cùng. Nói cách khác, bộ định tỉ lệ thời gian có thể được tạo cấu hình để dịch chuyển thời gian khối các mẫu thứ hai so với khối các mẫu thứ nhất, và để chồng lấp và cộng khối các mẫu thứ nhất và khối các mẫu thứ hai được dịch chuyển thời gian, để nhờ đó thu được phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào, nếu sự tính toán hoặc ước lượng chất lượng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bằng sự định tỉ lệ thời gian biểu thị chất lượng mà lớn hơn hoặc bằng giá trị ngưỡng chất lượng. Bộ định tỉ lệ thời gian có thể được tạo cấu hình để xác định sự dịch chuyển thời gian (ứng cử) của khối các mẫu thứ hai so với khối các mẫu thứ nhất phụ thuộc vào sự xác định mức tương tự, được đánh giá sử dụng số đo độ tương tự thứ nhất, giữa khối các mẫu thứ nhất, hoặc phần (ví dụ, phần phía bên phải) của khối các mẫu thứ nhất, và khối các mẫu thứ hai, hoặc phần (ví dụ, phần phía bên trái) của khối các mẫu thứ hai. Ngoài ra, bộ định tỉ lệ thời gian có thể được tạo cấu hình để tính toán hoặc ước lượng chất lượng phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bằng sự định tỉ lệ thời gian của tín hiệu âm thanh đầu vào trên cơ sở thông tin về mức tương tự, được đánh giá sử dụng số đo độ tương tự thứ hai, giữa khối các mẫu thứ nhất, hoặc phần (ví dụ, phần phía bên

phải) của khối các mẫu thứ nhất, và khối các mẫu thứ hai, được dịch chuyển thời gian bằng sự dịch chuyển thời gian (ứng cử) được xác định, hoặc phần (ví dụ, phần phía bên trái) của khối các mẫu thứ hai, được dịch chuyển thời gian bằng sự dịch chuyển thời gian (ứng cử) được xác định. Ví dụ, số đo độ tương tự thứ hai có thể phức tạp về tính toán hơn là số đo độ tương tự thứ nhất. Khái niệm như vậy là hữu ích, vì nó thường cần thiết để tính toán số đo độ tương tự thứ nhất nhiều lần mỗi thuật toán định tỉ lệ thời gian (để xác định sự dịch chuyển thời gian "ứng cử" giữa khối các mẫu thứ nhất và khối các mẫu thứ hai trong số nhiều giá trị dịch chuyển thời gian khả thi giữa khối các mẫu thứ nhất và khối các mẫu thứ hai). Ngược lại, số đo độ tương tự thứ hai thường chỉ cần để được tính toán một lần mỗi thuật toán dịch chuyển thời gian, ví dụ như sự kiểm tra chất lượng "cuối cùng" liệu sự dịch chuyển thời gian "ứng cử" được xác định sử dụng số đo chất lượng (ít phức tạp về tính toán) thứ nhất có thể được mong muốn có kết quả là chất lượng âm thanh đủ tốt. Kết quả là, sẽ khả thi để vẫn tránh sự thực hiện sự chồng lấp và cộng, nếu số đo độ tương tự thứ nhất biểu thị sự tương tự (hoặc ít nhất đủ) tốt một cách hợp lý giữa khối các mẫu thứ nhất (hoặc phần của chúng) và khối các mẫu thứ hai được dịch chuyển thời gian (hoặc phần của chúng) cho sự dịch chuyển thời gian "ứng cử" nhưng độ tương tự (và thường ý nghĩa hơn hoặc chính xác hơn) thứ hai biểu thị rằng sự định tỉ lệ thời gian sẽ không dẫn đến kết quả là chất lượng âm thanh đủ tốt. Do đó, sự ứng dụng sự điều khiển chất lượng (sử dụng số đo độ tương tự thứ hai) giúp tránh khỏi độ méo có thể nghe được trong sự định tỉ lệ thời gian.

Ví dụ, số đo độ tương tự thứ nhất có thể là sự tương quan chéo hoặc sự tương quan chéo được tiêu chuẩn hóa hoặc hàm chênh lệch độ lớn trung bình hoặc tổng các sai số được bình phương. Các số đo độ tương tự như vậy có thể thu được bằng cách hiệu quả về tính toán và đủ để tìm ra "sự thích hợp tốt nhất" giữa khối các mẫu thứ nhất (hoặc phần của nó) và khối các mẫu thứ hai (được dịch chuyển thời gian) (hoặc phần của nó) tức là, để xác định sự dịch chuyển thời gian "ứng cử". Ngoài ra, số đo độ tương tự thứ hai có thể, ví dụ, là tổ hợp các giá trị tương quan chéo hoặc các giá trị tương quan chéo được tiêu chuẩn hóa cho nhiều sự dịch chuyển thời gian khác nhau. Số đo độ tương tự như vậy cung cấp sự chính xác và giúp cân nhắc các thành phần tín hiệu bổ sung (như, ví dụ, sóng hài) hoặc sự ổn định của tín hiệu âm thanh khi đánh giá

chất lượng (mong muốn) của sự định tỉ lệ thời gian. Tuy nhiên, số đo độ tương tự thứ hai được yêu cầu về tính toán hơn số đo độ tương tự thứ nhất, sao cho nó không hiệu quả về tính toán để áp dụng số đo độ tương tự thứ hai khi tìm kiếm sự dịch chuyển thời gian "ứng cử".

Dưới đây, một số tùy chọn cho sự xác định số đo độ tương tự thứ hai sẽ được mô tả. Trong một số phương án, số đo độ tương tự thứ hai có thể là tổ hợp của các sự tương quan chéo cho ít nhất bốn sự dịch chuyển thời gian khác nhau. Ví dụ, số đo độ tương tự thứ hai có thể là tổ hợp của giá trị tương quan chéo thứ nhất và giá trị tương quan chéo thứ hai, mà thu được cho sự dịch chuyển thời gian mà được tạo khoảng cách bằng bội số nguyên của thời khoảng chu kỳ của tần số cơ bản của nội dung âm thanh của khối các mẫu thứ nhất hoặc khối các mẫu thứ hai, và của giá trị tương quan chéo thứ ba và giá trị tương quan chéo thứ tư, mà thu được cho các sự dịch chuyển thời gian mà được chèn khoảng trống bởi bội số nguyên của thời khoảng chu kỳ của tần số cơ bản của nội dung âm thanh. Sự dịch chuyển thời gian mà giá trị tương quan chéo thứ nhất có thể được chèn khoảng trống từ sự dịch chuyển thời gian mà giá trị tương quan chéo thứ ba thu được, bởi bội số lẻ của một nửa thời khoảng chu kỳ của tần số cơ bản của nội dung âm thanh. Nếu nội dung âm thanh (được biểu diễn bởi tín hiệu âm thanh đầu vào) về cơ bản là ổn định, và được chiếm ưu thế bởi tần số cơ bản, có thể mong muốn rằng giá trị tương quan chéo thứ nhất và giá trị tương quan chéo thứ hai mà có thể, ví dụ, được thông thường hóa, đều gần 1. Tuy nhiên, vì giá trị tương quan chéo thứ ba và giá trị tương quan chéo thứ tư đều thu được cho sự dịch chuyển thời gian mà được chèn khoảng trống, bởi bội số lẻ của một nửa thời khoảng chu kỳ của tần số cơ bản, từ các sự dịch chuyển thời gian mà giá trị tương quan chéo thứ nhất và giá trị tương quan chéo thứ hai thu được, có thể mong muốn rằng giá trị tương quan chéo thứ ba và giá trị tương quan chéo thứ tư là ngược đối với giá trị tương quan chéo thứ nhất và giá trị tương quan chéo thứ hai trong trường hợp nội dung âm thanh về cơ bản ổn định và được chiếm ưu thế bởi tần số cơ bản. Theo đó, sự tổ hợp ý nghĩa có thể được tạo ra trên cơ sở giá trị tương quan chéo thứ nhất, giá trị tương quan chéo thứ hai, giá trị tương quan chéo thứ ba và giá trị tương quan chéo thứ tư, mà biểu thị liệu tín hiệu âm thanh là đủ ổn định và được chiếm ưu thế bởi tần số cơ bản trong vùng chồng lấp và cộng (ứng cử).

Nên lưu ý rằng các số đo độ tương tự đặc biệt ý nghĩa có thể thu được bằng cách tính toán số đo độ tương tự q theo

$$q = c(p) * c(2*p) + c(3/2*p) * c(1/2*p)$$

hoặc theo

$$q = c(p) * c(-p) + c(-1/2*p) * c(1/2*p).$$

Như nêu trên, $c(p)$ là giá trị tương quan chéo giữa khối các mẫu thứ nhất (hoặc phần của nó) và khối các mẫu thứ hai (hoặc phần của nó), mà được dịch chuyển trong thời gian (ví dụ, đối với vị trí theo thời gian ban đầu nằm trong nội dung âm thanh đầu vào) bằng thời khoảng chu kì p của tần số cơ bản của nội dung âm thanh của khối các mẫu thứ nhất và/hoặc khối các mẫu thứ hai (trong đó tần số cơ bản của nội dung âm thanh thường về cơ bản giống nhau trong khối các mẫu thứ nhất và trong khối các mẫu thứ hai). Nói cách khác, giá trị tương quan chéo được tính toán trên cơ sở khối các mẫu mà được lấy từ nội dung âm thanh đầu vào và ngoài ra được dịch chuyển thời gian so với nhau bởi thời khoảng chu kỳ p của tần số cơ bản của nội dung âm thanh đầu vào (trong đó thời khoảng chu kỳ p của tần số cơ bản có thể thu được, ví dụ, trên cơ sở sự ước lượng tần số cơ bản, sự tự tương quan, hoặc tương tự). Tương tự, $c(2*p)$ là giá trị tương quan chéo giữa khối các mẫu thứ nhất (hoặc phần của nó) và khối các mẫu thứ hai (hoặc phần của nó) mà được dịch chuyển trong thời gian bởi $2 * p$. Các định nghĩa tương tự cũng áp dụng cho $c(3/2*p)$, $c(1/2*p)$, $c(-p)$ và $c(-1/2*p)$, trong đó đối số của $c(.)$ chỉ rõ sự dịch chuyển thời gian.

Dưới đây, một số cơ cấu cho việc quyết định liệu sự định tỉ lệ thời gian nên được thực hiện hay không sẽ được giải thích, mà có thể tùy chọn được áp dụng trong bộ định tỉ lệ thời gian 200. Trong sự thực hiện, bộ định tỉ lệ thời gian 200 có thể được tạo cấu hình để so sánh giá trị chất lượng, mà được dựa trên sự tính toán hoặc ước lượng của chất lượng (mong muốn) của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bởi sự định tỉ lệ thời gian, với giá trị ngưỡng biến thiên, để quyết định liệu sự định tỉ lệ thời gian nên được thực hiện hay không. Theo đó, quyết định có thực hiện sự định tỉ lệ thời gian hay không có thể cũng được thực hiện phụ thuộc vào các trường hợp, như, ví dụ, lịch sử biểu diễn các sự định tỉ lệ thời gian trước đó.

Ví dụ, bộ định tỉ lệ thời gian có thể được tạo cấu hình để giảm giá trị ngưỡng có thể biến thiên, để nhờ đó tăng yêu cầu chất lượng (mà phải đạt được để cho phép sự định tỉ lệ thời gian), phản hồi việc tìm ra rằng chất lượng của sự định tỉ lệ thời gian sẽ là không đủ cho một hoặc nhiều khối các mẫu trước đó. Theo đó, được đảm bảo rằng sự định tỉ lệ thời gian không được ngăn ngừa cho chuỗi các khung (hoặc các khối các mẫu) dài mà có thể gây ra sự chạy tràn vùng đệm hoặc sự chạy non vùng đệm. Hơn nữa, bộ định tỉ lệ thời gian có thể được tạo cấu hình để tăng giá trị ngưỡng biến thiên, để nhờ đó tăng sự yêu cầu chất lượng (mà phải đạt được để cho phép sự định tỉ lệ thời gian), phản hồi sự thật rằng sự định tỉ lệ thời gian đã được áp dụng cho các khối hoặc các mẫu trước đó. Theo đó, có thể ngăn ngừa việc quá nhiều khối hoặc mẫu tiếp theo được định tỉ lệ thời gian, trừ khi chất lượng rất tốt (được tăng với sự yêu cầu chất lượng thông thường) của sự định tỉ lệ thời gian có thể thu được. Theo đó, các thành phần lạ có thể được tránh, điều này có thể được gây ra nếu các điều kiện cho chất lượng của sự định tỉ lệ thời gian quá thấp.

Trong một số phương án, bộ định tỉ lệ thời gian có thể bao gồm bộ đếm thứ nhất được giới hạn phạm vi để đếm số lượng khối các mẫu hoặc số lượng khung mà đã được định tỉ lệ thời gian vì sự yêu cầu chất lượng tương ứng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bằng sự định tỉ lệ thời gian đã đạt được. Hơn nữa, bộ định tỉ lệ thời gian có thể cũng bao gồm bộ đếm thứ hai được giới hạn phạm vi để đếm số lượng khối các mẫu hoặc số lượng khung mà chưa được định tỉ lệ thời gian vì sự yêu cầu chất lượng tương ứng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bằng sự định tỉ lệ thời gian chưa đạt được. Trong trường hợp này, bộ định tỉ lệ thời gian có thể được tạo cấu hình để tính toán giá trị ngưỡng biến thiên phụ thuộc vào giá trị bộ đếm thứ nhất và phụ thuộc vào giá trị của bộ đếm thứ hai. Theo đó, "lịch sử" của sự định tỉ lệ thời gian (và cũng là lịch sử "chất lượng" có thể được xem như nỗ lực tính toán bớt đi.

Ví dụ, bộ định tỉ lệ thời gian có thể được tạo cấu hình để cộng giá trị mà tỉ lệ với giá trị của bộ đếm thứ nhất vào giá trị ngưỡng ban đầu, và để trừ giá trị mà tỉ lệ với giá trị của bộ đếm thứ hai từ đó (ví dụ, từ kết quả của việc cộng), để thu được giá trị ngưỡng biến thiên.

Dưới đây, một số chức năng quan trọng, mà có thể được cung cấp trong một số phương án của bộ định tỉ lệ thời gian 200 sẽ được tóm tắt. Tuy nhiên, nên lưu ý rằng chức năng được mô tả dưới đây không phải là các chức năng chủ yếu của bộ định tỉ lệ thời gian 200.

Trong sự thực hiện, bộ định tỉ lệ thời gian có thể được tạo cấu hình để thực hiện sự định tỉ lệ thời gian của tín hiệu âm thanh đầu vào phụ thuộc vào sự tính toán hoặc sự ước lượng chất lượng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bởi sự định tỉ lệ thời gian. Trong trường hợp này, sự tính toán hoặc sự ước lượng của chất lượng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào bao gồm sự tính toán hoặc sự ước lượng của các thành phần lạ trong phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào mà có thể sẽ gây ra bởi sự định tỉ lệ thời gian. Tuy nhiên, nên lưu ý rằng sự tính toán hoặc sự ước lượng của các thành phần lạ có thể được thực hiện bằng cách gián tiếp, ví dụ, bằng cách tính toán chất lượng của thuật toán chồng lấp và cộng. Nói cách khác, sự tính toán hoặc sự ước lượng chất lượng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể bao gồm sự tính toán hoặc sự ước lượng của các thành phần lạ trong phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào mà sẽ được gây ra bởi thuật toán chồng lấp và cộng của các khối các mẫu tiếp theo của tín hiệu âm thanh đầu vào (trong đó, về bản chất, một số sự dịch chuyển thời gian có thể được áp dụng cho các khối các mẫu tiếp theo).

Ví dụ, bộ định tỉ lệ thời gian có thể được tạo cấu hình để tính toán hoặc ước lượng chất lượng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bởi sự định tỉ lệ thời gian của tín hiệu âm thanh đầu vào phụ thuộc vào mức độ tương tự của các khối các mẫu (và có thể chồng lấp) tiếp theo của tín hiệu âm thanh đầu vào.

Trong phương án được ưu tiên, bộ định tỉ lệ thời gian có thể được tạo cấu hình để tính toán hoặc ước lượng liệu có các thành phần lạ có thể nghe được trong phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bởi sự định tỉ lệ thời gian của tín hiệu âm thanh đầu vào. Sự ước lượng các thành phần lạ có thể nghe được có thể được thực hiện bằng cách gián tiếp, như được nêu trên.

Như kết quả của việc điều khiển chất lượng âm thanh, sự định tỉ lệ thời gian có thể được thích hợp cho sự định tỉ lệ thời gian và được tránh tại các những thời điểm mà không thích hợp cho sự định tỉ lệ thời gian. Ví dụ, bộ định tỉ lệ thời gian có thể được tạo cấu hình để hoãn sự định tỉ lệ thời gian đến khung tiếp theo hoặc khỏi các mẫu tiếp theo nếu sự tính toán hoặc ước lượng chất lượng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bởi sự định tỉ lệ thời gian biểu thị chất lượng không đủ (ví dụ, chất lượng mà dưới giá trị ngưỡng chất lượng nhất định). Do đó, sự định tỉ lệ thời gian có thể được thực hiện tại thời điểm mà thích hợp hơn cho sự định tỉ lệ thời gian, sao cho ít thành phần lạ (cụ thể, các thành phần lạ có thể nghe được) được tạo ra. Nói cách khác, bộ định tỉ lệ thời gian có thể được tạo cấu hình để hoãn sự định tỉ lệ thời gian đến thời điểm mà sự định tỉ lệ thời gian ít có thể nghe được nếu sự tính toán hoặc sự ước lượng của chất lượng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bằng sự định tỉ lệ thời gian biểu thị chất lượng không đủ.

Kết luận lại, bộ định tỉ lệ thời gian 200 có thể được cải thiện bằng nhiều cách khác nhau, như được thảo luận trên đây.

Hơn nữa, nên lưu ý rằng bộ định tỉ lệ thời gian 200 có thể tùy chọn được tổ hợp với bộ điều khiển vùng đệm chấp chờn 100, trong đó bộ điều khiển vùng đệm chấp chờn 100 có thể quyết định liệu sự định tỉ lệ thời gian trên cơ sở mẫu, mà thường được thực hiện bởi bộ định tỉ lệ thời gian 200, nên được sử dụng hay không hoặc liệu sự định tỉ lệ thời gian trên cơ sở khung nên được sử dụng.

Bộ giải mã âm thanh theo Fig.3

Fig.3 thể hiện sơ đồ khối dạng giản lược của bộ giải mã âm thanh 300, theo phương án của sáng chế.

Bộ giải mã âm thanh 300 được tạo cấu hình để nhận nội dung âm thanh 310, mà có thể được xem như sự biểu diễn âm thanh đầu vào, và có thể, ví dụ, được biểu diễn dưới dạng các khung âm thanh. Hơn nữa, bộ giải mã 300 cung cấp, trên cơ sở của nó, nội dung âm thanh được giải mã 312, mà có thể, ví dụ, được biểu diễn dưới dạng các mẫu âm thanh được giải mã. Bộ giải mã âm thanh 300 có thể, ví dụ, bao gồm vùng đệm chấp chờn 320, mà được tạo cấu hình để nhận nội dung âm thanh đầu vào 310, ví

dụ, dưới dạng các khung âm thanh. Vùng đệm chấp chờn 320 được tạo cấu hình để làm đệm nhiều khung âm thanh biểu diễn các khối các mẫu âm thanh (trong đó khung đơn có thể biểu diễn một hoặc nhiều khối các mẫu âm thanh, và trong đó các mẫu âm thanh được biểu diễn bởi khung đơn có thể được chia nhỏ một cách hợp lý thành nhiều khối các mẫu chồng lấp hoặc không chồng lấp). Hơn nữa, vùng đệm chấp chờn 320 cung cấp các khung âm thanh 322 "được làm đệm", trong đó các khung âm thanh 322 có thể bao gồm cả hai khung âm thanh được bao gồm trong nội dung âm thanh đầu vào 310 và các khung âm thanh được tạo ra hoặc được chèn vào bởi vùng đệm chấp chờn (như, ví dụ, các khung âm thanh "không hoạt động" bao gồm thông tin tạo tín hiệu tạo tín hiệu sự tạo ra nhiều âm nền nhân tạo). Bộ giải mã âm thanh 300 còn bao gồm lõi bộ giải mã 330, mà nhận các khung âm thanh được làm đệm 322 từ vùng đệm chấp chờn 320 và cung cấp các mẫu âm thanh 332 (ví dụ, các khối với các mẫu âm thanh được kết hợp với các khung âm thanh) trên cơ sở các khung âm thanh 322 được nhận từ vùng đệm chấp chờn. Hơn nữa, bộ giải mã âm thanh 300 bao gồm bộ định tỉ lệ thời gian trên cơ sở mẫu 340, mà được tạo cấu hình để nhận các mẫu âm thanh 332 được cung cấp bởi lõi bộ giải mã 330 và để cung cấp, trên cơ sở của nó, các mẫu âm thanh được định tỉ lệ thời gian, mà tạo nên nội dung âm thanh được giải mã 312. Bộ định tỉ lệ thời gian trên cơ sở mẫu 340 được tạo cấu hình để cung cấp các mẫu âm thanh được định tỉ lệ thời gian (ví dụ, dưới dạng các khối các mẫu âm thanh) trên cơ sở các mẫu âm thanh 332 (tức là, trên cơ sở các khối các mẫu âm thanh được cung cấp bằng lõi các bộ giải mã). Hơn nữa, bộ giải mã âm thanh có thể bao gồm bộ điều khiển 350 tùy chọn. Bộ điều khiển vùng đệm chấp chờn 350, mà được sử dụng trong bộ giải mã âm thanh 300 có thể, ví dụ, giống hệt bộ điều khiển vùng đệm chấp chờn 100 theo điểm 1. Nói cách khác, bộ điều khiển vùng đệm chấp chờn 350 có thể được tạo cấu hình để lựa chọn sự định tỉ lệ thời gian trên cơ sở khung, mà được thực hiện bằng vùng đệm chấp chờn 320, hoặc sự định tỉ lệ thời gian trên cơ sở mẫu, mà được thực hiện bằng bộ định tỉ lệ thời gian trên cơ sở mẫu 340 theo cách thích ứng tín hiệu. Theo đó, bộ điều khiển vùng đệm chấp chờn 350 có thể nhận nội dung âm thanh đầu vào 310, hoặc thông tin về nội dung âm thanh đầu vào 310 như tín hiệu âm thanh 110, hoặc như thông tin về tín hiệu âm thanh 110. Hơn nữa, bộ điều khiển vùng đệm chấp chờn 350 có thể cung cấp thông tin điều khiển 112 (như được mô tả với bộ điều khiển vùng đệm

chập chờn 100) đến vùng đệm chập chờn 320, và bộ điều khiển vùng đệm chập chờn 350 có thể cung cấp thông tin điều khiển 114, như được mô tả với bộ điều khiển vùng đệm chập chờn 100, đến bộ định tỉ lệ thời gian trên cơ sở mẫu 140. Theo đó, vùng đệm chập chờn 320 có thể được tạo cấu hình để thả hoặc chèn các khung âm thanh để thực hiện sự định tỉ lệ thời gian trên cơ sở khung. Hơn nữa, lỗi bộ giải mã 330 có thể được tạo cấu hình để thực hiện sự tạo ra nhiều âm nền nhân tạo phản hồi sự mang khung thông tin tạo tín hiệu biểu thị sự tạo ra nhiều âm nền nhân tạo. Theo đó, nhiều âm nền nhân tạo có thể được tạo ra bởi lỗi bộ giải mã 330 phản hồi sự chèn khung "hoạt động" (bao gồm thông tin tạo tín hiệu biểu thị rằng nhiều âm nền nhân tạo nên được tạo ra) vào trong vùng đệm chập chờn 320. Nói cách khác, dạng đơn giản của sự định tỉ lệ thời gian có thể dẫn đến một cách hiệu quả sự tạo ra khung bao gồm nhiều âm nền nhân tạo, mà được khởi động bằng sự chèn khung "không hoạt động" vào vùng đệm chập chờn (mà có thể được thực hiện phản hồi thông tin điều khiển 112 được cung cấp bởi bộ điều khiển vùng đệm chập chờn). Hơn nữa, lỗi bộ giải mã có thể được tạo cấu hình để thực hiện "sự che giấu" phản hồi vùng đệm chập chờn trống. Sự che giấu như vậy có thể bao gồm sự tạo ra thông tin âm thanh cho khung "bị thiếu" (vùng đệm chập chờn trống) trên cơ sở thông tin âm thanh của một hoặc nhiều khung đứng trước khung âm thanh bị thiếu. Ví dụ, sự dự báo có thể được sử dụng, cho rằng nội dung âm thanh của khung âm thanh bị thiếu là "sự tiếp tục" của nội dung âm thanh của một hoặc nhiều khung âm thanh đứng trước khung âm thanh bị thiếu. Tuy nhiên, bất kỳ khái niệm che giấu sự mất khung đã biết trong tình trạng kỹ thuật có thể được sử dụng bởi lỗi bộ giải mã. Do đó, bộ điều khiển vùng đệm chập chờn 350 có thể chỉ dẫn vùng đệm chập chờn 320 (hoặc lỗi bộ giải mã 330) để bắt đầu việc che giấu trong trường hợp vùng đệm chập chờn 320 chạy hết. Tuy nhiên, lỗi bộ giải mã có thể thực hiện sự che giấu ngay cả khi không có tín hiệu điều khiển rõ ràng, dựa trên hiểu biết của nó.

