



(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)



1-0029643

(51)⁷ G10L 19/20

(13) B

(21) 1-2017-00216

(22) 23/07/2015

(86) PCT/EP2015/066953 23/07/2015

(87) WO 2016/016105 04/02/2016

(30) 14178830.7 28/07/2014 EP

(45) 25/10/2021 403

(43) 25/05/2017 350A

(73) FRAUNHOFER-GESELLSCHAFT ZUR FOERDERUNG DER ANGEWANDTEN
FORSCHUNG E.V. (DE)

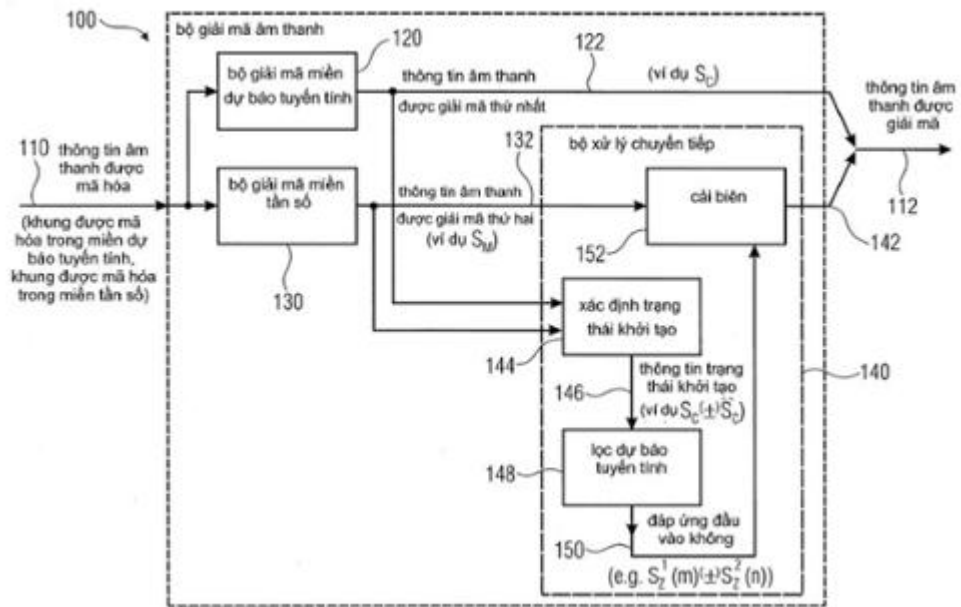
Hansastrasse 27c, 80686 Muenchen Germany

(72) RAVELLI, Emmanuel (FR); FUCHS, Guillaume (FR); DISCH, Sascha (DE);
MULTRUS, Makus (DE); PIETRZYK, Grzegorz (PL); SCHUBERT, Benjamin
(DE).

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

(54) BỘ GIẢI MÃ ÂM THANH VÀ PHƯƠNG PHÁP CUNG CẤP THÔNG TIN ÂM
THANH ĐƯỢC GIẢI MÃ TRÊN CƠ SỞ THÔNG TIN ÂM THANH ĐƯỢC MÃ
HÓA

(57) Sáng chế đề cập đến bộ giải mã âm thanh và phương pháp cung cấp thông tin âm thanh được giải mã trên cơ sở thông tin âm thanh được mã hóa. Bộ giải mã âm thanh (100;200;300) để cung cấp thông tin âm thanh được giải mã (112;212;312) trên cơ sở thông tin âm thanh được mã hóa (110;210;310), bộ giải mã âm thanh bao gồm bộ giải mã miền dự báo tuyến tính (120;220;320) được tạo cấu hình để cung cấp thông tin âm thanh được giải mã thứ nhất (122;222;322; $S_C(n)$) trên cơ sở khung âm thanh được mã hóa trong miền dự báo tuyến tính, bộ giải mã miền tần số (130;230;330) được tạo cấu hình để cung cấp thông tin âm thanh được giải mã thứ hai (132;232;332; $S_M(n)$) trên cơ sở khung âm thanh được mã hóa trong miền tần số, và bộ xử lý chuyển tiếp (140;240;340). Bộ xử lý chuyển tiếp được tạo cấu hình để thu được đáp ứng đầu vào không (150; 256;348) của phép lọc dự báo tuyến tính (148; 254; 346), trong đó trạng thái khởi tạo (146;252;344) của phép lọc dự báo tuyến tính được xác định phụ thuộc vào thông tin âm thanh được giải mã thứ nhất và thông tin âm thanh được giải mã thứ hai. Bộ xử lý chuyển tiếp còn được tạo cấu hình để cải biên thông tin âm thanh được giải mã thứ hai (132; 232;332; $S_M(n)$), mà được cung cấp trên cơ sở khung âm thanh được mã hóa trong miền tần số sau khung âm thanh được mã hóa trong miền dự báo tuyến tính, phụ thuộc vào đáp ứng đầu vào không, để thu được sự chuyển tiếp mượt giữa thông tin âm thanh được giải mã thứ nhất ($S_C(n)$) và thông tin âm thanh được giải mã thứ hai được cải biên ($\widetilde{S_M}(n)$).