Hơn nữa, nên lưu ý rằng bộ định tỉ lệ thời gian trên cơ sở mẫu 340 có thể bằng bộ định tỉ lệ thời gian 200 được mô tả với Fig.20 Theo đó, tín hiệu âm thanh đầu vào 210 có thể tương ứng với các mẫu âm thanh 332, và phiên bản được định tỉ lệ thời gian 212 của tín hiệu âm thanh đầu vào có thể tương ứng với các mẫu âm thanh được định tỉ lệ thời gian 342. Theo đó, bộ định tỉ lệ thời gian 340 có thể được tạo cấu hình để thực hiện sự định tỉ lệ thời gian của tín hiệu âm thanh đầu vào phụ thuộc vào sự tính toán

hoặc sự ước lượng chất lượng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bởi sự định tỉ lệ thời gian. Bộ định tỉ lệ thời gian trên cơ sở mẫu 340 có thể được điều khiển bởi bộ điều khiển vùng đệm chấp chèn 350, trong đó thông tin điều khiển 114 được cung cấp bởi bộ điều khiển vùng đệm chấp chèn đến bộ định tỉ lệ thời gian trên cơ sở mẫu 340 có thể biểu thị liệu sự định tỉ lệ thời gian trên cơ sở mẫu nên được thực hiện hay không. Trong đó, thông tin điều khiển 114 có thể, ví dụ, biểu thị lượng định tỉ lệ thời gian mong muốn để được thực hiện bởi bộ định tỉ lệ thời gian trên cơ sở mẫu 340.

Nên lưu ý rằng bộ định tỉ lệ thời gian 300 có thể được bổ sung bởi bất kỳ đặc điểm và chức năng nào được mô tả đối với bộ điều khiển vùng đệm chấp chèn 100 và/hoặc đối với bộ định tỉ lệ thời gian 200. Hơn nữa, bộ giải mã âm thanh 300 có thể cũng được bổ sung bằng các đặc điểm và chức năng khác được mô tả trong đây, ví dụ, đối với các hình từ Fig.4 đến Fig.15.

Bộ giải mã âm thanh theo Fig.4

Fig.4 thể hiện sơ đồ khối dạng giản lược của bộ giải mã âm thanh 400, theo phương án của sáng chế. Bộ giải mã âm thanh 400 được tạo cấu hình để nhận các gói 410, mà có thể bao gồm sự biểu diễn được gói của một hoặc nhiều khung âm thanh. Hơn nữa, bộ giải mã 400 cung cấp nội dung âm thanh được giải mã 412, ví dụ, dưới dạng các mẫu âm thanh. Các mẫu âm thanh có thể, ví dụ, được biểu diễn dưới định dạng "PCM" (tức là, trong dạng được điều biến mã xung, ví dụ, dưới dạng chuỗi các giá trị kỹ thuật số biểu diễn các mẫu của dạng sóng âm thanh).

Bộ giải mã âm thanh 400 bao gồm bộ mở gói 420, mà được tạo cấu hình để nhận các gói 410 và để cung cấp, trên cơ sở của nó, các khung được mở gói 422. Hơn nữa, bộ mở gói được tạo cấu hình để trích, từ các gói 410, cái gọi là "cờ hiệu SID", mà tạo tín hiệu khung âm thanh "không hoạt động" (tức là, khung âm thanh mà sự tạo nhiều âm nền nhân tạo nên được sử dụng, hơn là sự giải mã chi tiết "thông thường" của nội dung âm thanh). Thông tin cờ hiệu SID được chỉ rõ với 424. Hơn nữa, bộ mở gói cung cấp tem thời gian giao thức vận chuyển thời gian thực (cũng được chỉ rõ là "RTP TS") và tem thời gian đến (cũng được chỉ rõ như "TS đến"). Thông tin tem thời gian được chỉ rõ với 426. Hơn nữa, bộ giải mã âm thanh 400 bao gồm vùng đệm khử chấp chèn

430 (cũng được chỉ rõ vẫn tất là vùng đệm chập chờn 430), mà nhận các khung được mở gói 422 từ bộ mở gói 420, và cung cấp các khung được làm đệm 432 (và có khả năng cũng là các khung được chèn) cho lõi bộ giải mã 440. Hơn nữa, vùng đệm khử chập chờn 430 nhận thông tin điều khiển 434 cho sự định tỉ lệ (thời gian) trên cơ sở khung từ logic điều khiển. Ngoài ra, vùng đệm khử chập chờn 430 cung cấp thông tin phản hồi định tỉ lệ 436 cho sự ước lượng độ trễ phát ra. Bộ giải mã âm thanh 400 cũng bao gồm bộ định tỉ lệ thời gian (cũng được chỉ rõ như "TSM") 450, mà nhận các mẫu âm thanh được giải mã 442 (ví dụ, dưới dạng dữ liệu được điều biến mã xung) từ lõi bộ giải mã 440, trong đó lõi bộ giải mã 440 cung cấp các mẫu âm thanh được giải mã 442 trên cơ sở các khung được tạo đệm hoặc được chèn 432 được nhận từ vùng đệm khử chập chờn 430. Bộ định tỉ lệ thời gian 450 cũng nhận thông tin điều khiển 444 cho sự định tỉ lệ (thời gian) trên cơ sở mẫu từ logic điều khiển và cung cấp thông tin phản hồi sự định tỉ lệ 446 cho sự ước lượng độ trễ phát ra. Bộ định tỉ lệ thời gian 450 cũng cung cấp các mẫu được định tỉ lệ thời gian 448, mà có thể biểu diễn nội dung âm thanh được định tỉ lệ thời gian dưới dạng được điều biến mã xung. Bộ giải mã âm thanh 400 cũng bao gồm vùng đệm PCM 460, mà bao gồm các mẫu được định tỉ lệ thời gian 448 và làm đệm các mẫu được định tỉ lệ thời gian 448. Hơn nữa, vùng đệm PCM 460 cung cấp phiên bản được làm đệm của các mẫu được định tỉ lệ thời gian 448 như sự biểu diễn của nội dung âm thanh được giải mã 412. Hơn nữa, vùng đệm PCM 460 có thể cung cấp thông tin độ trễ 462 cho logic điều khiển.

Bộ giải mã âm thanh 400 cũng bao gồm sự ước lượng độ trễ đích 470, mà nhận thông tin 424 (ví dụ, cờ hiệu SID) cũng như thông tin tem thời gian 426 bao gồm tem thời gian RTP và tem thời gian đến. Trên cơ sở của thông tin này, sự ước lượng độ trễ đích 470 cung cấp thông tin độ trễ đích 472, mà mô tả độ trễ mong muốn, ví dụ, độ trễ mong muốn mà nên được gây ra bởi vùng đệm khử chập chờn 430, bộ giải mã 440, bộ định tỉ lệ thời gian 450 và vùng đệm PCM 460. Ví dụ, sự ước lượng độ trễ đích 470 có thể tính toán hoặc ước lượng thông tin độ trễ đích 472 sao cho độ trễ không được chọn không nhất thiết lớn nhưng đủ để bù cho một số sự chập chờn của các gói 410. Hơn nữa, bộ giải mã âm thanh 400 bao gồm sự xác định độ trễ phát ra 480, mà được tạo cấu hình để nhận thông tin phản hồi định tỉ lệ 436 từ vùng đệm khử chập chờn 430 và thông tin phản hồi định tỉ lệ 446 từ bộ định tỉ lệ thời gian 460. Ví dụ, thông tin phản

hỏi 436 có thể mô tả sự định tỉ lệ thời gian mà được thực hiện bởi vùng đệm khử chập chòn. Hơn nữa, thông tin phản hồi sự định tỉ lệ thời gian 446 mô tả sự định tỉ lệ thời gian mà được thực hiện bởi bộ định tỉ lệ thời gian 450. Đối với thông tin phản hồi sự định tỉ lệ thời gian 446, nên lưu ý rằng sự định tỉ lệ thời gian được thực hiện bởi bộ định tỉ lệ thời gian 450 thường thích ứng tín hiệu sao cho sự định tỉ lệ thời gian thực tế mà được mô tả bởi thông tin phản hồi định tỉ lệ 446 có thể khác với sự định tỉ lệ thời gian mong muốn mà có thể được mô tả bởi thông tin định tỉ lệ thời gian trên cơ sở mẫu 444. Kết luận, thông tin phản hồi định tỉ lệ 436 và thông tin phản hồi định tỉ lệ 446 có thể mô tả sự định tỉ lệ thời gian thực tế, mà có thể khác với sự định tỉ lệ thời gian mong muốn vì sự thích ứng tín hiệu được cung cấp phù hợp với một số khía cạnh của sáng chế.

Hơn nữa, bộ giải mã âm thanh 400 cũng bao gồm logic điều khiển 490, mà thực hiện sự điều khiển (chính) của bộ giải mã âm thanh. Logic điều khiển 490 nhận thông tin 424 (ví dụ, cờ hiệu SID) từ bộ mở gói 420. Ngoài ra, logic điều khiển 490 nhận thông tin độ trễ đích 472 từ sự ước lượng độ trễ đích 470, thông tin độ trễ phát ra 482 từ sự ước lượng độ trễ phát ra 480 (trong đó thông tin độ trễ phát ra 482 mô tả độ trễ thực sự, mà được suy ra từ sự ước lượng độ trễ phát ra 480 trên cơ sở thông tin phản hồi định tỉ lệ 436 và thông tin định tỉ lệ 446). Hơn nữa, logic điều khiển 490 (tùy chọn) nhận thông tin độ trễ 462 từ độ trễ PCM 460 (trong đó, mặt khác, thông tin độ trễ của vùng đệm PCM có thể là số lượng được định trước). Trên cơ sở thông tin nhận được, logic điều khiển 490 cung cấp thông tin định tỉ lệ trên cơ sở khung 434 và thông tin định tỉ lệ trên cơ sở mẫu 442 cho vùng đệm khử chập chòn 430 và cho bộ định tỉ lệ thời gian 450. Theo đó, logic điều khiển thiết lập thông tin định tỉ lệ trên cơ sở khung 434 và thông tin định tỉ lệ trên cơ sở mẫu 442 phụ thuộc vào thông tin độ trễ đích 472 và thông tin độ trễ phát ra 482 bằng cách thích ứng tín hiệu, xem xét một hoặc nhiều đặc điểm của nội dung âm thanh (như, ví dụ, câu hỏi liệu có khung "không hoạt động" mà sự tạo ra nhiều âm nền nhân tạo nên được thực hiện phù hợp với sự tạo tín hiệu được tạo ra bởi cờ hiệu SID).

Nên lưu ý rằng logic điều khiển 490 có thể thực hiện một số hoặc tất cả các chức năng của bộ điều khiển vùng đệm chập chòn 100, trong đó thông tin 424 có thể tương ứng với thông tin 110 về tín hiệu âm thanh, trong đó thông tin điều khiển 112 có thể

tương ứng thông tin định tỉ lệ trên cơ sở khung 434, và trong đó thông tin điều khiển 114 có thể tương ứng thông tin định tỉ lệ trên cơ sở mẫu 444. Tương tự, nên lưu ý rằng bộ định tỉ lệ thời gian 450 có thể thực hiện một số hoặc tất cả các chức năng của bộ định tỉ lệ thời gian 200 (hoặc ngược lại), trong đó tín hiệu âm thanh đầu vào 210 tương ứng các mẫu âm thanh được giải mã 442, và trong đó phiên bản được định tỉ lệ thời gian 212 của tín hiệu âm thanh đầu vào tương ứng với các mẫu âm thanh được định tỉ lệ thời gian 448.

Hơn nữa, nên lưu ý rằng bộ giải mã âm thanh 400 tương ứng bộ giải mã âm thanh 300, sao cho bộ giải mã âm thanh 300 có thể thực hiện một số hoặc tất cả các chức năng được mô tả đối với bộ giải mã âm thanh 400, và ngược lại. Vùng đệm chấp chờn 320 tương ứng vùng đệm khử chấp chờn 430, lõi bộ giải mã 330 tương ứng bộ giải mã 440, và bộ định tỉ lệ thời gian 340 tương ứng bộ định tỉ lệ thời gian 450. Bộ điều khiển 350 tương ứng logic bộ điều khiển 490.

Dưới đây, một số chi tiết bổ sung liên quan đến chức năng của bộ giải mã âm thanh 400 sẽ được cung cấp. Cụ thể, sự quản lý vùng đệm chấp chờn (jitter buffer management - JBM) được đề xuất sẽ được mô tả.

Giải pháp quản lý vùng đệm chấp chờn (jitter buffer management - JBM) được mô tả, mà có thể được sử dụng để nạp các gói đã nhận 410 với các khung, chứa dữ liệu tiếng nói hoặc âm thanh được mã hóa, vào trong bộ giải mã 440 trong khi duy trì sự phát ra liên tiếp. Trong sự truyền thông trên cơ sở gói, ví dụ, truyền giọng nói trên giao thức IP (voice-over-internet-protocol - VoIP), các gói (ví dụ, các gói 410) thường dễ bị thay đổi các thời gian truyền dẫn và bị mất trong khi truyền dẫn, điều này dẫn đến sự chấp chờn giữa các lần đến liên tiếp và các gói bị mất cho bộ nhận (ví dụ, bộ nhận bao gồm bộ giải mã âm thanh 400). Do đó, sự quản lý vùng đệm chấp chờn và các giải pháp che giấu sự mất gói được mong muốn cho phép tín hiệu đầu ra liên tiếp mà không bị lấp.

Sau đây, sự tổng quan giải pháp sẽ được đề xuất. Trong trường hợp sự quản lý vùng đệm chấp chờn được mô tả như trên, dữ liệu được mã hóa nằm trong các gói RTP (ví dụ, các gói 410) đầu tiên được mở gói (ví dụ, sử dụng bộ mở gói 420) và các khung thu được (ví dụ, các khung 422) với dữ liệu được giải mã (ví dụ, dữ liệu tiếng

nằm trong khung được mã hóa AMR-WB) được đưa vào trong vùng đệm khử chập chờn (ví dụ, vùng đệm khử chập chờn 430). Khi dữ liệu được điều biến mã xung được yêu cầu cho sự phát ra, nó cần được làm cho sẵn có bởi bộ giải mã (ví dụ, bởi bộ giải mã 440). Đối với mục đích này, các khung (ví dụ, các khung 432) được kéo từ vùng đệm khử chập chờn (ví dụ, từ vùng đệm khử chập chờn 430). Bằng cách sử dụng vùng đệm khử chập chờn, sự dao động trong thời gian đến có thể được bù. Để điều khiển độ sâu của vùng đệm, sự thay đổi sự định tỉ lệ thời gian (time scale modification - TSM) được áp dụng (trong đó sự thay đổi sự định tỉ lệ thời gian được chỉ rõ một cách vắn tắt là sự định tỉ lệ thời gian). Sự thay đổi sự định tỉ lệ thời gian có thể xảy ra trên cơ sở khung được mã hóa (ví dụ, nằm trong vùng đệm khử chập chờn 430) hoặc trong môđun riêng (Ví dụ, nằm trong bộ định tỉ lệ thời gian 450), cho phép sự thích ứng hạt mịn hơn của tín hiệu đầu ra PCM (ví dụ, của tín hiệu đầu ra PCM 448 hoặc của tín hiệu đầu ra PCM 412).

Khái niệm được mô tả như trên được minh họa, ví dụ, trên Fig.4 mà thể hiện tổng quan sự quản lý vùng đệm chập chờn. Để điều khiển độ sâu của vùng đệm khử chập chờn (ví dụ, vùng đệm khử chập chờn 430) và do đó cũng các mức độ của sự định tỉ lệ thời gian nằm trong vùng đệm khử chập chờn (ví dụ, vùng đệm khử chập chờn 430) và/hoặc môđun TSM (ví dụ, nằm trong bộ định tỉ lệ thời gian 450), logic điều khiển (ví dụ, logic điều khiển 490, mà được hỗ trợ bởi sự ước lượng độ trễ đích 470 và sự ước lượng độ trễ phát ra 480) được sử dụng. Nó sử dụng thông tin trên độ trễ đích (ví dụ, thông tin 472) và độ trễ phát ra (ví dụ, thông tin 482) và liệu sự truyền dẫn gián đoạn (discontinuous transmission - DTX) kết hợp với sự tạo ra nhiễu âm nền nhân tạo (comfort noise generation - CNG) hiện được sử dụng (ví dụ, thông tin 424). Các giá trị độ trễ được tạo ra, ví dụ, từ các môđun riêng (ví dụ, các môđun 470 và 480) cho sự ước lượng độ trễ đích và phát ra, và bit hoạt động/không hoạt động (cờ hiệu SID) được cung cấp, ví dụ, bởi môđun bộ mở gói (ví dụ, bộ mở gói 420).

Bộ mở gói

Sau đây, bộ mở gói 420 sẽ được mô tả. Môđun bộ mở gói chia các gói RTP 410 thành các khung đơn (các bộ phận truy cập) 422. Nó cũng tính toán tem thời gian RTP cho tất cả các khung mà không phải là khung duy nhất hoặc khung đầu tiên trong gói.

Ví dụ, tem thời gian được chứa trong gói RTP được gán cho khung thứ nhất của nó. Trong trường hợp gộp (tức là, đối với các gói RTP chứa một hoặc nhiều hơn một khung đơn), tem thời gian cho các khung sau được tăng bởi thời khoảng khung được chia bởi sự định tỉ lệ của các tem thời gian RTP. Ngoài ra, đối với tem thời gian RTP, mỗi khung cũng được ghi nhãn với thời gian hệ thống mà tại đó gói RTP được nhận ("tem thời gian đến"). Như có thể thấy, thông tin tem thời gian RTP và thông tin tem thời gian đến 426 có thể được cung cấp, ví dụ, cho sự ước lượng độ trễ đích 470. Môđun mở gói cũng xác định liệu khung có hoạt động hoặc chứa bộ mô tả chèn im lặng (silence insertion descriptor - SID). Nên lưu ý rằng trong các giai đoạn không hoạt động, chỉ các khung SID được nhận trong một số trường hợp. Theo đó, thông tin 424, mà có thể, ví dụ, bao gồm cờ hiệu SID, được cung cấp cho logic điều khiển 490.

Vùng đệm khử chập chờn

Môđun vùng đệm khử chập chờn 430 lưu trữ các khung 422 được nhận trên mạng lưới (ví dụ, qua mạng lưới loại TCP/IP) cho đến khi sự giải mã (ví dụ, bởi bộ giải mã 440). Các khung 422 được chèn vào hàng được sắp xếp theo thứ tự tem thời gian RTP tăng dần để hoàn tác sự sắp xếp lại mà có thể xảy ra trong mạng lưới. Khung tại đầu của hàng có thể được nạp cho bộ giải mã 440 và sau đó được loại bỏ (ví dụ, từ vùng đệm khử chập chờn 430). Nếu hàng trống hoặc khung bị mất theo sự khác biệt tem thời gian của khung tại phía trước (của hàng) và khung đã đọc trước đó, khung trống được quay lại (ví dụ, từ vùng đệm khử chập chờn 430 đến bộ giải mã 440) để bắt đầu sự che giấu sự mất gói (nếu khung cuối cùng là hoạt động) hoặc sự tạo ra nhiều âm nền nhân tạo (nếu khung cuối cùng là "SID" hoặc không hoạt động) trong môđun bộ giải mã 440.

Nói cách khác, bộ giải mã 440 có thể được tạo cấu hình để tạo ra nhiều âm nền nhân tạo trong trường hợp nó được tạo tín hiệu, trong khung, mà nhiều âm nền nhân tạo nên được sử dụng, ví dụ, sử dụng cờ hiệu "SID" hoạt động. Mặt khác, bộ giải mã có thể cũng được tạo cấu hình để thực hiện sự che giấu mất gói, ví dụ, bằng cách cung cấp các mẫu âm thanh được dự báo (hoặc được ngoại suy) trong trường hợp khung (cuối) trước là hoạt động (tức là, sự tạo ra nhiều âm nền nhân tạo bị vô hiệu hóa) và

vùng đệm chấp chừa chạy hết (sao cho khung trống được cung cấp cho bộ giải mã 440 bởi vùng đệm chấp chừa 430).

Môđun vùng đệm chấp chừa 430 cũng cung cấp sự định tỉ lệ theo khung bằng cách bổ sung khung trống về phía trước (ví dụ, của hàng của vùng đệm chấp chừa) cho sự kéo dài hoặc thả khung tại phía trước (ví dụ, của hàng của vùng đệm chấp chừa) cho sự rút ngắn thời gian. Trong trường hợp các giai đoạn không hoạt động, vùng đệm khứ chấp chừa có thể vận hành như thể các khung "NO_DATA" được cộng hoặc thả.

Sự thay đổi sự định tỉ lệ thời gian (Time Scale Modification - TSM)

Dưới đây, sự thay đổi sự định tỉ lệ thời gian (Time Scale Modification - TSM), mà được chỉ rõ một cách ngắn gọn là bộ định tỉ lệ thời gian hoặc bộ định tỉ lệ thời gian trên cơ sở mẫu, sẽ được mô tả. Thuật toán chồng lấp - cộng trên cơ sở sự tương tự dạng sóng (waveform-similarity-based-overlap-add - WSOLA) trên cơ sở gói được thay đổi (xem, ví dụ, [Lia01]) với sự điều khiển chất lượng có sẵn được sử dụng để thực hiện sự thay đổi định tỉ lệ thời gian (được chỉ ra một cách vắn tắt như sự định tỉ lệ thời gian) của tín hiệu. Một số chi tiết có thể được thể hiện, ví dụ, trên Fig.9, sẽ được giải thích dưới đây. Mức định tỉ lệ thời gian là phụ thuộc tín hiệu; các tín hiệu mà sẽ tạo ra các thành phần lạ rất xấu khi được định tỉ lệ được phát hiện bởi bộ điều khiển chất lượng và các tín hiệu mức độ thấp, mà gần với sự im lặng, được định tỉ lệ bởi phạm vi lớn nhất có thể. Các tín hiệu được định tỉ lệ thời gian tốt, như các tín hiệu tuần hoàn, được định tỉ lệ bởi sự dịch chuyển được lấy từ bên trong. Sự dịch chuyển được lấy từ số đo độ tương tự, như sự tương quan chéo được bình thường hóa. Với sự chồng lấp - cộng (overlap-add - OLA), đầu cuối của khung hiện thời (cũng được chỉ ra như "khối các mẫu thứ hai" tại đây) được dịch chuyển (ví dụ, đối với sự bắt đầu của khung hiện thời, mà cũng được chỉ ra như "khối các mẫu thứ nhất" ở đây) để hoặc rút ngắn hoặc kéo dài khung.

Như đã được đề cập, các chi tiết bổ sung đối với sự thay đổi định tỉ lệ thời gian (time scale modification - TSM) sẽ được mô tả dưới đây, viện dẫn đến Fig.9, mà thể hiện WSOLA được thay đổi với bộ điều khiển chất lượng, và cũng viện dẫn đến các hình vẽ Fig.10a, Fig.10b và Fig.11.

Vùng đệm PCM

Sau đây, vùng đệm PCM sẽ được mô tả. Môđun thay đổi định tỉ lệ thời gian 450 thay đổi thời khoảng của các khung PCM được phát ra bởi môđun bộ giải mã với sự định tỉ lệ thay đổi thời gian. Ví dụ, 1024 mẫu (hoặc 2048 mẫu) có thể được phát ra bởi bộ giải mã 440 mỗi khung âm thanh 432. Ngược lại, số lượng thay đổi các mẫu âm thanh có thể được phát ra bởi bộ định tỉ lệ thời gian 450 mỗi khung âm thanh 432 do sự định tỉ lệ thời gian trên cơ sở mẫu. Ngược lại, thẻ âm thanh loa phát thanh (hoặc, thông thường, thiết bị đầu ra âm thanh) thường đòi hỏi sự tạo khung cố định, ví dụ, 20 miligiây (milliseconds - ms). Do đó, vùng đệm bổ sung với sự vận hành vào trước ra trước được sử dụng để áp dụng sự tạo khung cố định trên các mẫu đầu ra bộ định tỉ lệ thời gian 448.

Khi nhìn vào toàn bộ chuỗi, vùng đệm PCM 460 này không tạo ra độ trễ bổ sung. Đúng hơn, độ trễ chỉ được chia sẻ giữa vùng đệm khử chập chờn 430 và vùng đệm PCM 460. Tuy nhiên, mục đích là để giữ số lượng các mẫu được lưu trữ trong vùng đệm PCM 460 thấp nhất có thể, vì điều này làm tăng số lượng khung được lưu trữ trong vùng đệm khử chập chờn 430 và do đó giảm xác suất mất chậm (trong đó bộ giải mã che giấu khung bị mất mà được nhận muộn hơn).

Mã chương trình giả được thể hiện trên Fig.5 thể hiện thuật toán để điều khiển mức vùng đệm PCM. Như được thể hiện từ mã chương trình giả của Fig.5, kích thước khung thẻ âm thanh ("soundCardFrameSize") được tính toán trên cơ sở tốc độ mẫu ("sampleRate"), trong đó nó được giả thiết rằng, như ví dụ, rằng thời khoảng khung là 20ms. Theo đó, số lượng các mẫu mỗi khung thẻ âm thanh được biết. Theo đó, vùng đệm PCM được lấp đầy bởi các khung âm thanh giải mã 432 (cũng được chỉ ra như "accessUnit") cho đến khi số lượng các mẫu trong vùng đệm PCM ("pcmBuffer_nReadableSamples()") không còn nhỏ hơn số lượng các mẫu mỗi khung thẻ âm thanh ("soundCardFrameSize"). Thứ nhất, khung (cũng được chỉ ra như "accessUnit") thu được (hoặc được đòi hỏi) từ vùng đệm khử chập chờn 430, như được thể hiện với số tham chiếu 510. Theo đó, "khung" của các mẫu âm thanh thu được bằng cách giải mã khung 432 được đòi hỏi từ vùng đệm khử chập chờn, như được thể hiện bởi số tham chiếu 512. Theo đó, khung của các mẫu âm thanh được giải mã (ví dụ, được chỉ ra với 442) thu được. Theo đó, sự thay đổi sự định tỉ lệ thời gian được áp dụng cho khung của các mẫu âm thanh được giải mã 442, sao cho "khung"

của các mẫu âm thanh được định tỉ lệ thời gian 448 thu được, mà có thể được thể hiện tại số tham chiếu 514. Nên lưu ý rằng khung của các mẫu âm thanh được định tỉ lệ thời gian có thể bao gồm số lượng các mẫu lớn hơn hoặc số lượng các mẫu nhỏ hơn khung của các mẫu âm thanh được giải mã 442 được đưa vào bộ định tỉ lệ thời gian 450. Tiếp theo, khung của các mẫu âm thanh được định tỉ lệ thời gian 448 được chèn vào vùng đệm PCM 460, như có thể được thể hiện tại số tham chiếu 516.

Quy trình này được lặp đi lặp lại, cho đến khi số lượng đủ các mẫu âm thanh (được định tỉ lệ thời gian) có sẵn trong vùng đệm PCM 460. Miễn là số lượng đủ các mẫu (được định tỉ lệ thời gian) có sẵn trong vùng đệm PCM, "khung" của các mẫu âm thanh được định tỉ lệ thời gian (có chiều dài khung như được yêu cầu bởi thiết bị phát lại âm thanh, như thẻ âm thanh) được đọc ra từ vùng đệm PCM 460 và được gửi chuyển tiếp đến thiết bị phát lại âm thanh (ví dụ, đến thẻ âm thanh), như được thể hiện tại các số tham chiếu 520 và 522.

Sự ước lượng độ trễ đích

Dưới đây, sự ước lượng độ trễ đích, mà có thể được thực hiện bằng bộ ước lượng độ trễ đích 470, sẽ được mô tả. Độ trễ đích chỉ rõ độ trễ vùng đệm mong muốn giữa thời gian khi khung phía trước được phát và thời gian mà khung này có thể được nhận nếu có độ trễ truyền dẫn thấp nhất trong mạng lưới so với tất cả các khung hiện được chứa trong lịch sử của môđun ước lượng độ trễ đích. Để ước lượng độ trễ đích, hai bộ ước lượng chấp chờn khác nhau được sử dụng, một bộ ước lượng chấp chờn dài hạn và một bộ ước lượng chấp chờn ngắn hạn.

Sự ước lượng chấp chờn dài hạn

Để tính toán sự chấp chờn dài hạn, cấu trúc dữ liệu FIFO có thể được sử dụng. Khoảng thời gian được lưu trữ trong FIFO có thể khác với số lượng các mục nhập được lưu trữ nếu chế độ truyền dẫn gián đoạn (discontinuous transmission mode - DTX) được sử dụng. Vì lý do này, kích thước cửa sổ của FIFO được giới hạn bằng hai cách. Nó có thể chứa nhiều nhất 500 mục nhập (bằng 10 giây tại 50 gói trên giây) và nhiều nhất khoảng thời gian (sự chênh lệch tem thời gian RTP giữa gói mới nhất và gói cũ nhất) của 10 giây. Nếu nhiều mục nhập là để được lưu trữ, mục nhập cũ nhất được loại bỏ. Đối với mỗi gói RTP được nhận trên mạng lưới, mục nhập sẽ được bỏ

sung cho FIFO. Mục nhập chứa ba giá trị: độ trễ, độ lệch và tem thời gian RTP. Các giá trị này được tính toán từ thời gian nhận (ví dụ, được biểu diễn bởi tem thời gian nhận) và tem thời gian RTP của gói RTP, được thể hiện trong mã giả của Fig.6.

Như có thể được thể hiện tại các số tham chiếu 610 và 612, sự chênh lệch thời gian giữa các tem thời gian RTP của hai gói (ví dụ, các gói tiếp theo) được tính toán (cho ra "rtpTimeDiff") và sự chênh lệch giữa các tem thời gian nhận của hai gói (ví dụ, các gói tiếp theo) được tính toán (cho ra "rcvTimeDiff"). Hơn nữa, tem thời gian RTP được chuyển đổi từ cơ sở thời gian của thiết bị truyền dẫn đến cơ sở thời gian của thiết bị nhận, như có thể được thể hiện tại số tham chiếu 614, cho ra "rtpTimeTicks". Tương tự, các sự chênh lệch thời gian RTP (sự chênh lệch giữa các tem thời gian RTP) được chuyển đổi đến sự định tỉ lệ thời gian. cơ sở thời gian bộ nhận của thiết bị nhận, như có thể được thể hiện tại số tham chiếu 616, cho ra "rtpTimeDiff".

Tiếp theo, thông tin độ trễ ("delay") được cập nhật trên cơ sở của thông tin độ trễ trước, như có thể được thể hiện tại số tham chiếu 618. Ví dụ, nếu sự chênh lệch thời gian nhận (tức là, sự chênh lệch về thời gian khi các gói đã được nhận) là lớn hơn sự chênh lệch thời gian RTP (tức là, sự chênh lệch giữa thời gian tại đó các gói đã được gửi ra ngoài), có thể kết luận rằng độ trễ đã tăng. Hơn nữa, thông tin thời gian độ lệch ("offset") được tính toán, như có thể được thể hiện tại số tham chiếu 620, trong đó thông tin thời gian độ lệch biểu diễn sự chênh lệch giữa thời gian nhận (tức là, thời gian tại đó gói đã được nhận) và thời gian tại đó gói đã được gửi (như được định nghĩa bởi tem thời gian RTP, được chuyển đổi thành sự định tỉ lệ thời gian bộ nhận). Hơn nữa, thông tin độ trễ, thông tin thời gian độ lệch và thông tin tem thời gian RTP (được chuyển hóa thành sự định tỉ lệ thời gian bộ nhận) được bổ sung vào FIFO dài hạn, như có thể được thể hiện tại số tham chiếu 622.

Tiếp theo, một số thông tin hiện thời được lưu trữ như thông tin "trước" cho phép lặp tiếp theo, như có thể được thể hiện tại số tham chiếu 624.

Sự chậm chèn dài hạn có thể được tính toán như sự chênh lệch giữa giá trị độ trễ tối đa hiện được lưu trữ trong FIFO và giá trị độ trễ tối thiểu:

```
longTermJitter = longTermFifo_getMaxDelay() – longTermFifo_getMinDelay();
```

Sự ước lượng chấp chờn ngắn hạn

Sau đây, sự ước lượng chấp chờn ngắn hạn sẽ được mô tả. Sự ước lượng chấp chờn ngắn hạn được thực hiện, ví dụ, trong hai bước. Trong bước thứ nhất, sự ước lượng chấp chờn tương tự như được thực hiện cho sự ước lượng dài hạn được sử dụng với các thay đổi sau: kích thước cửa sổ của FIFO được giới hạn nhiều nhất 50 mục nhập và nhiều nhất khoảng thời gian là 1 giây. Giá trị chấp chờn thu được được tính toán như sự chênh lệch giữa giá trị độ trễ phân vị 94% hiện được lưu trữ trong FIFO (ba giá trị cao nhất được bỏ qua) và giá trị độ trễ tối thiểu:

```
shortTermJitterTmp = shortTermFifo1_getPercentileDelay(94) –
                    shortTermFifo1_getMinDelay();
```

Trong bước thứ hai, đầu tiên, các độ lệch khác nhau giữa các FIFO dài hạn và ngắn hạn được bù cho kết quả này:

```
shortTermJitterTmp += shortTermFifo1_getMinOffset();
shortTermJitterTmp -= longTermFifo_getMinOffset();
```

Kết quả này được cộng vào FIFO khác với kích thước cửa sổ nhiều nhất là 200 mục nhập và khoảng thời gian nhiều nhất là bốn giây. Cuối cùng, giá trị tối đa được lưu trữ trong FIFO được tăng đến bội số nguyên của kích thước khung và được sử dụng như sự chấp chờn ngắn hạn:

```
shortTermFifo2_add( shortTermJitterTmp );
shortTermJitter = ceil( shortTermFifo2_getMax() / 20.f ) * 20;
```

Sự ước lượng độ trễ đích bởi sự tổ hợp của sự ước lượng chấp chờn dài hạn/ngắn hạn

Để tính toán độ trễ đích (ví dụ, thông tin độ trễ đích 472), các sự ước lượng chấp chờn dài hạn và ngắn hạn (ví dụ, như được định nghĩa như trên là “longTermJitter” và “shortTermJitter”) được tổ hợp bằng các cách khác nhau phụ thuộc vào trạng thái hiện thời. Đối với các tín hiệu hoạt động (hoặc các phần tín hiệu, mà sự tạo ra nhiều âm nền nhân tạo không được sử dụng cho các phần tín hiệu này), phạm vi (ví dụ, được xác định bởi “targetMin” và “targetMax”) được sử dụng như độ trễ đích. Trong suốt DTX

và cho sự khởi động sau DTX, hai giá trị khác nhau được tính toán như độ trễ đích (ví dụ, “targetDtx” và “targetStartUp”).

Các chi tiết về việc làm cách nào các giá trị độ trễ đích khác nhau có thể được tính toán có thể được thể hiện, ví dụ, trên Fig.7. Như có thể được thể hiện tại số tham chiếu 710 và 712, các giá trị “targetMin” và “targetMax”, mà gán phạm vi cho các tín hiệu hoạt động, được tính toán trên cơ sở sự chấp chừa ngắn hạn (“shortTermJitter”) và sự chấp chừa dài hạn (“longTermJitter”). Sự tính toán của độ trễ đích trong suốt DTX (“targetDtx”) được thể hiện tại số tham chiếu 714, và sự tính toán của giá trị độ trễ đích cho sự khởi động (ví dụ, sau DTX) (“targetStartUp”) được thể hiện tại số tham chiếu 716.

Sự ước lượng độ trễ phát ra

Dưới đây, sự ước lượng độ trễ phát ra, mà có thể được thực hiện bằng bộ ước lượng độ trễ phát ra 480, sẽ được mô tả. Độ trễ phát ra chỉ rõ độ trễ vùng đệm giữa thời gian khi khung phía trước được phát và thời gian mà khung này có thể được nhận nếu có độ trễ truyền dẫn thấp nhất có thể trong mạng lưới so với tất cả các khung hiện được chứa trong lịch sử của môđun ước lượng độ trễ đích. Nó được tính toán theo miligiây sử dụng công thức sau:

$$\text{playoutDelay} = \text{prevPlayoutOffset} - \text{longTermFifo_getMinOffset()} + \text{pcmBufferDelay};$$

Biến số “prevPlayoutOffset” được tính toán lại bất cứ khi nào khung được nhận được khung được nhận hiện ra từ môđun vùng đệm khử chấp chừa 430 sử dụng thời gian hệ thống hiện thời theo miligiây và tem thời gian của khung được chuyển đổi sang miligiây:

$$\text{prevPlayoutOffset} = \text{sysTime} - \text{rtpTimestamp}$$

Để tránh việc “prevPlayoutOffset” sẽ bị lỗi thời nếu khung không có sẵn, biến số được cập nhật trong trường hợp định tỉ lệ thời gian trên cơ sở khung. Đối với sự kéo dài thời gian trên cơ sở khung, “prevPlayoutOffset” được tăng bởi thời khoảng của khung, và đối với sự rút ngắn trên cơ sở khung, “PrevPlayoutOffset” được giảm bởi thời

khoảng của khung. Biến số “pcmBufferDelay” mô tả thời khoảng của thời gian được làm đệm trong môđun vùng đệm PCM.

Logic điều khiển

Dưới đây, sự điều khiển (ví dụ, logic điều khiển 490) sẽ được mô tả một cách cụ thể. Tuy nhiên, nên lưu ý rằng logic điều khiển 800 theo Fig.8 có thể được bổ sung bởi bất kỳ đặc điểm và chức năng nào được mô tả đối với bộ điều khiển vùng đệm chấp chờn 100 và ngược lại. Hơn nữa, nên lưu ý rằng logic điều khiển 800 có thể chiếm vị trí của logic điều khiển 490 theo Fig.4, nhưng có thể tùy chọn bao gồm các đặc điểm và chức năng bổ sung. Hơn nữa, không phải đòi hỏi rằng tất cả các đặc điểm và chức năng được mô tả như trên đối với Fig.4 cũng thể hiện trong logic điều khiển 800 theo Fig.8, và ngược lại.

Fig.8 thể hiện lưu đồ của logic điều khiển 800, mà cũng có thể được thực hiện một cách tự nhiên trong phân cứng.

Logic điều khiển 800 bao gồm việc kéo 810 khung cho việc giải mã. Nói cách khác, khung được lựa chọn cho sự giải mã, và nó được xác định như sau làm cách nào để sự giải mã này nên được thực hiện. Trong sự kiểm tra 814, nó được kiểm tra liệu khung trước đó (ví dụ, khung trước đứng trước khung được kéo cho sự giải mã trong bước 810) là hoạt động hay không. Nếu nó được tìm ra trong sự kiểm tra 814 rằng khung trước không hoạt động, đường quyết định thứ nhất (nhánh) 820 được lựa chọn mà sẽ được sử dụng để thích ứng tín hiệu không hoạt động. Ngược lại, nếu được tìm ra trong sự kiểm tra 814 rằng khung trước là hoạt động, đường quyết định (nhánh) 830 được chọn, mà được sử dụng để thích ứng tín hiệu hoạt động. Đường quyết định thứ nhất 820 bao gồm xác định giá trị "gap" trong bước 840, trong đó giá trị khoảng trống mô tả sự chênh lệch giữa độ trễ phát ra và độ trễ đích. Hơn nữa, đường quyết định thứ nhất 820 bao gồm sự quyết định 850 phép toán định tỉ lệ thời gian để được thực hiện trên cơ sở giá trị khoảng trống. Đường quyết định thứ hai 830 bao gồm sự lựa chọn 860 sự định tỉ lệ thời gian phụ thuộc vào việc liệu độ trễ phát ra thực sự nằm trong khoảng độ trễ đích.

Dưới đây, các chi tiết bổ sung đối với đường quyết định thứ nhất 820 và đường quyết định thứ hai 830 sẽ được mô tả.

Trong bước 840 của đường quyết định thứ nhất 820, sự kiểm tra 842 được thực hiện liệu khung tiếp theo có hoạt động. Ví dụ, sự kiểm tra 842 có thể kiểm tra liệu khung được kéo cho sự giải mã trong bước 810 có hoạt động hay không. Cách khác, sự kiểm tra 842 có thể kiểm tra liệu khung sau khung được kéo cho sự giải mã trong bước 810 có hoạt động hay không. Nếu được tìm ra, trong sự kiểm tra 842, rằng khung tiếp theo không hoạt động, hoặc khung tiếp theo chưa có sẵn, biến số "gap" được thiết lập, trong bước 844, như sự chênh lệch giữa độ trễ phát ra thực sự (được định nghĩa bởi biến số "payoutDelay") và độ trễ đích DTX (được biểu diễn bởi biến số "targetDtx"), được mô tả như trên trong phần "Target Delay Estimation". Ngược lại, nếu được tìm ra trong sự kiểm tra 840 rằng khung tiếp theo hoạt động, biến số "gap" được thiết lập đến sự chênh lệch giữa độ trễ phát ra (được biểu diễn bởi biến số "payoutDelay") và độ trễ đích khởi động (như được định nghĩa bởi biến số "targetStartUp") trong bước 846.

Trong bước 850, đầu tiên nó được kiểm tra liệu biên độ của biến số "gap" là lớn hơn (hoặc bằng với) ngưỡng. Điều này được thực hiện trong sự kiểm tra 852. Nếu tìm ra rằng biên độ của biến số "gap" là nhỏ hơn (hoặc bằng) giá trị ngưỡng, không có sự định tỉ lệ thời gian được thực hiện. Ngược lại, nếu tìm ra rằng sự kiểm tra 852 mà biên độ của biến số "gap" là lớn hơn ngưỡng (hoặc bằng các giá trị ngưỡng, phụ thuộc vào sự thực hiện), nó được quyết định rằng sự định tỉ lệ là cần. Trong sự kiểm tra 854 khác, nó được kiểm tra liệu giá trị của biến số "gap" là dương hay âm (tức là, nếu biến số "gap" là lớn hơn 0 hay không). Nếu tìm ra rằng giá trị của biến số "gap" không lớn hơn 0 (tức là âm), khung được chèn vào vùng đệm khứ chấp chừa (sự kéo dài thời gian trên cơ sở khung trong bước 856), sao cho sự định tỉ lệ thời gian trên cơ sở khung được thực hiện. Điều này có thể, ví dụ, được tạo tín hiệu bởi thông tin định tỉ lệ trên cơ sở khung 434. Ngược lại, nếu tìm ra rằng trong sự kiểm tra 854, mà giá trị của biến số "gap" là lớn hơn 0, tức là, dương, khung được thả từ vùng đệm khứ chấp chừa (sự rút ngắn thời gian trên cơ sở khung trong bước 856), sao cho sự định tỉ lệ thời gian trên cơ sở khung được thực hiện. Điều này có thể được tạo tín hiệu sử dụng thông tin định tỉ lệ trên cơ sở khung 434.