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến bộ giải mã âm thanh để cung cấp thông tin âm thanh được giải mã trên cơ sở thông tin âm thanh được mã hóa.

Sáng chế còn đề cập đến phương pháp cung cấp thông tin âm thanh được giải mã trên cơ sở thông tin âm thanh được mã hóa.

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

Nói chung, các phương án theo sáng chế được đề cập đến việc xử lý quá trình chuyển tiếp từ bộ mã hóa-giải mã CELP thành bộ mã hóa-giải mã trên cơ sở MDCT trong mã hóa âm thanh được chuyển.

Tình trạng kỹ thuật của sáng chế

Trong những năm gần đây, nhu cầu truyền và lưu trữ thông tin âm thanh được mã hóa tăng lên. Nhu cầu mã hóa âm thanh và giải mã âm thanh của các tín hiệu âm thanh bao gồm cả tiếng nói và âm thanh nói chung (như ví dụ, âm nhạc, tạp âm nền, và âm thanh tương tự) cũng tăng lên.

Để cải thiện chất lượng mã hóa và cũng để cải thiện hiệu quả tốc độ bit, các bộ mã hóa-giải mã âm thanh được chuyển (hoặc chuyển) được đưa ra mà chuyển giữa các sơ đồ mã hóa khác nhau, chẳng hạn như, ví dụ, khung thứ nhất được mã hóa sử dụng khái niệm mã hóa thứ nhất (ví dụ, khái niệm mã hóa trên cơ sở CELP), và sao cho khung âm thanh thứ hai tiếp theo được mã hóa sử dụng khái niệm mã hóa thứ hai khác (ví dụ, khái niệm mã hóa trên cơ sở MDCT). Nói cách khác, có thể có sự chuyển giữa mã hóa miền mã hóa dự báo tuyến tính (ví dụ, sử dụng khái niệm mã hóa trên cơ sở CELP) và mã hóa miền tần số (ví dụ, mã hóa dựa trên biến đổi miền thời gian sang miền tần số hoặc biến đổi miền tần số sang miền thời gian, như ví dụ, biến đổi FFT, biến đổi FFT nghịch đảo, biến đổi MDCT hoặc biến đổi MDCT nghịch đảo). Ví dụ, khái niệm mã hóa thứ nhất có thể là khái niệm mã hóa trên cơ sở CELP, khái niệm mã

hóa trên cơ sở ACELP, khái niệm mã hóa trên cơ sở miền dự báo tuyến tính kích thích được mã hóa biến đổi, hoặc tương tự. Khái niệm mã hóa thứ hai có thể là, ví dụ, khái niệm mã hóa trên cơ sở FFT, khái niệm mã hóa trên cơ sở MDCT, khái niệm mã hóa trên cơ sở AAC hoặc khái niệm mã hóa mà có thể được xem như khái niệm kế tiếp của khái niệm mã hóa trên cơ sở AAC.

Dưới đây, một số ví dụ về các bộ mã hóa-giải mã âm thanh truyền thống (các bộ mã hóa và/hoặc giải mã) sẽ được mô tả.

Các bộ mã hóa-giải mã âm thanh được chuyển, như ví dụ, MPEG USAC, dựa trên hai sơ đồ mã hóa âm thanh chính. Một sơ đồ mã hóa là, ví dụ, bộ mã hóa-giải mã CELP, được nhắm mục tiêu cho các tín hiệu tiếng nói. Sơ đồ mã hóa còn lại là, ví dụ, các bộ mã hóa-giải mã trên cơ sở MDCT (sau đây được gọi đơn giản là MDCT), được nhắm mục tiêu cho tất cả các tín hiệu âm thanh khác (ví dụ, âm nhạc, tạp âm nền). Trên các tín hiệu có nội dung hỗn hợp (ví dụ, tiếng nói trên âm nhạc), bộ mã hóa (và do đó cũng có bộ giải mã) thường chuyển giữa hai sơ đồ mã hóa. Sau đó cần phải tránh các giả âm bất kỳ (ví dụ, cú kích chuột do sự gián đoạn) khi chuyển từ một chế độ (hoặc sơ đồ mã hóa) sang chế độ kia.