Sau đây, nhánh quyết định thứ hai 860 sẽ được mô tả. Trong sự kiểm tra 862, nó được kiểm tra liệu độ trễ sự phát ra là lớn hơn (hoặc bằng) giá trị đích tối đa (tức là,

giới hạn trên của khoảng đích) mà được mô tả, ví dụ, bởi biến số “targetMax”). Nếu tìm ra rằng độ trễ phát ra là lớn hơn (hoặc bằng) giá trị đích tối đa, sự rút ngắn thời gian được thực hiện bởi bộ định tỉ lệ thời gian 450 (bước 866, sự rút ngắn thời gian trên cơ sở mẫu sử dụng TSM), sao cho sự định tỉ lệ thời gian trên cơ sở mẫu được thực hiện. Điều này có thể được tạo tín hiệu, ví dụ, bởi thông tin định tỉ lệ trên cơ sở mẫu 444. Tuy nhiên, nếu tìm ra rằng trong sự kiểm tra 862 rằng độ trễ phát ra là nhỏ hơn (hoặc bằng) độ trễ đích tối đa, sự kiểm tra 864 được thực hiện, trong đó nó được kiểm tra liệu độ trễ phát ra là nhỏ hơn (hoặc bằng) độ trễ đích tối thiểu, mà được mô tả, ví dụ, bởi biến số “targetMin”. Nếu tìm ra rằng độ trễ phát ra là nhỏ hơn (hoặc bằng) độ trễ đích tối thiểu, sự kéo dài thời gian được thực hiện bởi bộ định tỉ lệ thời gian 450 (bước 866, sự kéo dài thời gian trên cơ sở mẫu sử dụng TSM), sao cho sự định tỉ lệ thời gian trên cơ sở mẫu được thực hiện. Điều này có thể được tạo tín hiệu, ví dụ, bởi thông tin định tỉ lệ trên cơ sở mẫu 444. Tuy nhiên, nếu tìm ra rằng trong sự kiểm tra 864 mà độ trễ phát ra không nhỏ hơn (hoặc bằng) độ trễ đích tối thiểu, không có sự định tỉ lệ thời gian được thực hiện.

Tóm lại, môđun logic điều khiển (cũng được chỉ rõ như logic điều khiển sự quản lý vùng đệm chấp chờn) được thể hiện trên Fig.8 so sánh độ trễ thực tế (độ trễ phát ra) với độ trễ mong muốn (độ trễ đích). Trong trường hợp sự chênh lệch đáng kể, nó sẽ bắt đầu sự định tỉ lệ thời gian. Trong suốt nhiều âm nền nhân tạo (ví dụ, khi cờ hiệu SID hoạt động), sự định tỉ lệ thời gian trên cơ sở khung sẽ được gây ra và được thực hiện bởi môđun vùng đệm khử chấp chờn. Trong suốt các giai đoạn hoạt động, sự định tỉ lệ thời gian trên cơ sở mẫu được gây ra và được thực hiện bởi môđun TSM.

Fig.12 thể hiện ví dụ về sự ước lượng độ trễ đích và phát ra. Trục hoành 1210 của sự biểu diễn đồ thị 1200 mô tả thời gian, và trục tung 1212 của sự biểu diễn đồ thị 1200 mô tả độ trễ theo miligiây. Các dãy “targetMin” và “targetMax” tạo ra phạm vi độ trễ mong muốn bởi môđun ước lượng độ trễ đích tiếp theo sự chấp chờn mạng được tạo cửa sổ. Độ trễ phát ra “playoutDelay” thường nằm trong phạm vi, nhưng sự thích ứng có thể được làm trễ một chút vì sự thay đổi sự định tỉ lệ thời gian thích ứng tín hiệu.

Fig.13 thể hiện các thuật toán định tỉ lệ thời gian được thực hiện trong sự tạo vết của Fig.12. Trục hoành 1310 của sự biểu diễn đồ thị 1300 mô tả thời gian theo giây, và trục tung 1312 mô tả sự định tỉ lệ thời gian theo miligiây. Các giá trị dương biểu thị sự kéo dài thời gian, các giá trị âm biểu thị sự rút ngắn thời gian trong sự biểu diễn đồ thị 1300. Trong suốt khối tín hiệu truyền, cả hai vùng đệm chỉ bị trống một lần, và một khung được che giấu được chen cho sự kéo dài (thêm 20 miligiây tại 35 giây). Đối với các sự thích ứng khác, phương pháp định tỉ lệ thời gian trên cơ sở mẫu chất lượng cao hơn có thể được sử dụng mà dẫn đến việc thay đổi các tỉ lệ vì phương pháp thích ứng tín hiệu.

Tóm lại, độ trễ đích được thích ứng một cách động phản hồi sự tăng của sự chập chờn (và cũng phản hồi sự giảm của sự chập chờn) qua cửa sổ nhất định. Khi độ trễ đích tăng hoặc giảm, sự định tỉ lệ thời gian thường được thực hiện, trong đó sự quyết định về loại định tỉ lệ thời gian được thực hiện theo cách thích ứng tín hiệu. Miễn là khung hiện thời (hoặc khung trước đó) là hoạt động, sự định tỉ lệ thời gian trên cơ sở mẫu được thực hiện, trong đó độ trễ thực tế của sự định tỉ lệ thời gian trên cơ sở mẫu được thích ứng theo cách thích ứng tín hiệu để giảm các thành phần lạ. Theo đó, thường không có lượng cố định sự định tỉ lệ thời gian khi sự định tỉ lệ thời gian trên cơ sở mẫu được áp dụng. Tuy nhiên, khi vùng đệm chập chờn hết, cần thiết (hoặc có giá trị) - như sự tiến hành một cách ngoại lệ - để chen khung được che giấu (mà tạo thành sự định tỉ lệ thời gian trên cơ sở khung) thậm chí khung trước (hoặc khung hiện thời) hoạt động.

Sự thay đổi sự định tỉ lệ thời gian theo Fig.9

Sau đây, các chi tiết liên quan đến sự thay đổi định tỉ lệ thời gian sẽ được mô tả viện dẫn đến Fig.9. Nên lưu ý rằng sự thay đổi định tỉ lệ thời gian đã được mô tả một cách vắn tắt trong phần 5.4.3. Tuy nhiên, sự thay đổi sự định tỉ lệ thời gian, mà có thể, ví dụ, được thực hiện bởi bộ định tỉ lệ thời gian 150, sẽ được mô tả chi tiết hơn như sau.

Fig.9 thể hiện lưu đồ của WSOLA được thay đổi với sự điều khiển chất lượng, theo phương án của sáng chế này. Nên lưu ý rằng, sự định tỉ lệ thời gian 900 theo Fig.9 có thể được bổ sung bởi bất kỳ đặc điểm và chức năng nào được mô tả đối với

bộ định tỉ lệ thời gian 200 theo Fig.2 và ngược lại. Hơn nữa, nên lưu ý rằng, sự định tỉ lệ thời gian 900 theo Fig.9 có thể liên quan đến sự định tỉ lệ thời gian trên cơ sở mẫu 340 theo Fig.3 và bộ định tỉ lệ thời gian 450 theo Fig.4. Hơn nữa, sự định tỉ lệ thời gian 900 theo Fig.9 có thể diễn ra sự định tỉ lệ thời gian trên cơ sở mẫu 866.

Sự định tỉ lệ thời gian (hoặc bộ định tỉ lệ thời gian, hoặc bộ thay đổi bộ định tỉ lệ thời gian) 900 nhận các mẫu (âm thanh) được giải mã 910, ví dụ, dưới dạng điều biến mã xung (pulse-code-modulated - PCM). Các mẫu được giải mã 910 có thể tương ứng với các mẫu được giải mã 442, đến các mẫu âm thanh 332 hoặc đến tín hiệu âm thanh đầu vào 210. Hơn nữa, bộ định tỉ lệ thời gian 900 nhận thông tin điều khiển 912, mà có thể, ví dụ, tương ứng với thông tin định tỉ lệ trên cơ sở mẫu 444. Thông tin điều khiển 912 có thể, ví dụ, mô tả sự định tỉ lệ đích và/hoặc kích thước khung tối thiểu (ví dụ, số lượng tối thiểu các mẫu của khung của các mẫu âm thanh 448 để được cung cấp cho vùng đệm PCM 460). Bộ định tỉ lệ thời gian 900 bao gồm sự chuyển đổi (hoặc sự lựa chọn) 920, trong đó nó được quyết định, trên cơ sở thông tin về sự định tỉ lệ đích, liệu sự rút ngắn thời gian nên được thực hiện, liệu sự kéo dài thời gian nên được thực hiện hoặc liệu không có sự định tỉ lệ thời gian nên được thực hiện. Ví dụ, sự chuyển đổi (hoặc kiểm tra, hoặc sự lựa chọn) 920 có thể được dựa trên cơ sở thông tin định tỉ lệ trên cơ sở mẫu 444 nhận được từ logic điều khiển 490.

Nếu tìm ra, trên cơ sở thông tin định tỉ lệ đích, rằng không có sự định tỉ lệ nên được thực hiện, các mẫu được giải mã được nhận 910 được chuyển tiếp dưới dạng không thay đổi như đầu ra của bộ định tỉ lệ thời gian 900. Ví dụ, các mẫu được giải mã 910 được chuyển tiếp, dưới dạng không thay đổi, đến vùng đệm PCM 460 như các mẫu “được định tỉ lệ thời gian” 448.

Dưới đây, lưu đồ xử lý sẽ được mô tả cho trường hợp mà sự rút ngắn thời gian để được thực hiện (mà có thể được tìm thấy, bằng sự kiểm tra 920, trên cơ sở thông tin định tỉ lệ đích 912). Trong trường hợp sự rút ngắn thời gian được mong muốn, sự tính toán năng lượng 930 được thực hiện. Trong sự tính toán năng lượng 930 này, năng lượng của khối các mẫu (ví dụ, của khung bao gồm số lượng các mẫu đã cho) được tính toán. Tiếp theo sự tính toán năng lượng 930, sự lựa chọn (hoặc chuyển đổi, hoặc kiểm tra) 936 được thực hiện. Nếu tìm ra rằng giá trị năng lượng 932 được cung cấp

bởi sự tính toán năng lượng 930 là lớn hơn (hoặc bằng) giá trị ngưỡng năng lượng (ví dụ, giá trị ngưỡng năng lượng Y), đường xử lý thứ nhất 940 được chọn, mà bao gồm sự xác định thích ứng tín hiệu của lượng sự định tỉ lệ thời gian nằm trong sự định tỉ lệ thời gian trên cơ sở mẫu. Ngược lại, nếu tìm ra rằng giá trị năng lượng 932 được cung cấp bởi sự xử lý năng lượng 930 là nhỏ hơn (hoặc bằng) giá trị ngưỡng (ví dụ, giá trị ngưỡng Y), đường xử lý thứ hai 960 được chọn, trong đó lượng cố định sự dịch chuyển thời gian được áp dụng trong sự định tỉ lệ thời gian trên cơ sở mẫu. Trong đường xử lý 940, trong đó lượng dịch chuyển thời gian được xác định bằng cách thích ứng tín hiệu, sự ước lượng độ tương tự 942 được thực hiện trên cơ sở các mẫu âm thanh. Sự ước lượng độ tương tự 942 có thể xem xét thông tin kích thước khung tối thiểu 944 và có thể cung cấp thông tin 946 về độ tương tự cao nhất (hoặc về vị trí của độ tương tự cao nhất). Nói cách khác, sự ước lượng độ tương tự 942 có thể xác định vị trí nào (ví dụ, vị trí mẫu nào nằm trong khối các mẫu) là thích hợp nhất cho thuật toán chồng lấp và cộng rút ngắn thời gian. Thông tin 946 về độ tương tự cao nhất được chuyển tiếp đến bộ điều khiển chất lượng 950, mà tính toán hoặc ước lượng liệu thuật toán chồng lấp và cộng sử dụng thông tin 946 về độ tương tự cao nhất sẽ dẫn đến chất lượng âm thanh mà lớn hơn (hoặc bằng) giá trị ngưỡng chất lượng X (mà có thể không đổi hoặc có thể biến thiên). Nếu tìm ra rằng, bằng bộ điều khiển chất lượng 950, rằng chất lượng của thuật toán chồng lấp và cộng (hoặc tương đương, của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bởi thuật toán chồng lấp và cộng) sẽ nhỏ hơn (hoặc bằng) giá trị ngưỡng chất lượng X , sự định tỉ lệ thời gian được loại bỏ và các mẫu âm thanh không được định tỉ lệ được phát ra bởi bộ định tỉ lệ thời gian 900. Ngược lại, nếu tìm ra rằng, bằng bộ điều khiển chất lượng 950, rằng chất lượng của thuật toán chồng lấp và cộng sử dụng thông tin 946 về độ tương tự cao nhất (hoặc về vị trí của độ tương tự cao nhất) sẽ lớn hơn hoặc bằng với giá trị ngưỡng chất lượng X , thuật toán chồng lấp và cộng 954 được thực hiện, trong đó sự dịch chuyển, mà được áp dụng trong thuật toán chồng lấp và cộng, được mô tả bởi thông tin 946 về độ tương tự cao nhất (hoặc về vị trí của sự tương tự cao nhất). Theo đó, khối (hoặc khung) các mẫu âm thanh được định tỉ lệ được cung cấp bởi thuật toán chồng lấp và cộng.

Khối (hoặc khung) các mẫu âm thanh được định tỉ lệ thời gian 956 có thể, ví dụ, tương ứng với các mẫu được định tỉ lệ thời gian 448. Tương tự, khối (hoặc khung) của các mẫu âm thanh không được định tỉ lệ 952, mà được cung cấp nếu bộ điều khiển chất lượng 950 tìm ra rằng chất lượng có thể thu được sẽ nhỏ hơn hoặc bằng giá trị ngưỡng chất lượng X, có thể cũng tương ứng với các mẫu “được định tỉ lệ thời gian” 448 (trong đó không thực sự có sự định tỉ lệ thời gian trong trường hợp này).

Ngược lại, nếu tìm ra rằng trong sự lựa chọn 936 mà năng lượng của khối (hoặc khung) các mẫu âm thanh đầu vào 910 là nhỏ hơn (hoặc bằng) giá trị ngưỡng năng lượng Y, thuật toán chồng lấp và cộng 962 được thực hiện, trong đó sự dịch chuyển, mà được sử dụng trong thuật toán chồng lấp và cộng, được xác định bởi kích thước khung tối thiểu (được mô tả bởi thông tin kích thước khung tối thiểu), và trong đó khối (hoặc khung) của các mẫu âm thanh được định tỉ lệ 964 thu được, mà có thể tương ứng với các mẫu được định tỉ lệ thời gian 448.

Hơn nữa, nên được lưu ý rằng sự xử lý, mà được thực hiện trong trường hợp kéo dài thời gian, là tương tự với sự xử lý được thực hiện trong sự rút ngắn thời gian với độ tương tự được thay đổi và sự chồng lấp và cộng.

Tóm lại, nên lưu ý rằng ba trường hợp khác nhau được phân biệt trong sự định tỉ lệ thời gian trên cơ sở mẫu thích ứng tín hiệu khi sự rút ngắn thời gian hoặc sự kéo dài thời gian được lựa chọn. Nếu năng lượng của khối (hoặc khung) của các mẫu âm thanh đầu vào bao gồm năng lượng tương đối nhỏ (ví dụ, nhỏ hơn (hoặc bằng) giá trị ngưỡng năng lượng Y), thuật toán rút ngắn thời gian hoặc kéo dài thời gian được thực hiện với sự dịch chuyển thời gian cố định (tức là, với lượng cố định rút ngắn thời gian hoặc kéo dài thời gian). Ngược lại, nếu năng lượng của khối (hoặc khung) của các mẫu âm thanh đầu vào là lớn hơn (hoặc bằng với) giá trị ngưỡng năng lượng Y, lượng “tối ưu” (thỉnh thoảng cũng được chỉ ra trong đây như “ứng cử”) sự rút ngắn thời gian hoặc kéo dài thời gian được xác định bởi sự ước lượng độ tương tự (sự ước lượng độ tương tự 942). Trong bước điều khiển chất lượng tiếp theo, được xác định liệu chất lượng đủ sẽ thu được bởi thuật toán chồng lấp và cộng như vậy sử dụng lượng rút ngắn thời gian hoặc kéo dài thời gian “tối ưu” được xác định trước đó. Nếu tìm ra rằng chất lượng đủ có thể đạt đến, thuật toán chồng lấp và cộng được thực hiện sử dụng lượng rút ngắn

thời gian hoặc kéo dài thời gian "tối ưu" được xác định. Ngược lại, nếu tìm ra rằng chất lượng đủ có thể không được đạt đến sử dụng thuật toán chồng lấp và cộng sử dụng lượng rút ngắn thời gian và kéo dài thời gian "tối ưu" được định trước đó, sự rút ngắn thời gian và kéo dài thời gian được bỏ qua (hoặc hoãn lại đến thời điểm sau, ví dụ, đến khung sau).

Dưới đây, một số chi tiết nữa liên quan đến sự định tỉ lệ thời gian thích ứng chất lượng, mà có thể được thực hiện bởi bộ định tỉ lệ thời gian 900 (hoặc bộ định tỉ lệ thời gian 200), hoặc bởi bộ định tỉ lệ thời gian 340, hoặc bởi bộ định tỉ lệ thời gian 450), sẽ được mô tả. Các phương pháp định tỉ lệ thời gian sử dụng sự chồng lấp và cộng (overlap-and-add - OLA) là sẵn có rộng rãi, nhưng nhìn chung không thực hiện các kết quả định tỉ lệ thời gian thích ứng tín hiệu. Trong giải pháp này, mà có thể được sử dụng trong các bộ định tỉ lệ thời gian được mô tả ở đây, lượng định tỉ lệ thời gian không chỉ phụ thuộc vào vị trí được tách bởi sự ước lượng độ tương tự (ví dụ, bởi sự ước lượng độ tương tự 942), mà dường như tối ưu cho sự định tỉ lệ thời gian chất lượng cao, nhưng cũng trên chất lượng mong muốn của sự chồng lấp - cộng (ví dụ, của sự chồng lấp - cộng 954). Do đó, hai bước điều khiển chất lượng được đưa vào trong mô đun định tỉ lệ thời gian (ví dụ, trong bộ định tỉ lệ thời gian 900, hoặc trong các bộ định tỉ lệ thời gian khác được mô tả trong đây), để quyết định liệu sự định tỉ lệ thời gian sẽ dẫn đến các thành phần lạ có thể nghe được hay không. Trong trường hợp các thành phần lạ tiềm năng, sự định tỉ lệ thời gian được hoãn đến thời điểm mà nó ít có thể nghe được.

Bước điều khiển chất lượng thứ nhất tính toán số đo chất lượng mục tiêu sử dụng vị trí p được tách bởi số đo độ tương tự (ví dụ, bởi sự ước lượng độ tương tự 942) như đầu vào. trong trường hợp tín hiệu chu kì, p sẽ là tần số cơ bản của khung hiện thời. Sự tương quan chéo được tiêu chuẩn hóa $c()$ được tính toán cho các vị trí p , $2*p$, $3/2*p$, và $1/2*p$. $c(p)$ được cho là giá trị dương và $c(1/2*p)$ có thể là dương hoặc âm. Đối với các tín hiệu sóng hài, dấu của $c(2p)$ nên là dương và dấu của $c(3/2*p)$ nên bằng với dấu của $c(1/2*p)$. Hệ thức này có thể được sử dụng để tạo ra số đo chất lượng mục tiêu q :

$$q = c(p) * c(2*p) + c(3/2*p) * c(1/2*p).$$

Phạm vi các giá trị cho q là $[-2; +2]$. Tín hiệu sóng hài lý tưởng có thể dẫn đến $q = 2$, trong đó các tín hiệu rất động và băng rộng có thể tạo ra các thành phần lạ có thể nghe được trong suốt sự định tỉ lệ thời gian sẽ tạo ra giá trị thấp hơn. Do thực tế rằng sự định tỉ lệ thời gian được thực hiện trên cơ sở khung cạnh khung, toàn bộ tín hiệu để tính toán $c(2*p)$ và $c(3/2*p)$ có thể chưa có sẵn. Tuy nhiên, sự đánh giá có thể được thực hiện bằng việc xem các ví dụ trước. Do đó, $c(-p)$ có thể được sử dụng thay vì $c(2*p)$, và tương tự $c(-1/2*p)$ có thể được sử dụng thay vì $c(3/2*p)$.

Bước điều khiển chất lượng thứ hai so sánh giá trị hiện thời của số đo chất lượng mục tiêu q với giá trị chất lượng tối thiểu động q_{Min} (mà có thể tương ứng với giá trị ngưỡng chất lượng X) để xác định nếu sự định tỉ lệ thời gian nên được áp dụng cho khung hiện thời.

Có các dự định khác cho việc có giá trị chất lượng tối thiểu động: nếu q có giá trị thấp vì tín hiệu được đánh giá là kém để định tỉ lệ trong suốt chu kỳ dài, q_{Min} nên được giảm chậm để đảm bảo rằng sự định tỉ lệ mong muốn sẽ được thực hiện tại một số thời điểm với chất lượng được mong muốn thấp hơn. Mặt khác, các tín hiệu với giá trị cao cho q không nên dẫn đến sự định tỉ lệ nhiều khung theo hàng mà sẽ giảm chất lượng đối với các đặc điểm tín hiệu dài hạn (ví dụ, nhịp điệu).

Do đó, công thức sau được sử dụng để tính toán chất lượng tối thiểu động q_{Min} (mà có thể, ví dụ, tương đương với giá trị ngưỡng chất lượng X):

$$q_{Min} = q_{MinInitial} - (nNotScaled * 0,1) + (nScaled * 0,2)$$

$q_{MinInitial}$ là giá trị kết cấu để tối ưu hóa giữa chất lượng nhất định và độ trễ cho đến khi khung có thể được định tỉ lệ với chất lượng được yêu cầu, mà giá trị là 1 là sự thỏa hiệp tốt. $nNotScaled$ là bộ đếm các khung mà đã được định tỉ lệ vì chất lượng không đủ ($q < q_{Min}$). $nScaled$ đếm số lượng các khung mà đã được định tỉ lệ vì sự yêu cầu chất lượng đã đạt được ($q \geq q_{Min}$). Phạm vi của cả hai bộ đếm được giới hạn: chúng sẽ không được giảm về các giá trị âm và sẽ không tăng trên giá trị được chỉ ra mà được thiết lập mặc định là 4 (ví dụ).

Khung hiện thời sẽ được định tỉ lệ thời gian bởi vị trí p nếu $q \geq q_{Min}$, nếu không thì sự định tỉ lệ thời gian sẽ được hoãn đến khung tiếp theo tại đó điều kiện này

được đáp ứng. Mã giả của Fig.11 minh họa bộ điều khiển chất lượng cho sự định tỉ lệ thời gian.

Như có thể nhìn thấy, giá trị ban đầu cho $qMin$ được thiết lập về 1, trong đó giá trị ban đầu đã nêu được chỉ ra với “ $qMinInitial$ ” (được cho số tham chiếu là 1110). Tương tự, giá trị bộ đếm tối đa của $nScaled$ (được chỉ ra như “biến số $qualityRise$ ”) được khởi tạo là 4, như được thấy tại số tham chiếu 1112. Giá trị tối đa của bộ đếm $nNotScaled$ được khởi tạo là 4 (biến số “ $qualityRed$ ”), được cho số tham chiếu 1114. Tiếp theo, thông tin vị trí p được tách bởi số đo độ tương tự, như có thể được thấy tại số tham chiếu 1116. Tiếp theo, giá trị chất lượng q được tính toán cho vị trí được mô tả bởi giá trị vị trí p phù hợp với phương trình mà có thể được thấy tại số tham chiếu 1116. Giá trị ngưỡng chất lượng $qMin$ được tính toán phụ thuộc vào biến số $qMinInitial$, và cũng phụ thuộc vào các giá trị bộ đếm $nNotScaled$ và $nScaled$, như có thể được thấy tại số tham chiếu 1118. Như có thể được thấy, giá trị ban đầu $qMinInitial$ cho giá trị ngưỡng chất lượng $qMin$ được giảm bởi giá trị mà tỉ lệ với giá trị của bộ đếm $nNotScaled$, và tăng bởi giá trị mà tỉ lệ với giá trị $nScaled$. Như được nhìn thấy, các giá trị tối đa cho các giá trị bộ đếm $nNotScaled$ và $nScaled$ cũng xác định sự tăng tối đa của giá trị ngưỡng chất lượng $qMin$ và sự giảm tối đa của giá trị ngưỡng chất lượng $qMin$. Tiếp theo, sự kiểm tra được thực hiện liệu giá trị chất lượng q là lớn hơn hoặc bằng giá trị ngưỡng chất lượng $qMin$, có thể được thấy tại số tham chiếu 1120.

Nếu đây là trường hợp, thuật toán chồng lấp - cộng được thực hiện, như có thể thấy tại số tham chiếu 1122. Hơn nữa, biến số bộ đếm $nNotScaled$ được giảm, trong đó nó đảm bảo rằng biến số bộ đếm đã nêu không âm. Hơn nữa, biến số bộ đếm $nScaled$ tăng, trong đó nó được đảm bảo rằng $nScaled$ không vượt quá giới hạn trên được xác định bởi biến số (hoặc hằng số) $qualityRise$. Sự thích ứng của các biến số bộ đếm có thể được thấy tại các số tham chiếu 1124 và 1126.

Ngược lại, nếu tìm ra rằng sự so sánh được thể hiện tại số tham chiếu 1120 rằng giá trị chất lượng q là nhỏ hơn ngưỡng chất lượng $qMin$, sự thực hiện thuật toán chồng lấp và cộng được loại bỏ, biến số bộ đếm $nNotScaled$ tăng, xem xét rằng biến số bộ đếm $nNotScaled$ không vượt quá ngưỡng được xác định bởi biến số (hoặc hằng số)

qualityRed, và biến số bộ đếm nScaled được giảm, xem xét rằng biến số bộ đếm nScaled không trở thành âm. Sự thích ứng các biến số bộ đếm cho trường hợp mà chất lượng là không đủ được thể hiện tại các tham số 1128 và 1130.

Bộ định tỉ lệ thời gian theo Fig.10a và Fig.10b

Dưới đây, bộ định tỉ lệ thời gian thích ứng tín hiệu sẽ được giải thích viện dẫn đến Fig.10a và Fig.10b. Fig.10a và Fig.10b thể hiện lưu đồ của sự định tỉ lệ thời gian thích ứng tín hiệu. Nên lưu ý rằng sự định tỉ lệ thời gian thích ứng tín hiệu, như thể hiện trên Fig.10a và Fig.10b, ví dụ, được áp dụng trong bộ định tỉ lệ thời gian 200, trong bộ định tỉ lệ thời gian 340, trong bộ định tỉ lệ thời gian 450 hoặc trong bộ định tỉ lệ thời gian 900.

Bộ định tỉ lệ thời gian 1000 theo Fig.10a và Fig.10b, bao gồm sự tính toán năng lượng 1010, trong đó năng lượng khung (hoặc phân, hoặc khối) các mẫu âm thanh được tính toán. Ví dụ, sự tính toán năng lượng 1010 có thể tương ứng với sự tính toán năng lượng 930. Tiếp theo, sự kiểm tra 1014 được thực hiện, trong đó nó được kiểm tra liệu giá trị năng lượng thu được trong sự tính toán năng lượng 1010 là lớn hơn (hoặc bằng) giá trị ngưỡng năng lượng (mà có thể, ví dụ, là giá trị ngưỡng năng lượng cố định). Tìm ra rằng, trong sự kiểm tra 1014, giá trị năng lượng thu được trong sự tính toán 1010 là nhỏ hơn (hoặc bằng) giá trị ngưỡng năng lượng, có thể giả thiết rằng chất lượng đủ có thể thu được bởi thuật toán chồng lấp - cộng, và thuật toán chồng lấp - cộng được thực hiện với sự dịch chuyển thời gian tối đa (để nhờ đó thu được sự định tỉ lệ thời gian tối đa) trong bước 1018. Ngược lại, nếu tìm ra rằng sự kiểm tra 1014 mà giá trị năng lượng thu được trong sự tính toán 1010 là không nhỏ hơn (hoặc bằng với giá trị ngưỡng năng lượng, sự tìm kiếm cho sự kết hợp tốt nhất của phân đoạn mẫu nằm trong vùng tìm kiếm được thực hiện sử dụng số đo độ tương tự. Ví dụ, số đo độ tương tự có thể là sự tương quan chéo, sự tương quan chéo được tiêu chuẩn hóa hoặc hàm chênh lệch độ lớn trung bình hoặc tổng các sai số được bình phương. Dưới đây, một số chi tiết về sự tìm kiếm này cho sự kết hợp tốt nhất sẽ được mô tả, và nó cũng sẽ được giải thích làm thế nào sự kéo dài thời gian hoặc sự rút ngắn thời gian có thể thu được.

Bây giờ sự viện dẫn được thực hiện đến sự biểu diễn biểu đồ tại số tham chiếu 1040. Sự biểu diễn thứ nhất 1042 thể hiện khối (hoặc khung) của các mẫu mà bắt đầu tại thời gian t_1 và kết thúc tại thời gian t_2 . Như được thấy, khối các mẫu mà bắt đầu tại thời gian t_1 và kết thúc tại thời gian t_2 có thể được chia ra một cách logic thành khối các mẫu thứ nhất, mà bắt đầu tại thời gian t_1 và kết thúc tại thời gian t_3 và khối các mẫu thứ hai mà bắt đầu tại thời gian t_4 và kết thúc tại thời gian t_2 . Tuy nhiên, khối các mẫu thứ hai sau đó được dịch chuyển thời gian đối với khối các mẫu thứ nhất, mà có thể được thấy tại số tham chiếu 1044. Ví dụ, kết quả của sự dịch chuyển thời gian thứ nhất, khối các mẫu thứ hai được dịch chuyển thời gian bắt đầu tại thời gian t_4' và kết thúc tại thời gian t_2' . Theo đó, sự chồng lấp theo thời gian giữa khối các mẫu thứ nhất và khối các mẫu thứ hai được dịch chuyển thời gian giữa các thời gian t_4' và t_3 . Tuy nhiên, có thể thấy được, không có sự kết hợp tốt (tức là, không có độ tương tự cao) giữa khối các mẫu thứ nhất và phiên bản được dịch chuyển thời gian của khối các mẫu thứ hai, ví dụ, trong vùng chồng lấp giữa các thời gian t_4' và t_3 (hoặc nằm trong vùng quá khu vực chồng lấp đã nêu giữa các thời gian t_4' và t_3). Nói cách khác, bộ định tỉ lệ thời gian có thể, ví dụ, dịch chuyển thời gian khối các mẫu thứ hai, như thể hiện tại số tham chiếu 1044, và xác định số đo độ tương tự cho vùng chồng lấp (hoặc cho phần vùng chồng lấp) giữa các thời gian t_4' và t_3 . Hơn nữa, bộ định tỉ lệ thời gian có thể cũng được áp dụng sự dịch chuyển thời gian bổ sung cho khối các mẫu thứ hai, như thể hiện tại số tham chiếu 1046, sao cho phiên bản được định tỉ lệ thời gian (hai lần) của khối các mẫu thứ hai bắt đầu tại thời gian t_4'' và kết thúc tại thời gian t_2'' (với $t_2'' > t_2' > t_2$ và tương tự $t_4'' > t_4' > t_4$). Bộ định tỉ lệ thời gian có thể cũng xác định thông tin độ tương tự (định lượng) biểu thị độ tương tự giữa khối các mẫu thứ nhất và phiên bản được dịch chuyển hai lần của khối các mẫu thứ hai, ví dụ, giữa các thời gian t_4'' và t_3 (hoặc, ví dụ, nằm trong phần giữa t_4'' và t_3). Theo đó, bộ định tỉ lệ thời gian ước lượng sự dịch chuyển thời gian nào của phiên bản được dịch chuyển thời gian của khối các mẫu thứ hai độ tương tự, trong vùng chồng lấp với khối các mẫu thứ nhất, được tối đa hóa (hoặc cuối cùng lớn hơn giá trị ngưỡng). Theo đó, sự dịch chuyển thời gian có thể được xác định kết quả nào trong "sự kết hợp tốt nhất" trong đó độ tương tự giữa khối các mẫu thứ nhất và phiên bản được dịch chuyển thời gian của khối các mẫu thứ hai được tối đa hóa (hoặc ít nhất đủ lớn). Theo đó, nếu có độ tương tự đủ giữa khối

các mẫu thứ nhất và phiên bản được dịch chuyển hai lần của khối các mẫu thứ hai nằm trong vùng chồng lấp theo thời gian (ví dụ giữa các thời gian t_4'' và t_3), nó có thể được mong đợi, với độ tin cậy được xác định bởi số đo đã được sử dụng, rằng thuật toán chồng lấp và cộng chồng lấp và cộng khối các mẫu thứ nhất và phiên bản được dịch chuyển thời gian hai lần của khối các mẫu thứ hai dẫn đến tín hiệu âm thanh mà không có thành phần lạ có thể nghe được đáng kể. Hơn nữa, nên lưu ý rằng sự chồng lấp và cộng giữa khối các mẫu thứ nhất và phiên bản được dịch chuyển thời gian hai lần của khối các mẫu thứ hai dẫn đến phần tín hiệu âm thanh mà sự mở rộng tạm thời giữa các thời gian t_1 và t_2'' , mà dài hơn tín hiệu âm thanh "gốc", được mở rộng từ thời gian t_1 đến thời gian t_2 . Theo đó, sự kéo dài thời gian có thể đạt được bằng cách chồng lấp và cộng khối các mẫu thứ nhất và phiên bản được dịch chuyển thời gian hai lần của khối các mẫu thứ hai.

Tương tự, sự rút gọn thời gian có thể đạt được, như được giải thích viện dẫn đến sự biểu diễn đồ thị tại số tham chiếu 1050. Có thể nhìn thấy tại số tham chiếu 1052, khối (hoặc khung) các mẫu ban đầu mà kéo dài giữa các thời gian t_{11} và t_{12} . Khối (hoặc khung) các mẫu ban đầu có thể được chia ra, ví dụ, thành khối các mẫu thứ nhất mà mở rộng từ thời gian t_{11} đến thời gian t_{13} và khối các mẫu thứ hai mà mở rộng từ thời gian t_{13} đến thời gian t_{12} . Khối các mẫu thứ hai được dịch chuyển thời gian về bên trái, mà có thể được thấy tại số tham chiếu 1054. Do đó, phiên bản được dịch chuyển thời gian của khối các mẫu thứ hai bắt đầu tại thời gian t_{13}' và kết thúc tại thời gian t_{12}' . Tương tự, có sự chồng lấp theo thời gian giữa khối các mẫu thứ nhất và phiên bản được dịch chuyển thời gian lần một của khối các mẫu thứ hai giữa các thời gian t_{13}' và t_{13} . Tuy nhiên, bộ định tỉ lệ thời gian có thể xác định thông tin độ tương tự (định lượng) biểu diễn độ tương tự của khối các mẫu thứ nhất và của phiên bản được dịch chuyển thời gian (một lần) của khối các mẫu thứ hai giữa các thời gian t_{13}' và t_{13} (hoặc cho phần của thời gian giữa các thời gian t_{13}' và t_{13}) và tìm ra độ tương tự là không đặc biệt tốt. Hơn nữa, bộ định tỉ lệ thời gian có thể còn dịch chuyển thời gian khối các mẫu thứ hai, để nhờ đó thu được phiên bản được dịch chuyển thời gian hai lần của khối các mẫu thứ hai, mà được thể hiện tại số tham chiếu 1056, và bắt đầu tại thời gian t_{13}'' và kết thúc tại thời gian t_{12}'' . Tương tự, có sự chồng lấp theo thời gian giữa khối các mẫu thứ nhất và phiên bản được dịch chuyển thời gian lần một của

khối các mẫu thứ hai giữa các thời gian $t13'$ và $t13$. Nó có thể được tìm ra, bởi bộ định tỉ lệ thời gian, rằng thông tin độ tương tự (định lượng) biểu thị độ tương tự cao giữa khối các mẫu thứ nhất và phiên bản được định tỉ lệ thời gian hai lần của khối các mẫu thứ hai giữa các thời gian $t13''$ và $t13$. Theo đó, nó có thể được bao gồm, bởi bộ định tỉ lệ thời gian, mà thuật toán chồng lấp và cộng có thể được thực hiện với chất lượng tốt và ít các thành phần lạ có thể nghe được giữa khối các mẫu thứ nhất và phiên bản được dịch chuyển thời gian hai lần của khối các mẫu thứ hai (ít nhất với độ tin cậy được cung cấp bởi số đo độ tương tự được sử dụng). Hơn nữa, phiên bản được dịch chuyển thời gian ba lần của khối các mẫu thứ hai, mà được thể hiện tại số tham chiếu 1058 mà có thể cũng được xem xét. Phiên bản được dịch chuyển thời gian ba lần của khối các mẫu thứ hai có thể bắt đầu tại thời gian $t13'''$ và kết thúc tại thời gian $t12'''$. Tuy nhiên, phiên bản được dịch chuyển thời gian ba lần của khối các mẫu thứ hai có thể không bao gồm độ tương tự tốt với khối các mẫu thứ nhất trong vùng chồng lấp giữa các thời gian $t13'''$ và $t13$, vì sự dịch chuyển thời gian không thích hợp. Thông thường, bộ định tỉ lệ thời gian có thể tìm ra rằng phiên bản được dịch chuyển thời gian hai lần của khối các mẫu thứ hai bao gồm sự kết hợp tốt nhất (độ tương tự tốt nhất trong vùng chồng lấp, và/hoặc trong môi trường của vùng chồng lấp, và/hoặc trong phần của vùng chồng lấp) với khối các mẫu thứ nhất. Theo đó, bộ định tỉ lệ thời gian có thể thực hiện sự chồng lấp và cộng của khối các mẫu thứ nhất và phiên bản được dịch chuyển thời gian hai lần của khối các mẫu thứ hai, được cung cấp sự kiểm tra chất lượng bổ sung (mà có thể dựa vào số đo độ tương tự thứ hai có ý nghĩa hơn) biểu thị chất lượng đủ. Kết quả của thuật toán chồng lấp và cộng, khối các mẫu được tổ hợp thu được, mà mở rộng từ thời gian $t11$ đến thời gian $t12''$, và ngắn hơn theo thời gian hơn khối các mẫu ban đầu từ thời gian $t11$ đến thời gian $t12$. Theo đó, sự rút ngắn thời gian có thể được thực hiện.

Nên lưu ý rằng các chức năng nêu trên, mà đã được mô tả viện dẫn đến các sự biểu diễn đồ thị tại các số tham chiếu 1040 và 1050, có thể được thực hiện bởi sự tìm kiếm 1030, trong đó được thực hiện bởi sự tìm kiếm 1030, trong đó thông tin về vị trí của độ tương tự cao nhất được cung cấp như kết quả của sự tìm kiếm cho sự kết hợp tốt nhất (trong đó thông tin hoặc giá trị mô tả vị trí của độ tương tự cao nhất cũng được chỉ ra với p ở đây). Độ tương tự giữa khối các mẫu thứ nhất và phiên bản được dịch

chuyển thời gian của khối các mẫu thứ hai nằm trong các vùng chồng lấp tương ứng có thể được xác định sử dụng sự tương quan chéo, sử dụng sự tương quan chéo được tiêu chuẩn hóa, sử dụng hàm chênh lệch biên độ trung bình hoặc sử dụng tổng của các sai số được bình phương.

Một khi thông tin về vị trí độ tương tự cao nhất (p) được xác định, sự tính toán 1060 của chất lượng kết hợp cho vị trí được xác định (p) của độ tương tự cao nhất được thực hiện. Sự tính toán này có thể được thực hiện, ví dụ, như được thể hiện tại số tham chiếu 1116 trên Fig.11. Nói cách khác, thông tin (định lượng) về chất lượng kết hợp (mà có thể, ví dụ, được chỉ ra với q) có thể được tính toán sử dụng sự tổ hợp của bốn giá trị tương quan, mà có thể thu được cho các sự dịch chuyển thời gian khác nhau (ví dụ, các sự dịch chuyển thời gian p , $2*p$, $3/2*p$ và $1/2*p$). Theo đó, thông tin (định lượng) (q) biểu diễn chất lượng kết hợp có thể thu được.

Bây giờ viện dẫn đến Fig.10b, sự kiểm tra 1064 được thực hiện, trong đó thông tin định lượng q mô tả chất lượng kết hợp được so sánh với giá trị ngưỡng chất lượng q_{Min} . Sự kiểm tra hoặc so sánh 1064 này có thể đánh giá liệu chất lượng kết hợp, được biểu diễn bởi biến số q , là lớn hơn (hoặc bằng) giá trị ngưỡng biến thiên q_{Min} . Nếu tìm ra rằng trong sự kiểm tra 1064 rằng chất lượng kết hợp là đủ (tức là, lớn hơn hoặc bằng giá trị ngưỡng chất lượng biến thiên), thuật toán chồng lấp - cộng được áp dụng (bước 1068) sử dụng vị trí của độ tương tự cao nhất (mà được mô tả, ví dụ, bởi biến số p). Theo đó, thuật toán chồng lấp và cộng được thực hiện, ví dụ, giữa khối các mẫu thứ nhất và phiên bản được dịch chuyển thời gian của khối các mẫu thứ hai mà dẫn đến kết quả là "sự kết hợp tốt nhất" (tức là, trong giá trị cao nhất của thông tin độ tương tự). Cụ thể hơn, sự tham chiếu được thực hiện, ví dụ, cho sự giải thích được thực hiện đối với sự biểu diễn đồ thị 1040 và 1050. Sự áp dụng sự chồng lấp và cộng cũng được thể hiện tại số tham chiếu 1122 trên Fig.11. Hơn nữa, sự cập nhật bộ đếm khung được thực hiện trong bước 1072. Ví dụ, biến số bộ đếm "nNotScaled" và biến số bộ đếm "nScaled", được cập nhật, ví dụ, như được mô tả với sự viện dẫn đến Fig.11 tại các số tham chiếu 1124 và 1126. Ngược lại, nếu tìm ra trong sự kiểm tra 1064 rằng chất lượng kết hợp là không đủ (ví dụ, nhỏ hơn (hoặc bằng) giá trị ngưỡng chất lượng biến thiên q_{min}), thuật toán chồng lấp và cộng được tránh (ví dụ, được hoãn lại), mà được biểu thị tại số tham chiếu 1076. Trong trường hợp này, các bộ đếm khung cũng

được cập nhật, như được thể hiện trong bước 1080. Sự cập nhật các bộ đếm khung có thể được thực hiện, ví dụ, như được thể hiện tại các số tham chiếu 1128 và 1130 trên Fig.11. Hơn nữa, bộ định tỉ lệ thời gian được mô tả với sự viện dẫn đến Fig.10a và Fig.10b có thể cũng tính toán giá trị ngưỡng chất lượng biến thiên q_{Min} , mà được thể hiện tại số tham chiếu 1084. Sự tính toán của giá trị ngưỡng chất lượng biến thiên q_{Min} có thể được thực hiện, ví dụ, như được thể hiện tại số tham chiếu 1118 trên Fig.11.

Tóm lại, bộ định tỉ lệ thời gian 1000, chức năng của nó đã được mô tả viện dẫn đến Fig.10a và Fig.10b dưới dạng lưu đồ, có thể thực hiện sự định tỉ lệ thời gian trên cơ sở mẫu sử dụng cơ cấu điều khiển chất lượng (các bước từ 1060 đến 1084).

Phương pháp theo Fig.14

Fig.14 thể hiện lưu đồ của phương pháp để điều khiển sự cung cấp nội dung âm thanh được giải mã trên cơ sở nội dung âm thanh đầu vào. Phương pháp 1400 theo Fig.14 bao gồm lựa chọn 1410 sự định tỉ lệ thời gian trên cơ sở khung hoặc sự định tỉ lệ thời gian trên cơ sở mẫu theo cách thích ứng tín hiệu.

Ngoài ra, nên lưu ý rằng phương pháp 1400 có thể được bổ sung bởi bất kỳ đặc điểm và chức năng được mô tả trong đây, ví dụ, đối với bộ điều khiển vùng đệm chấp chờn.