Các bộ mã hóa-giải mã âm thanh được chuyển có thể, ví dụ, chứa các vấn đề mà được gây ra bởi các chuyển tiếp CELP sang MDCT.

Các chuyển tiếp CELP sang MDCT thường đưa ra hai vấn đề. Răng cưa có thể bị tạo ra do mất khung MDCT trước đó. Sự gián đoạn có thể bị tạo ra tại biên giữa khung CELP và khung MDCT, do bản chất mã hóa dạng sóng không hoàn hảo của hai sơ đồ mã hóa hoạt động tại các tốc độ bit thấp/trung bình.

Nhiều phương thức tiếp cận đã tồn tại để giải quyết các vấn đề được đưa ra bởi các chuyển tiếp CELP thành MDCT, và sẽ được thảo luận sau đây.

Phương thức tiếp cận có khả năng được mô tả trong bài báo "các cửa sổ nhiễu xuyên kênh hiệu quả cho sự chuyển tiếp giữa mã hóa âm thanh trên cơ sở LPC và không trên cơ sở LPC" bởi JeremieLecomte, Philippe Gournay, Ralf Geiger, Bruno Bessette và Max Neuendorf (được trình bày tại hội nghị AES lần 126, tháng 5 2009, trang 771). Mục này mô tả phương thức tiếp cận trong phần 4.4.2 "ACELP thành chế

độ không LPD". Sự viện dẫn cũng được thực hiện, ví dụ, đến Fig.8 của bài báo đã nêu. Vấn đề răng cưa được giải quyết trước tiên bằng việc tăng chiều dài MDCT (ở đây từ 1024 thành 1152) sao cho điểm gấp bên trái MDCT được di chuyển tại bên trái của biên giữa các khung CELP và MDCT, sau đó bằng cách thay đổi phần bên trái của cửa sổ MDCT sao cho sự chồng lấp giảm, và cuối cùng bằng cách đưa ra một cách giả răng cưa khuyết sử dụng tín hiệu CELP và thao tác chồng lấp và cộng. Vấn đề gián đoạn được giải quyết tại cùng thời gian nhờ thao tác chồng lấp và cộng.

Phương thức tiếp cận này hoạt động tốt nhưng có nhược điểm là tạo ra độ trễ trong bộ giải mã CELP, độ trễ này bằng chiều dài chồng lấp (ở đây là 128 mẫu).

Phương thức tiếp cận khác được mô tả trong Bằng sáng chế Mỹ số 8,725,503 B2 ngày 13 tháng 5 năm 2014 với tiêu đề "Xóa bỏ răng cưa miền thời gian chuyển tiếp với ứng dụng trong miền tín hiệu ban đầu hoặc được gán trọng số" bởi Bruno Bessette.

Trong phương thức tiếp cận này, chiều dài MDCT không được thay đổi (và không là hình dạng cửa sổ MDCT). Vấn đề răng cưa được giải quyết ở đây bằng việc mã hóa tín hiệu hiệu chỉnh răng cưa với bộ mã hóa trên cơ sở biến đổi riêng rẽ. Các bit thông tin phụ bổ sung được gửi vào trong dòng bit. Bộ giải mã khôi phục tín hiệu hiệu chỉnh răng cưa và cộng nó vào khung MDCT được giải mã. Ngoài ra, đáp ứng đầu vào không (zero input response - ZIR) của bộ lọc tổng hợp CELP được sử dụng để giảm biên độ của tín hiệu hiệu chỉnh răng cưa và để cải thiện hiệu quả mã hóa.

ZIR cũng giúp giảm đáng kể vấn đề gián đoạn.

Phương thức tiếp cận này cũng hoạt động tốt nhưng nhược điểm là nó cần lượng thông tin phụ bổ sung đáng kể và số lượng bit cần thiết thường thay đổi mà không phù hợp cho bộ mã hóa-giải mã tốc độ bit không đổi.