Phương pháp theo Fig.15

Fig.15 thể hiện sơ đồ khối dạng giản lược của phương pháp 1500 để cung cấp phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào. Phương pháp bao gồm tính toán hoặc ước lượng 1510 chất lượng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bởi sự định tỉ lệ thời gian của tín hiệu âm thanh đầu vào. Hơn nữa, phương pháp 1500 bao gồm thực hiện 1520 sự định tỉ lệ thời gian của tín hiệu âm thanh đầu vào phụ thuộc vào sự tính toán hoặc ước lượng của chất lượng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bằng sự định tỉ lệ thời gian.

Phương pháp 1500 có thể được bổ sung bởi bất kỳ đặc điểm và chức năng được mô tả trong đây, ví dụ, đối với bộ định tỉ lệ thời gian.

Kết luận

Tóm lại, các phương án theo sáng chế tạo ra phương pháp quản lý vùng đệm chập chờn và thiết bị cho sự truyền thông tiếng nói và âm thanh chất lượng cao. Phương pháp và thiết bị có thể được sử dụng cùng nhau với các bộ mã hóa-giải mã truyền thông, như MPEG ELD, AMR-WB, hoặc các bộ mã hóa-giải mã tương lai. Nói cách khác, các phương án theo sáng chế tạo ra phương pháp và thiết bị cho sự bù của sự chập chờn giữa các lần liên tiếp trong sự truyền thông trên cơ sở gói.

Các phương án của sáng chế có thể được áp dụng, ví dụ, trong công nghệ được gọi là “3GPP EVS”.

Dưới đây, một số khía cạnh của phương án theo sáng chế sẽ được mô tả một cách vắn tắt.

Giải pháp sự quản lý vùng đệm chập chờn được mô tả trong đây tạo ra hệ thống, trong đó lượng môđun được mô tả có sẵn và được tổ hợp theo cách được mô tả như trên. Hơn nữa, nên lưu ý rằng các khía cạnh của sáng chế cũng liên quan đến các đặc điểm của bản thân các môđun.

Khía cạnh quan trọng của sáng chế là sự lựa chọn thích ứng tín hiệu của phương pháp định tỉ lệ thời gian cho sự quản lý vùng đệm chập chờn thích ứng. Giải pháp được mô tả kết hợp sự định tỉ lệ trên cơ sở khung và sự định tỉ lệ trên cơ sở mẫu trong logic điều khiển để lợi ích của cả hai phương pháp được tổ hợp. Các phương pháp định tỉ lệ thời gian có sẵn là:

- Sự chèn/xóa nhiễu âm nền nhân tạo trong DTX
- Sự chồng lấp và cộng (Overlap-and-add - OLA) mà không có sự tương quan trong năng lượng thấp (ví dụ, đối với các khung có năng lượng tín hiệu thấp);
- WSOLA cho các tín hiệu hoạt động;
- Sự chèn khung được che giấu cho sự kéo dài trong trường hợp vùng đệm chập chờn trống.

Giải pháp được mô tả trong đây mô tả cơ cấu để tổ hợp các phương pháp theo khung (sự chèn và xóa nhiễu âm nền nhân tạo, và sự chèn các khung được che giấu cho sự kéo dài) với các phương pháp trên cơ sở mẫu (WSOLA đối với các tín hiệu

hoạt động, và sự chồng lấp - cộng (overlap-add - OLA) cho các tín hiệu năng lượng thấp). Trên Fig.8, logic điều khiển được minh họa mà lựa chọn công nghệ tối ưu cho sự thay đổi sự định tỉ lệ thời gian theo phương án của sáng chế.

Theo khía cạnh nữa được mô tả ở đây, nhiều đích cho sự quản lý vùng đệm chấp chờn thích ứng được sử dụng. Trong giải pháp được mô tả, sự ước lượng độ trễ đích sử dụng các tiêu chuẩn tối ưu khác nhau để tính toán độ trễ sự phát lại đích đơn. Các khía cạnh này dẫn đến đích khác nhau lúc đầu, được tối ưu hóa cho chất lượng cao hoặc độ trễ thấp.

Nhiều đích cho sự tính toán độ trễ phát ra đích là:

- Chất lượng: tránh sự mất muộn (đánh giá sự chấp chờn);
- Độ trễ: độ trễ giới hạn (đánh giá sự chấp chờn).

Khía cạnh (tối ưu) của giải pháp được mô tả để tối ưu hóa sự ước lượng độ trễ đích để độ trễ được giới hạn nhưng các sự mất muộn cũng được tránh và hơn nữa sự dự trữ nhỏ trong vùng đệm chấp chờn được giữ xác suất của phép nội suy để cho phép sự che giấu lỗi chất lượng cao cho bộ giải mã.

Khía cạnh (tùy chọn) khác liên quan đến sự khôi phục sự che giấu TCX với các khung muộn. Các khung mà đến muộn được loại bỏ bởi hầu hết giải pháp quản lý vùng đệm chấp chờn hiện nay. Các cơ cấu đã được mô tả để sử dụng các khung muộn trong bộ giải mã trên cơ sở ACELP [Lef03]. Theo khía cạnh, cơ cấu như vậy cũng được sử dụng cho các khung hơn là các khung ACELP, ví dụ, các khung được mã hóa miền tần số như TCX, nhìn chung để giúp đỡ việc khôi phục của trạng thái giải mã. Do đó, các khung mà được nhận muộn và đã được che giấu vẫn được đưa vào bộ giải mã để cải thiện sự khôi phục trạng thái bộ giải mã.

Khía cạnh quan trọng khác theo sáng chế là sự định tỉ lệ thích ứng tín hiệu, mà được mô tả như trên.

Để kết luận, các phương án theo sáng chế tạo ra giải pháp quản lý vùng đệm chấp chờn hoàn toàn mà có thể được sử dụng cho kinh nghiệm người dùng được cải thiện trong sự truyền thông trên cơ sở gói. Đó là sự quan sát rằng các giải pháp được trình

bày thực hiện tốt hơn bất kỳ giải pháp quản lý vùng đệm chấp chừa đã biết đối với các tác giả.

Các thay thế việc thực thi

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. 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.

Tín hiệu âm thanh đã được mã hoá có thể được lưu trữ trên vật ghi lưu trữ số hoặc có thể được truyền trên môi trường truyền như môi trường truyền dẫn không dây hoặc môi trường truyền dẫn có dây như Liên mạng.

Phụ thuộc vào các yêu cầu thực hiện đã biết, 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 lưu trữ số, ví dụ, đĩa mềm, DVD, Blu-Ray, CD, ROM, PROM, EPROM, EEPROM hoặc bộ nhớ FLASH, có các tín hiệu điều khiển có thể đọc được bằng điện được lưu trữ trên đó, mà kết hợp (hoặc có thể kết hợp) với hệ thống máy tính khả trình sao cho 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 bằng máy.

Các phương án khác bao gồm chương trình máy tính để thực hiện một trong số 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.

Phương án khác của các phương pháp theo sáng chế là, vật mang dữ liệu (hoặc vật ghi lưu trữ số, hoặc vật ghi đọc được bằng máy tính) gồm có, 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 đã được ghi thường là hữu hình và/hoặc không chuyển tiếp.

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

Phương án khác gồm có phương tiện xử lý, ví dụ, máy tính, hoặc thiết bị logic lập trình được tạo cấu hình để hoặc được làm thích ứng 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. Ví dụ, bộ thu có thể 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.

Theo một số phương án, thiết bị logic lập trình được (ví dụ, mảng công 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 công 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 tốt hơn là được thực hiện bởi thiết bị phần cứng bất kỳ.

Thiết bị được mô tả ở đây có thể được thực hiện sử dụng thiết bị phần cứng, hoặc sử dụng máy tính, hoặc sử dụng tổ hợp của thiết bị phần cứng và máy tính.

Các phương pháp được mô tả ở đây có thể được thực hiện sử dụng thiết bị phần cứng, hoặc sử dụng máy tính, hoặc sử dụng tổ hợp của thiết bị phần cứng và máy tính.

Các phương án được mô tả ở 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 hiểu biết 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.

Các phương án khác

Phương án thứ nhất đề cập đến bộ định tỉ lệ thời gian (200; 340; 450; 866; 900; 1000) để cung cấp phiên bản được định tỉ lệ thời gian (212;312;448;956) của tín hiệu âm thanh đầu vào (210; 332;442;910),

trong đó bộ định tỉ lệ thời gian được tạo cấu hình để tính toán hoặc ước lượng (950; 1060) chất lượng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bởi sự định tỉ lệ thời gian của tín hiệu âm thanh đầu vào, và

trong đó bộ định tỉ lệ thời gian được tạo cấu hình để thực hiện (954; 1068) sự định tỉ lệ thời gian của tín hiệu âm thanh đầu vào phụ thuộc vào sự tính toán hoặc sự ước lượng chất lượng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bởi sự định tỉ lệ thời gian.

Phương án thứ hai đề cập đến bộ định tỉ lệ thời gian (200; 340; 450; 866; 900; 1000) theo phương án 1, trong đó bộ định tỉ lệ thời gian được tạo cấu hình để thực hiện thao tác chồng lấp và cộng (954;1068) sử dụng khối các mẫu thứ nhất của tín hiệu âm thanh đầu vào và khối các mẫu thứ hai của tín hiệu âm thanh đầu vào.

trong đó bộ định tỉ lệ thời gian được tạo cấu hình để dịch chuyển thời gian khỏi các mẫu thứ hai so với khối các mẫu thứ nhất, và để chồng lấp và cộng khối các mẫu thứ nhất và khối các mẫu thứ hai được dịch chuyển thời gian, để nhờ đó thu được phiên bản định tỉ lệ thời gian của tín hiệu âm thanh đầu vào.

Phương án thứ ba đề cập đến bộ định tỉ lệ thời gian (200; 340; 450; 866; 900; 1000) theo phương án 2, trong đó bộ định tỉ lệ thời gian được tạo cấu hình để tính toán hoặc ước lượng (950;1060) chất lượng của thao tác chồng lấp và cộng giữa khối các mẫu thứ nhất và khối các mẫu thứ hai được dịch chuyển thời gian, để tính toán hoặc ước lượng chất lượng của phiên bản của tín hiệu âm thanh đầu vào có thể thu được bằng sự định tỉ lệ thời gian.

Phương án thứ tư đề cập đến bộ định tỉ lệ thời gian (200; 340; 450; 866; 900; 1000) theo phương án 2 hoặc phương án 3, trong đó bộ định tỉ lệ thời gian được tạo cấu hình để xác định (942; 1030) sự dịch chuyển thời gian (p) của khối các mẫu thứ hai so với khối các mẫu thứ nhất phụ thuộc vào sự xác định mức tương tự, được đánh giá sử dụng số đo độ tương tự thứ nhất, giữa khối các mẫu thứ nhất, hoặc phần của khối các mẫu thứ nhất, và khối các mẫu thứ hai, hoặc phần của khối các mẫu thứ hai.

Phương án thứ năm đề cập đến bộ định tỉ lệ thời gian (200; 340; 450; 866; 900; 1000) theo phương án 4, trong đó bộ định tỉ lệ thời gian được tạo cấu hình để xác định thông tin về mức tương tự giữa khối các mẫu thứ nhất, hoặc phần của khối các mẫu thứ nhất, và khối các mẫu thứ hai, hoặc phần của khối các mẫu thứ hai, so với nhiều sự dịch chuyển thời gian khác nhau giữa khối các mẫu thứ nhất và khối các mẫu thứ hai, và để xác định sự dịch chuyển thời gian (ứng cử), để được sử dụng cho thao tác chồng lấp và cộng trên cơ sở thông tin về mức tương tự đối với nhiều sự dịch chuyển thời gian khác nhau.

Phương án thứ sáu đề cập đến bộ định tỉ lệ thời gian (200; 340; 450; 866; 900; 1000) theo phương án 4 hoặc phương án 5, trong đó bộ định tỉ lệ thời gian được tạo cấu hình để xác định sự dịch chuyển thời gian (p) của khối các mẫu thứ hai so với khối các mẫu thứ nhất, mà sự dịch chuyển thời gian được sử dụng cho thao tác chồng lấp và cộng, phụ thuộc vào thông tin dịch chuyển thời gian đích.

Phương án thứ bảy đề cập đến bộ định tỉ lệ thời gian (200; 340; 450; 866; 900; 1000) theo một trong số các phương án từ 4 đến 6, trong đó bộ định tỉ lệ thời gian được tạo cấu hình để tính toán hoặc ước lượng (950;1060) chất lượng (q) của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bởi sự định tỉ lệ thời gian của tín hiệu âm thanh đầu vào trên cơ sở thông tin về mức tương tự, giữa khối các mẫu thứ nhất, hoặc phần của khối các mẫu thứ nhất, và khối các mẫu thứ hai, được dịch chuyển thời gian bởi sự dịch chuyển thời gian được xác định (p), hoặc phần của khối các mẫu thứ hai, được dịch chuyển thời gian bởi sự dịch chuyển thời gian được xác định (p).

Phương án thứ tám đề cập đến bộ định tỉ lệ thời gian (200; 340; 450; 866; 900; 1000) theo phương án 7, trong đó bộ định tỉ lệ thời gian được tạo cấu hình để quyết định (1064) trên cơ sở thông tin về mức tương tự, giữa khối các mẫu thứ nhất, hoặc phần của khối các mẫu thứ nhất, và khối các mẫu thứ hai, được dịch chuyển thời gian bởi sự dịch chuyển thời gian được xác định (p), hoặc phần của khối các mẫu thứ hai, được dịch chuyển thời gian bởi sự dịch chuyển thời gian được xác định (p), liệu sự định tỉ lệ thời gian thực sự được thực hiện.

Phương án thứ chín đề cập đến bộ định tỉ lệ thời gian (200; 340; 450; 866; 900; 1000) theo một trong số các phương án từ 1 đến 8,

trong đó bộ định tỉ lệ thời gian được tạo cấu hình để dịch chuyển thời gian khối các mẫu thứ hai so với khối các mẫu thứ nhất, và để chồng lấp và cộng (954;1068) khối các mẫu thứ nhất và khối các mẫu thứ hai được dịch chuyển thời gian, để nhờ đó thu được phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào, nếu sự tính toán hoặc ước lượng chất lượng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bằng sự định tỉ lệ thời gian biểu thị chất lượng mà lớn hơn hoặc bằng giá trị ngưỡng chất lượng (q_{min}); và

trong đó bộ định tỉ lệ thời gian được tạo cấu hình để xác định sự dịch chuyển thời gian (p) của khối các mẫu thứ hai so với khối các mẫu thứ nhất phụ thuộc vào sự xác định mức tương tự, được đánh giá sử dụng số đo độ tương tự thứ nhất, giữa khối các mẫu thứ nhất, hoặc phần của khối các mẫu thứ nhất, và khối các mẫu thứ hai, hoặc phần của khối các mẫu thứ hai; và

trong đó bộ định tỉ lệ thời gian được tạo cấu hình để tính toán hoặc ước lượng (950;1060) chất lượng (q) của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bởi sự định tỉ lệ thời gian của tín hiệu âm thanh đầu vào trên cơ sở thông tin về mức tương tự, được đánh giá sử dụng số đo độ tương tự thứ hai, giữa khối các mẫu thứ nhất, hoặc phần của khối các mẫu thứ nhất, và khối các mẫu thứ hai, được dịch chuyển thời gian bởi sự dịch chuyển thời gian được xác định, hoặc phần của khối các mẫu thứ hai, được dịch chuyển thời gian bởi sự dịch chuyển thời gian được xác định.

Phương án thứ mười đề cập đến bộ định tỉ lệ thời gian (200; 340; 450; 866; 900; 1000) theo phương án 9, trong đó số đo độ tương tự thứ hai (q) phức tạp hơn về tính toán hơn là số đo độ tương tự thứ nhất.

Phương án thứ mười một đề cập đến bộ định tỉ lệ thời gian (200; 340; 450; 866; 900; 1000) theo phương án 9 hoặc phương án 10, trong đó số đo độ tương tự thứ nhất là sự tương quan chéo hoặc sự tương quan chéo được tiêu chuẩn hóa, hoặc hàm chênh lệch biên độ trung bình hoặc tổng các sai số được bình phương, và

trong đó số đo độ tương tự thứ hai (q) là tổ hợp các sự tương quan chéo hoặc của các sự tương quan được tiêu chuẩn hóa cho nhiều sự dịch chuyển thời gian khác nhau.

Phương án thứ mười hai đề cập đến bộ định tỉ lệ thời gian (200; 340; 450; 866; 900; 1000) theo một trong số các phương án từ 9 đến 11, trong đó số đo độ tương tự thứ hai (q) là tổ hợp của các sự tương quan chéo cho ít nhất bốn sự dịch chuyển thời gian khác nhau.

Phương án thứ mười ba đề cập đến bộ định tỉ lệ thời gian theo phương án 12, trong đó số đo độ tương tự thứ hai (q) là tổ hợp của giá trị tương quan chéo thứ nhất và giá trị tương quan chéo thứ hai, mà thu được cho sự dịch chuyển thời gian mà được tạo khoảng cách bằng bội số nguyên của thời khoảng chu kỳ (p) của tần số cơ bản của nội dung âm thanh của khối các mẫu thứ nhất hoặc khối các mẫu thứ hai, và của giá trị tương quan chéo thứ ba và giá trị tương quan chéo thứ tư, mà thu được cho các sự dịch chuyển thời gian mà được chèn khoảng trống bởi bội số nguyên của thời khoảng chu kỳ (p) của tần số cơ bản của nội dung âm thanh,

trong đó sự dịch chuyển thời gian mà giá trị tương quan chéo thứ nhất được chèn khoảng trống từ sự dịch chuyển thời gian mà giá trị tương quan chéo thứ ba thu được, bởi bội số lẻ của một nửa thời khoảng chu kỳ (p) của tần số cơ bản của nội dung âm thanh.

Phương án thứ mười bốn đề cập đến bộ định tỉ lệ thời gian theo một trong số các phương án từ 9 đến 13, trong đó số đo độ tương tự thứ hai q thu được theo

$$q=c(p)*c(2*p)+c(3/2*p)*c(1/2*p)$$

hoặc theo

$$q=c(p)*c(-p)+c(-1/2*p)*c(1/2*p),$$

trong đó $c(p)$ là giá trị tương quan chéo giữa khối các mẫu thứ nhất và khối các mẫu thứ hai, mà được dịch chuyển trong thời gian bởi thời khoảng chu kỳ p của tần số cơ bản của nội dung âm thanh của khối các mẫu thứ nhất hoặc của khối các mẫu thứ hai;

trong đó $c(2*p)$ là giá trị tương quan chéo giữa khối các mẫu thứ nhất và khối các mẫu thứ hai, mà được dịch chuyển trong thời gian bởi $2*p$;

trong đó $c(3/2*p)$ là giá trị tương quan chéo giữa khối các mẫu thứ nhất và khối các mẫu thứ hai, mà được dịch chuyển trong thời gian bởi $3/2*p$;

trong đó $c(1/2*p)$ là giá trị tương quan chéo giữa khối các mẫu thứ nhất và khối các mẫu thứ hai, mà được dịch chuyển trong thời gian bởi $1/2*p$;

trong đó $c(-p)$ là giá trị tương quan chéo giữa khối các mẫu thứ nhất và khối các mẫu thứ hai, mà được dịch chuyển trong thời gian bởi $-p$; và

trong đó $c(-1/2*p)$ là giá trị tương quan chéo giữa khối các mẫu thứ nhất và khối các mẫu thứ hai, mà được dịch chuyển trong thời gian bởi $-1/2*p$.

Phương án thứ mười lăm đề cập đến bộ định tỉ lệ thời gian (200; 340; 450; 866; 900; 1000) theo một trong số các phương án từ 1 đến 14,

trong đó bộ định tỉ lệ thời gian được tạo cấu hình để so sánh (1064) giá trị chất lượng (q), mà dựa trên sự tính toán hoặc sự ước lượng chất lượng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bằng sự định tỉ lệ

thời gian, với giá trị ngưỡng có thể biến thiên (q_{min}), để quyết định liệu việc định tỉ lệ thời gian nên được thực hiện hay không.

Phương án thứ mười sáu đề cập đến bộ định tỉ lệ thời gian (200; 340; 450; 866; 900; 1000) theo phương án 15, trong đó bộ định tỉ lệ thời gian được tạo cấu hình để giảm giá trị ngưỡng có thể thay đổi (q_{min}), để nhờ đó giảm yêu cầu chất lượng, phản hồi sự tìm ra rằng chất lượng của sự định tỉ lệ thời gian sẽ không đủ cho một hoặc nhiều khối các mẫu trước đó.

Phương án thứ mười bảy đề cập đến bộ định tỉ lệ thời gian (200; 340; 450; 866; 900; 1000) theo phương án 15 hoặc phương án 16, trong đó bộ định tỉ lệ thời gian được tạo cấu hình để tăng giá trị ngưỡng có thể thay đổi (q_{min}), để nhờ đó tăng yêu cầu chất lượng, phản hồi thực tế rằng chất lượng của sự định tỉ lệ thời gian đã được áp dụng cho một hoặc nhiều khối các mẫu trước đó.

Phương án thứ mười tám đề cập đến bộ định tỉ lệ thời gian (200; 340; 450; 866; 900; 1000) theo một trong số các phương án từ 15 đến 17,

trong đó bộ định tỉ lệ thời gian bao gồm bộ đếm thứ nhất được giới hạn phạm vi (n_{Scaled}) để đếm số lượng khối các mẫu hoặc số lượng khung mà đã được định tỉ lệ thời gian vì sự yêu cầu chất lượng tương ứng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bằng sự định tỉ lệ thời gian đã đạt được, và

trong đó bộ định tỉ lệ thời gian bao gồm bộ đếm thứ hai được giới hạn phạm vi ($n_{NotScaled}$) để đếm số lượng khối các mẫu hoặc số lượng khung mà chưa được định tỉ lệ thời gian vì sự yêu cầu chất lượng tương ứng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bằng sự định tỉ lệ thời gian chưa đạt được; và

trong đó bộ định tỉ lệ thời gian được tạo cấu hình để tính toán giá trị ngưỡng biến thiên (q_{min}) phụ thuộc vào giá trị bộ đếm thứ nhất (n_{Scaled}) và phụ thuộc vào giá trị của bộ đếm thứ hai ($n_{NotScaled}$).

Phương án thứ mười chín đề cập đến bộ định tỉ lệ thời gian (200; 340; 450; 866; 900; 1000) theo phương án 18, trong đó bộ định tỉ lệ thời gian được tạo cấu hình để

cộng giá trị mà tỉ lệ với giá trị bộ đếm thứ nhất (nScaled) đến giá trị ngưỡng ban đầu, và để trừ giá trị mà tỉ lệ với giá trị của bộ đếm thứ hai (nNotScaled) từ đó, để thu giá trị ngưỡng biến thiên (qmin).

Phương án thứ hai mươi đề cập đến bộ định tỉ lệ thời gian (200; 340; 450; 866; 900; 1000) theo một trong số các phương án từ 1 đến 19, trong đó bộ định tỉ lệ thời gian được tạo cấu hình để thực hiện sự định tỉ lệ thời gian của tín hiệu âm thanh đầu vào phụ thuộc vào sự tính toán hoặc ước lượng (950;1060) chất lượng (q) của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bằng sự định tỉ lệ thời gian, trong đó sự tính toán hoặc sự ước lượng chất lượng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào bao gồm sự tính toán hoặc ước lượng của các thành phần lạ trong phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào mà sẽ được gây ra bởi sự định tỉ lệ thời gian.

Phương án thứ hai mươi một đề cập đến bộ định tỉ lệ thời gian (200; 340; 450; 866; 900; 1000) theo phương án 20, trong đó sự tính toán hoặc ước lượng (950;1060) chất lượng âm thanh (q) của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh bao gồm sự tính toán hoặc sự ước lượng các thành phần lạ trong phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào mà sẽ được gây ra bởi thao tác chồng lấp và cộng (954;1068) của các khối các mẫu tiếp theo của tín hiệu âm thanh đầu vào.

Phương án thứ hai mươi hai đề cập đến bộ định tỉ lệ thời gian (200; 340; 450; 866; 900; 1000) theo một trong số các phương án từ 1 đến 21, trong đó bộ định tỉ lệ thời gian được tạo cấu hình để tính toán hoặc ước lượng (950;1060) chất lượng (q) của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bởi sự định tỉ lệ thời gian của tín hiệu âm thanh đầu vào phụ thuộc vào mức tương tự của các khối các mẫu tiếp theo của tín hiệu âm thanh.

Phương án thứ hai mươi ba đề cập đến bộ định tỉ lệ thời gian (200; 340; 450; 866; 900; 1000) theo một trong số các phương án từ 1 đến 22, trong đó bộ định tỉ lệ thời gian được tạo cấu hình để tính toán hoặc ước lượng liệu có các thành phần lạ có thể nghe được trong phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bởi tín hiệu âm thanh đầu vào.

Phương án thứ hai mươi tư đề cập đến bộ định tỉ lệ thời gian (200; 340; 450; 866; 900; 1000) theo một trong số các phương án từ 1 đến 23, trong đó bộ định tỉ lệ thời gian được tạo cấu hình để hoãn (1076) sự định tỉ lệ thời gian đến khung tiếp theo hoặc khỏi các mẫu tiếp theo nếu sự tính toán hoặc ước lượng chất lượng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bởi sự định tỉ lệ thời gian biểu thị chất lượng không đủ.

Phương án thứ hai mươi lăm đề cập đến bộ định tỉ lệ thời gian (200; 340; 450; 866; 900; 1000) theo một trong số các phương án từ 1 đến 24, trong đó bộ định tỉ lệ thời gian được tạo cấu hình để hoãn sự định tỉ lệ thời gian đến thời gian khi sự định tỉ lệ thời gian ít có thể nghe được nếu sự tính toán hoặc sự ước lượng của chất lượng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bởi sự định tỉ lệ thời gian biểu thị chất lượng không đủ.

Phương án thứ hai mươi sáu đề cập đến bộ giải mã âm thanh (300) để cung cấp nội dung âm thanh được giải mã (312) trên cơ sở nội dung âm thanh đầu vào (310), bộ giải mã âm thanh bao gồm:

vùng đệm chấp chờn (320) được tạo cấu hình để làm đệm nhiều khung âm thanh biểu diễn các khối các mẫu âm thanh;

lỗi bộ giải mã (330) được tạo cấu hình để cung cấp các khối (332) của các mẫu âm thanh trên cơ sở các khung âm thanh (322) được nhận từ vùng đệm chấp chờn;

bộ định tỉ lệ thời gian trên cơ sở mẫu (200; 340; 450; 866; 900; 1000) theo một trong số các phương án từ 1 đến 25, trong đó bộ định tỉ lệ thời gian trên cơ sở mẫu được tạo cấu hình để cung cấp các khối các mẫu âm thanh được định tỉ lệ thời gian (342) trên cơ sở các khối các mẫu âm thanh (332) được cung cấp bởi lỗi bộ giải mã.

Phương án thứ hai mươi bảy đề cập đến bộ giải mã (300) theo phương án 26, trong đó bộ giải mã âm thanh còn bao gồm bộ điều khiển vùng đệm chấp chờn (100;350;490;800),

trong đó bộ điều khiển vùng đệm chấp chờn được tạo cấu hình để cung cấp thông tin điều khiển (114; 444) cho bộ định tỉ lệ thời gian trên cơ sở mẫu (200; 340; 450; 866; 900; 1000), trong đó thông tin điều khiển biểu thị liệu sự định tỉ lệ thời gian

trên cơ sở mẫu nên được thực hiện hay không, và/hoặc trong đó thông tin điều khiển biểu thị lượng định tỉ lệ thời gian mong muốn.

Phương án thứ hai mươi tám đề cập đến phương pháp (1500) cung cấp phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào,

trong đó phương pháp bao gồm tính toán hoặc ước lượng (1510) chất lượng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bởi sự định tỉ lệ thời gian của tín hiệu âm thanh đầu vào, và

trong đó phương pháp bao gồm thực hiện (1520) sự định tỉ lệ thời gian của tín hiệu âm thanh đầu vào phụ thuộc vào sự tính toán hoặc ước lượng của chất lượng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bằng sự định tỉ lệ thời gian.

Phương án thứ hai mươi chín đề cập đến chương trình máy tính thực hiện phương pháp theo phương án 28 khi chương trình máy tính chạy trên máy tính.

YÊU CẦU BẢO HỘ

1. Bộ định tỉ lệ thời gian (200; 340; 450; 866; 900; 1000) để cung cấp phiên bản được định tỉ lệ thời gian (212;312;448;956) của tín hiệu âm thanh đầu vào (210; 332;442;910),

trong đó bộ định tỉ lệ thời gian được tạo cấu hình để tính toán hoặc ước lượng (950; 1060) chất lượng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bởi sự định tỉ lệ thời gian của tín hiệu âm thanh đầu vào, và

trong đó bộ định tỉ lệ thời gian được tạo cấu hình để thực hiện (954; 1068) sự định tỉ lệ thời gian của tín hiệu âm thanh đầu vào phụ thuộc vào sự tính toán hoặc sự ước lượng chất lượng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bởi sự định tỉ lệ thời gian;

trong đó bộ định tỉ lệ thời gian được tạo cấu hình để so sánh (1064) giá trị chất lượng (q), mà dựa trên sự tính toán hoặc sự ước lượng chất lượng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bằng sự định tỉ lệ thời gian, với giá trị ngưỡng có thể biến thiên (qmin), để quyết định liệu việc định tỉ lệ thời gian nên được thực hiện hay không;

trong đó bộ định tỉ lệ thời gian được tạo cấu hình để tăng giá trị ngưỡng biến thiên (qmin), để nhờ đó tăng sự yêu cầu chất lượng, phản hồi thực tế rằng sự định tỉ lệ thời gian đã được áp dụng cho một hoặc nhiều khối các mẫu trước đó, sao cho đảm bảo rằng khối các mẫu tiếp theo chỉ được định tỉ lệ thời gian nếu mức chất lượng tương đối cao, cao hơn mức chất lượng thông thường, có thể đạt được.

2. Bộ giải mã âm thanh (300) để cung cấp nội dung âm thanh được giải mã (312) trên cơ sở nội dung âm thanh đầu vào (310), bộ giải mã âm thanh bao gồm:

vùng đệm chấp chờn (320) được tạo cấu hình để làm đệm nhiều khung âm thanh biểu diễn các khối các mẫu âm thanh;

lỗi bộ giải mã (330) được tạo cấu hình để cung cấp các khối (332) của các mẫu âm thanh trên cơ sở các khung âm thanh (322) được nhận từ vùng đệm chấp chờn;

bộ định tỉ lệ thời gian trên cơ sở mẫu (200; 340; 450; 866; 900; 1000) theo điểm 1, trong đó bộ định tỉ lệ thời gian trên cơ sở mẫu được tạo cấu hình để cung cấp các

khối các mẫu âm thanh được định tỉ lệ thời gian (342) trên cơ sở các khối các mẫu âm thanh (332) được cung cấp bởi lõi bộ giải mã.

3. Bộ giải mã (300) theo điểm 2, trong đó bộ giải mã âm thanh còn bao gồm bộ điều khiển vùng đệm chấp chờn (100;350;490;800),

trong đó bộ điều khiển vùng đệm chấp chờn được tạo cấu hình để cung cấp thông tin điều khiển (114; 444) cho bộ định tỉ lệ thời gian trên cơ sở mẫu (200; 340; 450; 866; 900; 1000), trong đó thông tin điều khiển biểu thị liệu sự định tỉ lệ thời gian trên cơ sở mẫu nên được thực hiện hay không, và/hoặc trong đó thông tin điều khiển biểu thị lượng định tỉ lệ thời gian mong muốn.

4. Phương pháp cung cấp phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào (1500),

trong đó phương pháp bao gồm tính toán hoặc ước lượng (1510) chất lượng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bởi sự định tỉ lệ thời gian của tín hiệu âm thanh đầu vào, và

trong đó phương pháp bao gồm thực hiện (1520) sự định tỉ lệ thời gian của tín hiệu âm thanh đầu vào phụ thuộc vào sự tính toán hoặc ước lượng của chất lượng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bằng sự định tỉ lệ thời gian;

trong đó phương pháp bao gồm so sánh (1064) giá trị chất lượng (q), mà dựa trên sự tính toán hoặc sự ước lượng chất lượng của phiên bản được định tỉ lệ thời gian của tín hiệu âm thanh đầu vào có thể thu được bằng sự định tỉ lệ thời gian, với giá trị ngưỡng biến thiên (q_{min}), để quyết định liệu việc định tỉ lệ thời gian nên được thực hiện hay không;

trong đó phương pháp bao gồm tăng giá trị ngưỡng biến thiên (q_{min}), để nhờ đó tăng sự yêu cầu chất lượng, phản hồi thực tế rằng sự định tỉ lệ thời gian đã được áp dụng cho một hoặc nhiều khối các mẫu trước đó, sao cho đảm bảo rằng khối các mẫu tiếp theo chỉ được định tỉ lệ thời gian nếu mức chất lượng tương đối cao, cao hơn mức chất lượng thông thường, có thể đạt được.

5. Vật ghi có thể đọc được bằng máy tính bao gồm chương trình máy tính thực hiện phương pháp theo điểm 4 khi chương trình máy tính chạy trên máy tính.

1/14

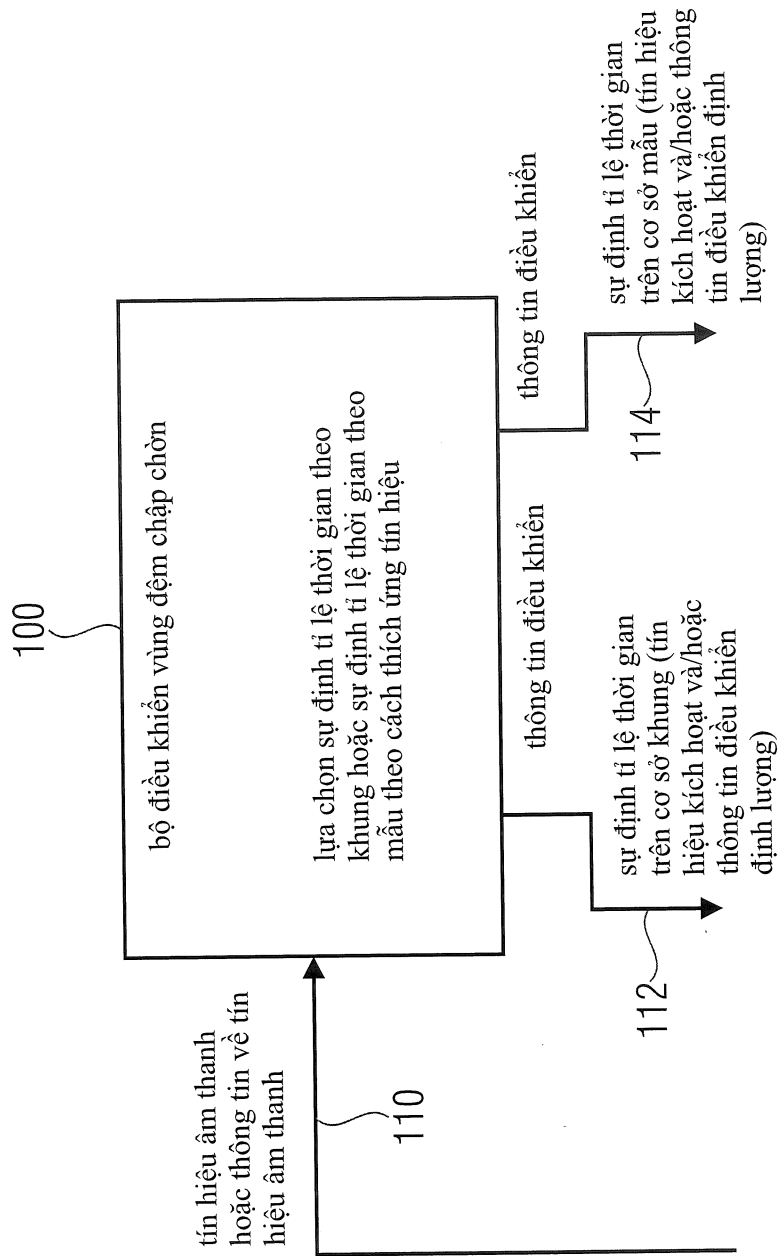


FIG 1

2/14

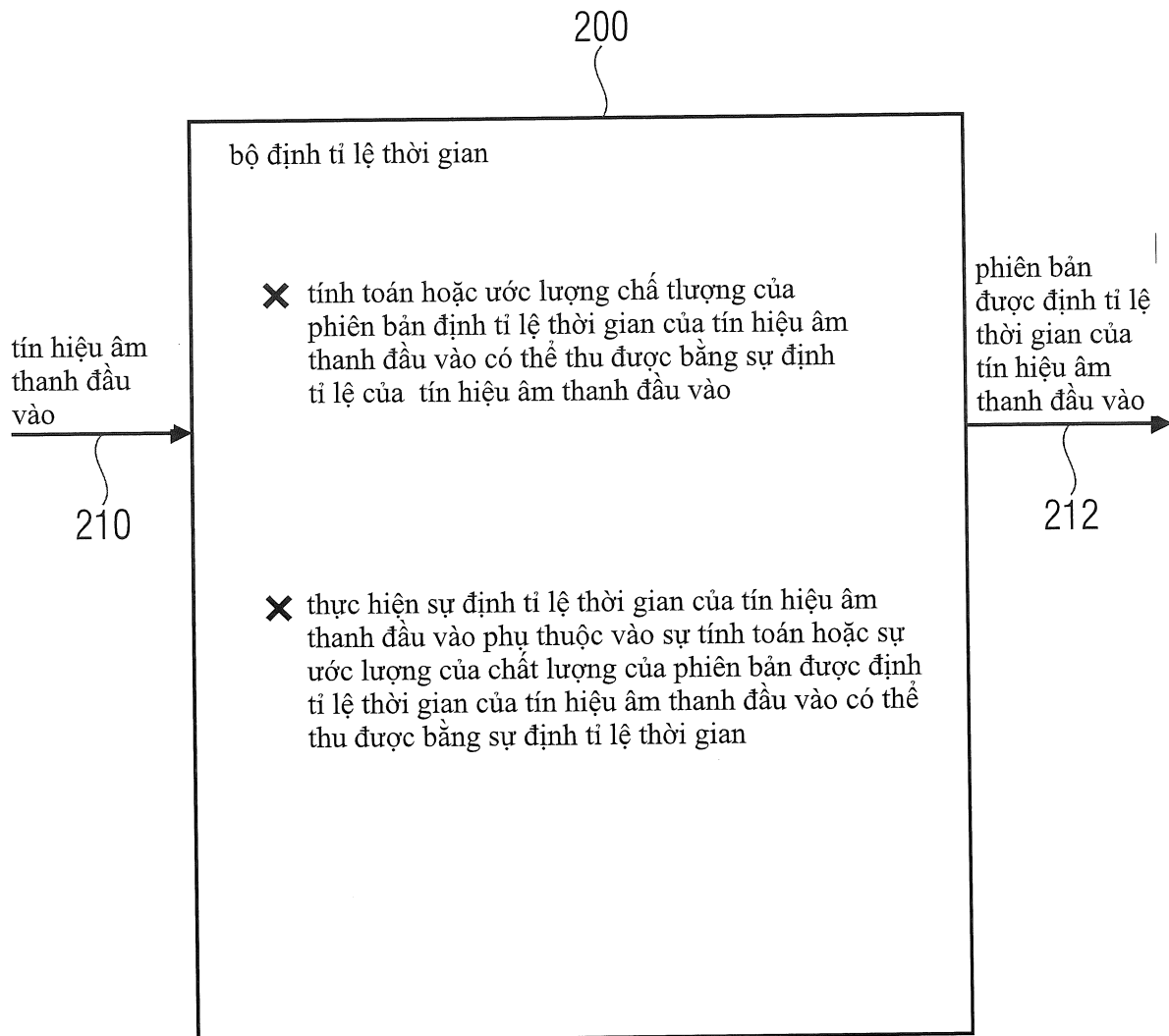


FIG 2

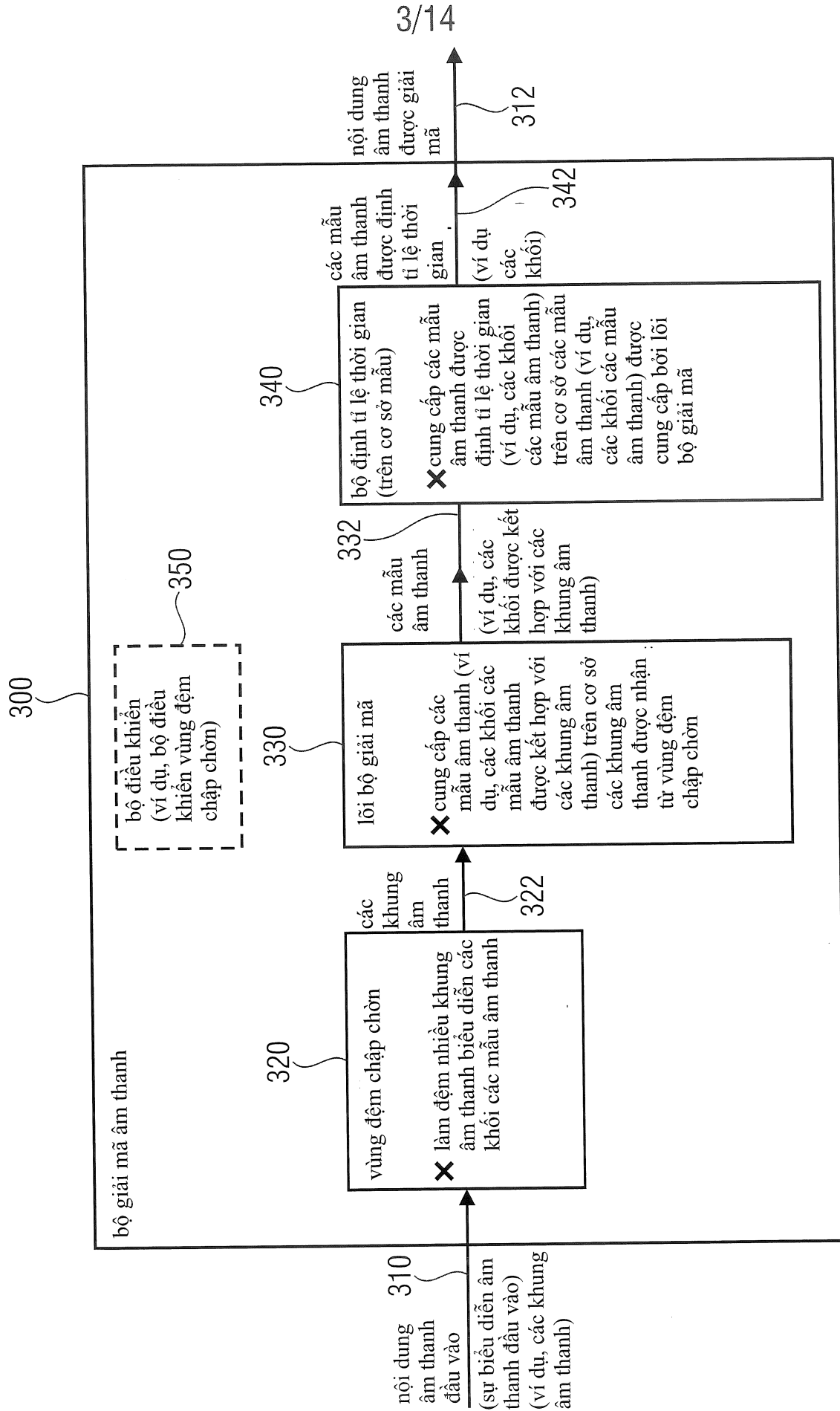


FIG 3

5/14

```

soundCardFrameSize = sampleRate * 20 / 1000;
while(pcmBuffer_nReadableSamples() < soundCardFrameSize) {
    accessUnit = deJitterBuffer_pop();           ← 510
    audioSamples = decoder_decodeOrConceal(accessUnit); ← 512
    tsm_scale(audioSamples);                     ← 514
    pcmBuffer_write(audioSamples);               ← 516
}
audioSamples = pcmBuffer_read(soundCardFrameSize); ← 520
soundcard_push(audioSamples);                   ← 522

```

FIG 5

```

rtpTimeDiff = rtpTimeStamp - prevRtpTimeStamp; ← 610
rcvTimeDiff = rcvTime - prevRcvTime;           ← 612
/* convert time base to milliseconds */
rtpTimeTicks = rtpTimeStamp * sysTimeScale / rtpTimeScale; ← 614
rtpTimeDiff = rtpTimeDiff * sysTimeScale / rtpTimeScale; ← 616
delay = rcvTimeDiff - rtpTimeDiff + prevDelay; ← 618
offset = rcvTime - rtpTimeTicks;                ← 620
longTermFifo_add( delay, offset, rtpTimeTicks ); ← 622
/* backup values of current packet */
prevRtpTimeStamp = rtpTimeStamp;
prevRcvTime      = rcvTime;
prevDelay        = delay;                       } ← 624