Phương thức tiếp cận khác được mô tả trong đơn đăng kí sáng chế Mỹ số US 2013/0289981 A1 ngày 31 tháng 10 năm 2013 với tiêu đề "Mã hóa âm thanh độ trễ thấp thay thế giữa mã hóa dự báo và mã hóa biến đổi" bởi Stephane Ragot, BalazsKovesi và Pierre Berthet. Theo phương thức tiếp cận đã nêu, MDCT không thay đổi, nhưng phần bên trái của cửa sổ MDCT được thay đổi để giảm chiều dài chồng lấp. Để giải quyết vấn đề răng cưa, phần bắt đầu của khung MDCT được mã hóa sử

dụng bộ mã hóa-giải mã CELP, và sau đó tín hiệu CELP được sử dụng để xóa bỏ răng cửa, hoặc bằng việc thay thế hoàn toàn tín hiệu MDCT hoặc bằng cách đưa vào một cách giả thành phần răng cửa khuyết (tương tự bài báo được đề cập ở trên bởi JeremieLecomte và cộng sự). Vấn đề gián đoạn được giải quyết bởi thao tác chồng lấp và cộng nếu phương thức tiếp cận tương tự bài báo của JeremieLecomte và cộng sự được sử dụng, nếu không thì nó được giải quyết bởi thao tác nhiễu xuyên kênh đơn giản giữa tín hiệu CELP và tín hiệu MDCT.

Tương tự với bằng sáng chế Mỹ số US 8,725,503 B2, phương thức tiếp cận này thường hoạt động tốt nhưng nhược điểm là nó cần lượng thông tin phụ đáng kể do CELP bổ sung.

Bản chất kỹ thuật của sáng chế

Theo quan điểm của các vấn đề truyền thống được mô tả ở trên, mục đích của sáng chế là đề xuất bộ giải mã âm thanh và phương pháp cung cấp thông tin âm thanh được giải mã trên cơ sở thông tin âm thanh được mã hóa bao gồm các đặc tính được cải thiện (ví dụ, sự cân bằng được cải thiện giữa tốc độ bit trên đầu, độ trễ và độ phức tạp) để chuyển giữa các chế độ mã hóa khác nhau.

Phương án theo sáng chế tạo ra các bộ giải mã âm thanh để cung cấp thông tin âm thanh được giải mã trên cơ sở thông tin âm thanh được mã hóa. Bộ giải mã âm thanh bao gồm bộ giải mã miền dự báo tuyến tính được tạo cấu hình để cung cấp thông tin âm thanh được giải mã thứ nhất trên cơ sở khung âm thanh được mã hóa trong miền dự báo tuyến tính và bộ giải mã miền tần số được tạo cấu hình để cung cấp thông tin âm thanh được giải mã thứ hai trên cơ sở khung âm thanh được mã hóa trong miền tần số. Bộ giải mã âm thanh cũng bao gồm bộ xử lý chuyển tiếp. Bộ xử lý chuyển tiếp được tạo cấu hình để thu được đáp ứng đầu vào không của phép lọc dự báo tuyến tính, trong đó trạng thái khởi tạo của phép lọc dự báo tuyến tính được xác định phụ thuộc vào thông tin âm thanh được giải mã thứ nhất và thông tin âm thanh được giải mã thứ hai. Bộ xử lý chuyển tiếp cũng được tạo cấu hình để cải biên thông tin âm thanh được giải mã thứ hai, mà được cung cấp trên cơ sở khung âm thanh được mã hóa trong miền tần số theo khung âm thanh được mã hóa trong miền dự báo tuyến tính, phụ thuộc vào

đáp ứng đầu vào không, để thu được sự chuyển tiếp mượt giữa thông tin âm thanh được giải mã thứ nhất và thông tin âm thanh được giải mã thứ hai được cải biên.

Bộ giải mã âm thanh này được dựa trên cơ sở tìm kiếm mà sự chuyển tiếp mượt giữa khung âm thanh được mã hóa trong miền dự báo tuyến tính và khung âm thanh sau được mã hóa trong miền tần số có thể đạt được bằng cách sử dụng đáp ứng đầu vào không của bộ lọc dự báo tuyến tính để cải biên thông tin âm thanh được giải mã, mà trạng thái khởi tạo của phép lọc dự báo tuyến tính xem xét cả hai thông tin âm thanh được giải mã thứ nhất và thông tin âm thanh được giải mã thứ hai. Theo đó, thông tin âm thanh được giải mã thứ hai có thể được làm thích ứng (được cải biên) sao cho phần bắt đầu của thông tin âm thanh được giải mã thứ hai được cải biên để giống với phần kết thúc của thông tin âm thanh được giải mã thứ nhất, mà giúp giảm, hoặc thậm chí tránh các gián đoạn cơ bản giữa khung âm thanh thứ nhất và khung âm thanh thứ hai. Khi được so sánh với bộ giải mã âm thanh được mô tả ở trên, khái niệm nói chung áp dụng được ngay cả khi thông tin âm thanh được giải mã thứ hai không chứa bất kỳ răng cưa nào. Hơn nữa, cần lưu ý rằng thuật ngữ "lọc dự báo tuyến tính" có thể chỉ cả ứng dụng đơn lẻ của bộ lọc dự báo tuyến tính và nhiều ứng dụng của các bộ lọc dự báo tuyến tính, trong đó cần lưu ý rằng ứng dụng đơn lẻ của phép lọc dự báo tuyến tính tương đương một các điển hình với nhiều ứng dụng của các bộ lọc dự báo tuyến tính đồng nhất, vì các bộ lọc dự báo tuyến tính là tuyến tính điển hình.