```

FIG 6

```

targetMax      = shortTermJitter + 60;           ← 710
targetMin      = min( longTermJitter + 20, targetMax ); ← 712
targetDtx      = min( longTermJitter, shortTermJitter ); ← 714
targetStartUp  = ( targetMin + targetMax ) / 2;   ← 716

```

FIG 7

6/14

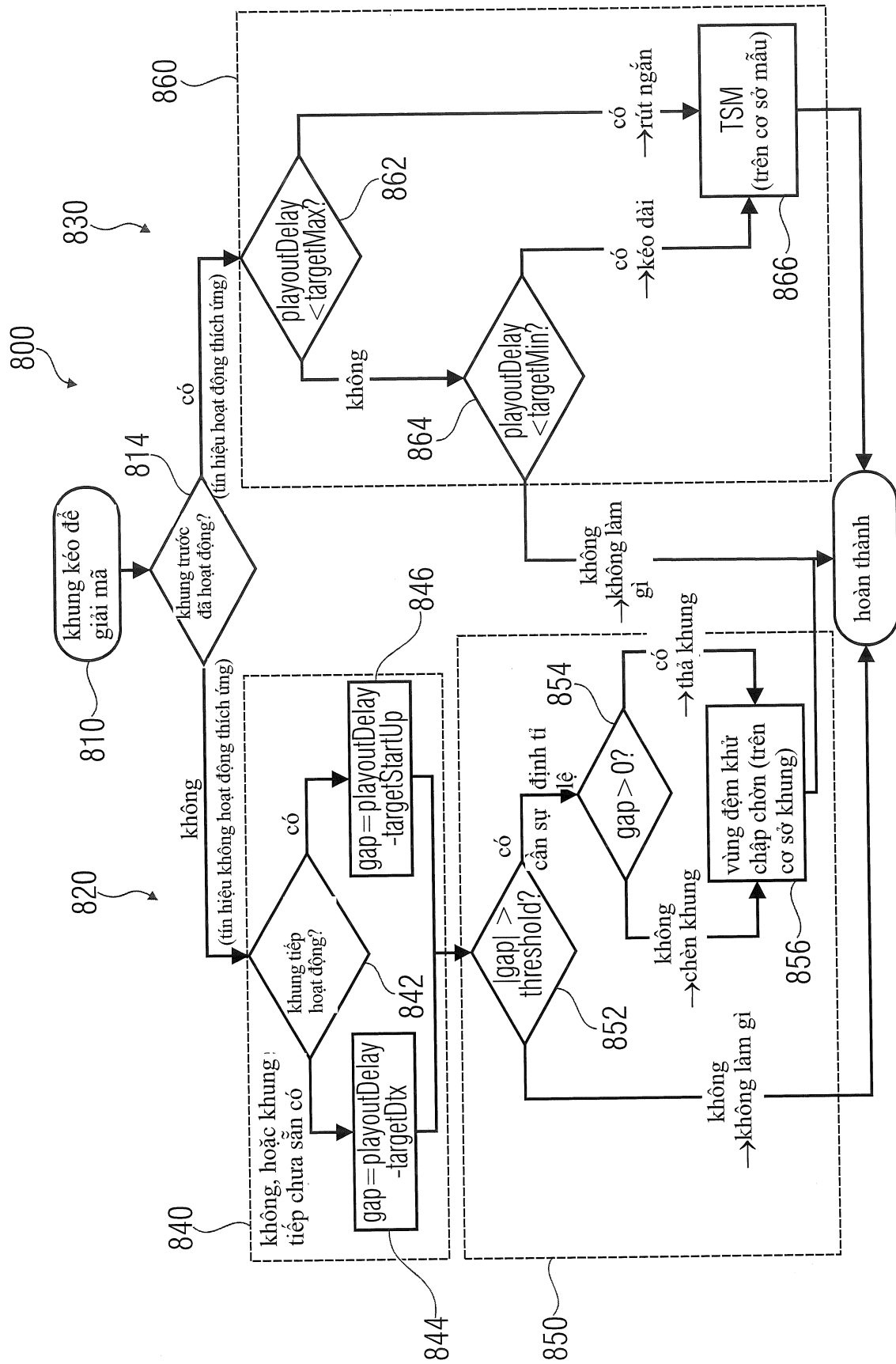


FIG 8

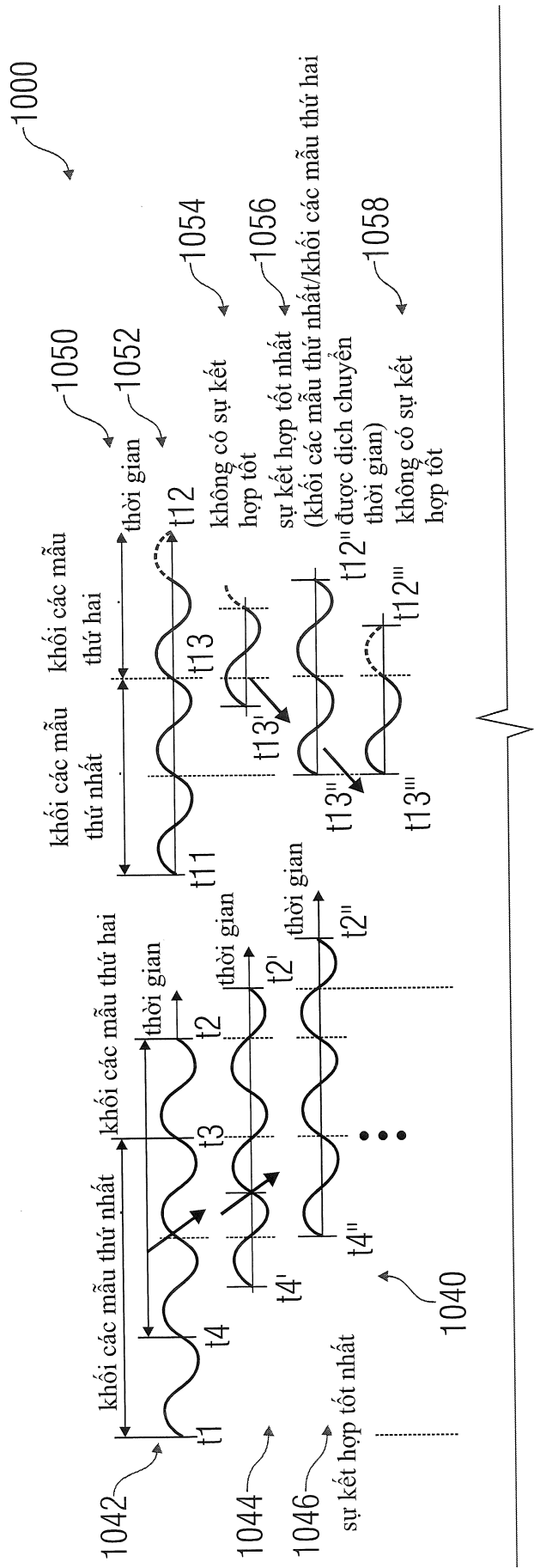


FIG 10A-1

FIG 10A	FIG 10A-1
	FIG 10A-2

FIG 10A-2

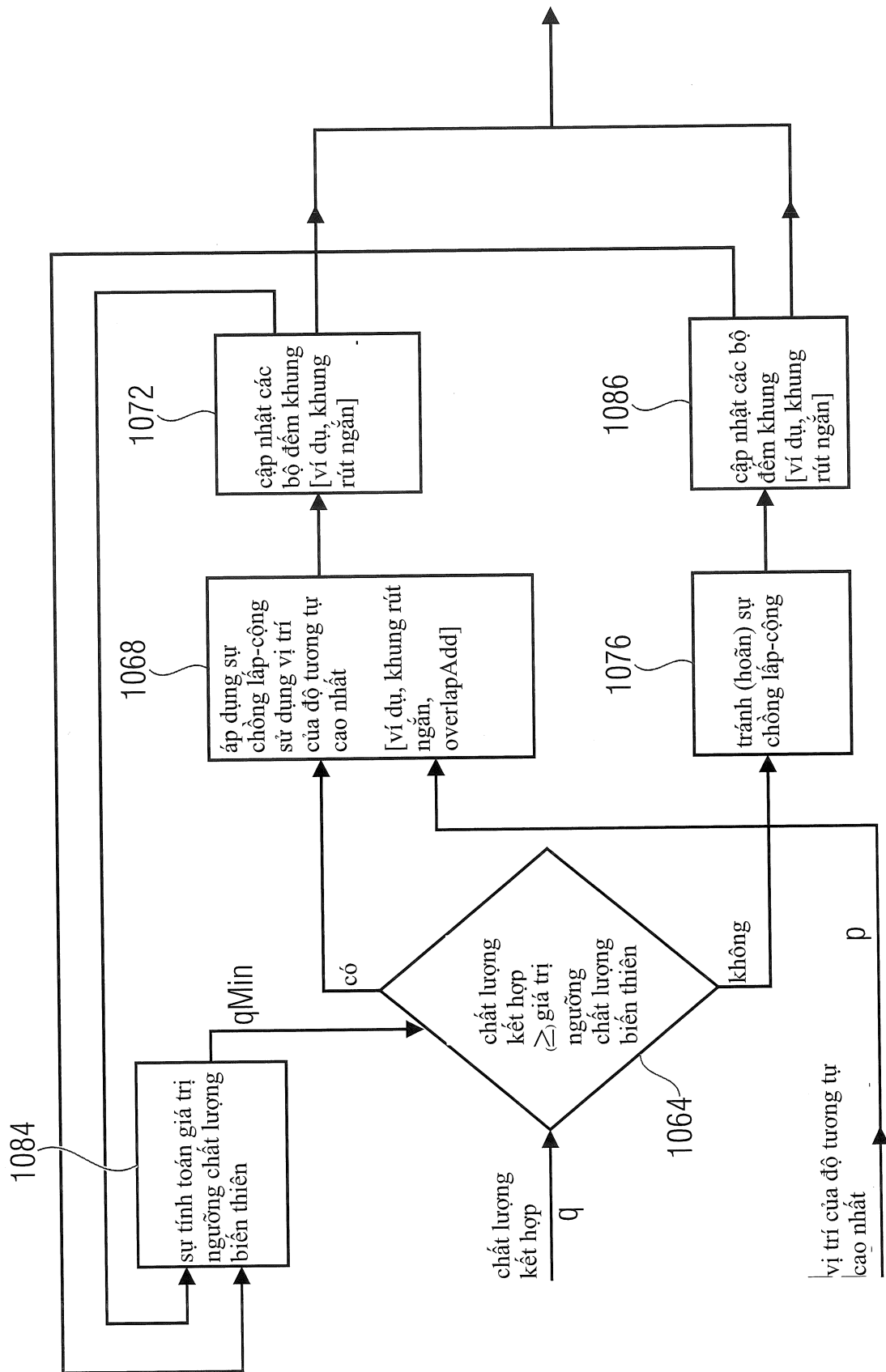


FIG 10B

11/14

```

qMinInitial  = 1;                               ← 1110
qualityRise  = 4;                               ← 1112
qualityRed   = 4;                               ← 1114
p = similarityEstimation();                      ← 1116
q = c(p) * c(2*p) + c(3/2*p) * c(1/2*p);
qMin = qMinInitial - (nNotScaled * 0.1) + (nScaled * 0.2); ← 1118
if(q >= qMin) { /* sufficient quality, apply time-scaling */ ← 1120
    applyOverlapAdd();                            ← 1122
    if( nNotScaled > 0 )                          } ← 1124
        nNotScaled = nNotScaled - 1;
    if( nScaled < qualityRise )                    } ← 1126
        nScaled = nScaled + 1;
}
else { /* not sufficient, postpone time-scaling */
    if( nNotScaled < qualityRed )                  } ← 1128
        nNotScaled = nNotScaled + 1;
    if( nScaled > 0 )                              } ← 1130
        nScaled = nScaled - 1;
}

```

FIG 11

12/14

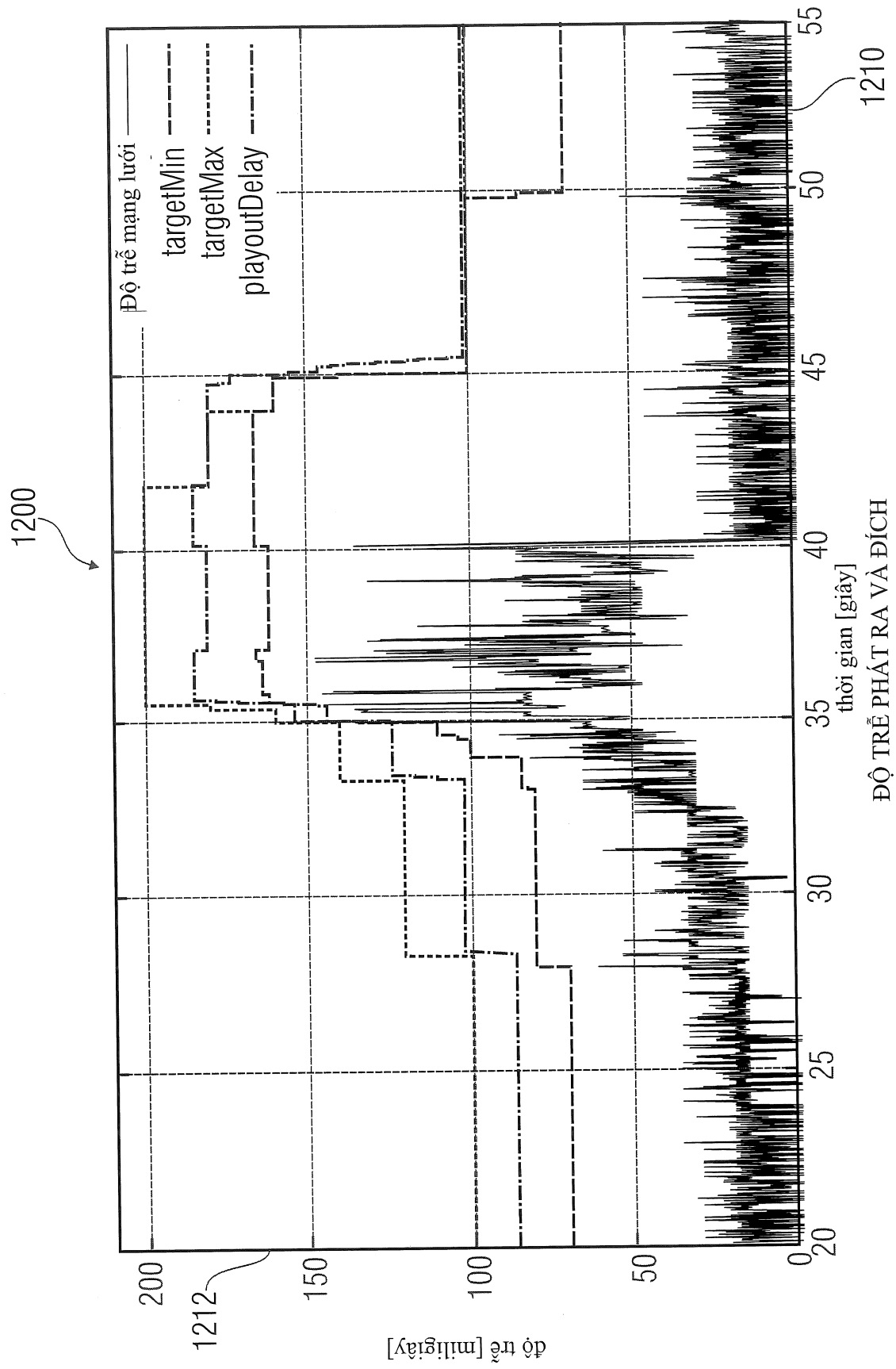
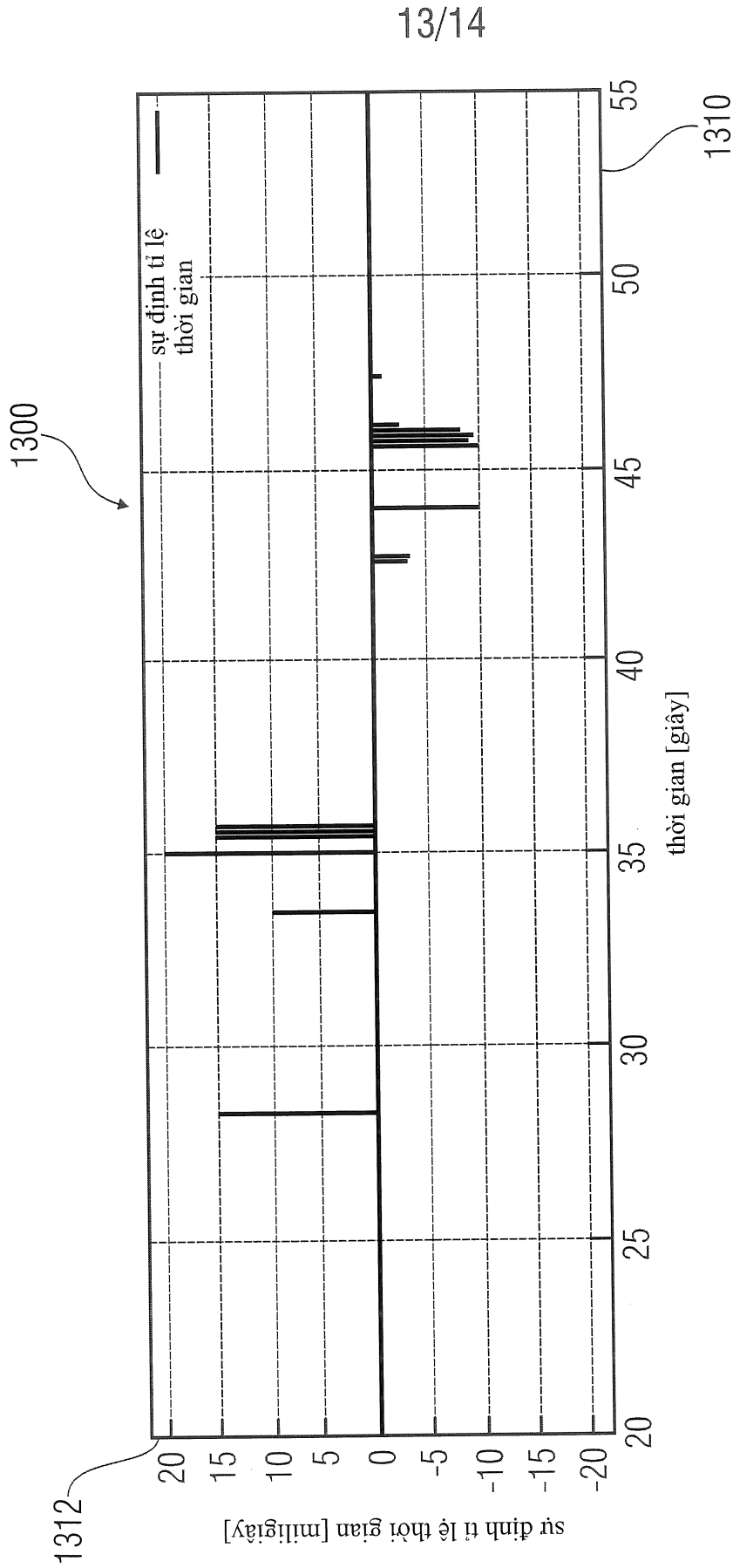


FIG 12



SỰ ĐỊNH TỈ LỆ THỜI GIAN

FIG 13

14/14

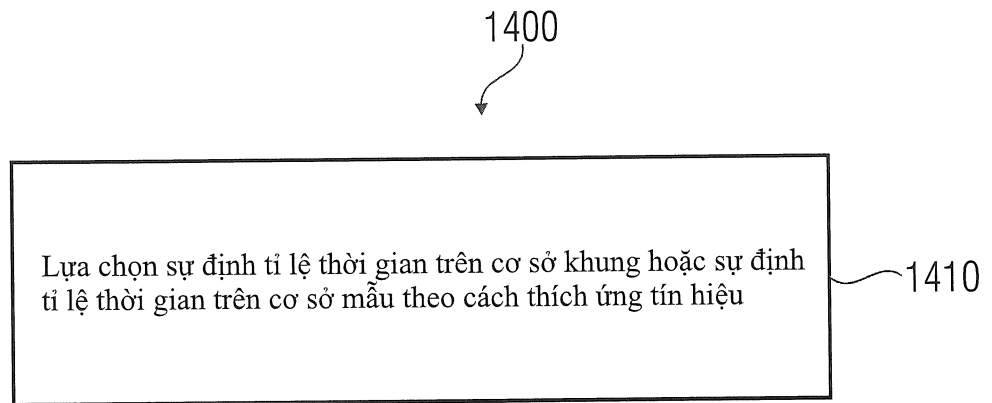


FIG 14

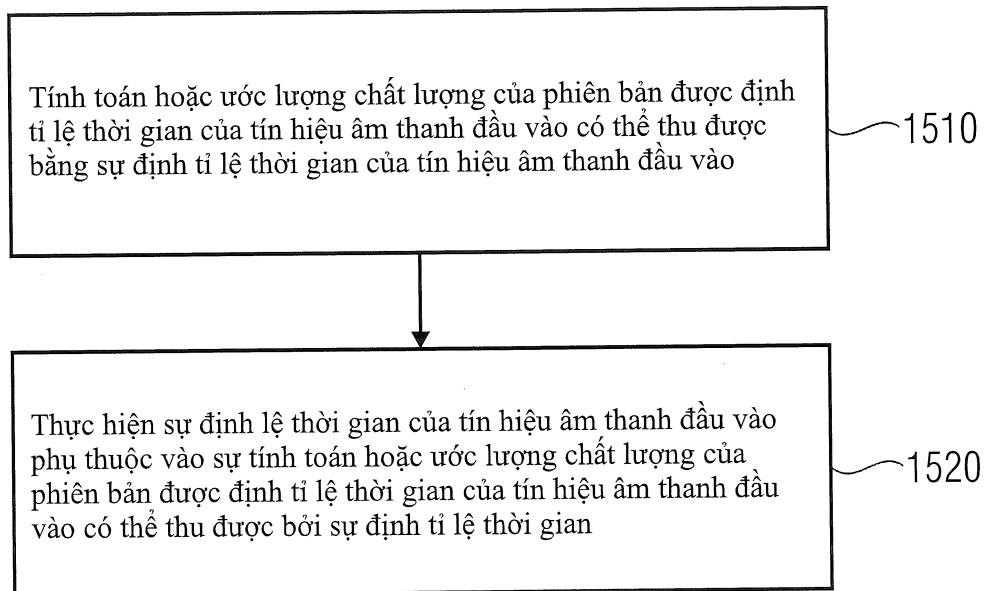


FIG 15