Nói tóm lại, bộ giải mã âm thanh nêu trên cho phép thu được sự chuyển tiếp mượt giữa khung âm thanh thứ nhất được mã hóa trong miền dự báo tuyến tính và khung âm thanh thứ hai tiếp theo được mã hóa trong miền tần số (hoặc miền biến đổi), trong đó không có độ trễ nào được tạo ra, và trong đó nỗ lực tính toán là tương đối nhỏ.

Phương án khác theo sáng chế tạo ra các bộ giải mã âm thanh để cung cấp thông tin âm thanh được giải mã trên cơ sở thông tin âm thanh được mã hóa. Bộ giải mã âm thanh bao gồm bộ giải mã miền dự báo tuyến tính được tạo cấu hình để cung cấp thông tin âm thanh được giải mã thứ nhất trên cơ sở khung âm thanh được mã hóa trong miền dự báo tuyến tính (hoặc, một cách tương đương, trong phép biểu diễn miền dự báo tuyến tính). Bộ giải mã âm thanh cũng bao gồm bộ giải mã miền tần số được tạo

cấu hình để cung cấp thông tin âm thanh được giải mã thứ hai trên cơ sở khung âm thanh được mã hóa trong miền tần số (hoặc, một cách tương đương, trong phép biểu diễn miền tần số). Bộ giải mã âm thanh cũng bao gồm bộ xử lý chuyển tiếp. Bộ xử lý chuyển tiếp được tạo cấu hình để thu được đáp ứng đầu vào không của bộ lọc dự báo tuyến tính đáp ứng với trạng thái khởi tạo thứ nhất của bộ lọc dự báo tuyến tính được xác định bởi thông tin âm thanh được giải mã thứ nhất, và để thu được đáp ứng đầu vào không thứ hai của bộ lọc dự báo tuyến tính đáp ứng với trạng thái khởi tạo thứ hai của bộ lọc dự báo tuyến tính được xác định bởi phiên bản được cải biên của thông tin âm thanh được giải mã thứ nhất, mà được cung cấp với răng cưa giả, mà bao gồm sự đóng góp của một phần thông tin âm thanh được giải mã thứ hai. Một cách thay thế, bộ xử lý chuyển tiếp được tạo cấu hình để thu được đáp ứng đầu vào không được kết hợp của bộ lọc dự báo tuyến tính đáp ứng với trạng thái khởi tạo của bộ lọc dự báo tuyến tính được xác định bởi việc kết hợp thông tin âm thanh được giải mã thứ nhất và phiên bản được cải biên của thông tin âm thanh được giải mã thứ nhất, mà được cung cấp với răng cưa giả, mà bao gồm sự đóng góp của một phần thông tin âm thanh được giải mã thứ hai. Bộ xử lý chuyển tiếp cũng được tạo cấu hình để cải biên thông tin âm thanh được giải mã thứ hai, mà được cung cấp trên cơ sở khung âm thanh được mã hóa trong miền tần số theo khung âm thanh được mã hóa trong miền dự báo tuyến tính, phụ thuộc vào đáp ứng đầu vào không thứ nhất và đáp ứng đầu vào không thứ hai, hoặc phụ thuộc vào đáp ứng đầu vào không được kết hợp, để thu được sự chuyển tiếp mượt giữa thông tin âm thanh được giải mã thứ nhất và thông tin âm thanh được giải mã thứ hai được cải biên.

Phương án này theo sáng chế dựa trên việc tìm kiếm sự chuyển tiếp mượt giữa khung âm thanh được mã hóa trong miền dự báo tuyến tính và khung âm thanh tiếp theo được mã hóa trong miền tần số (hoặc, thông thường, trong miền biến đổi) có thể thu được bằng cách cải biên thông tin âm thanh được giải mã thứ hai trên cơ sở tín hiệu mà là đáp ứng đầu vào không của bộ lọc dự báo tuyến tính, trạng thái khởi tạo mà cả hai được xác định bởi thông tin âm thanh được giải mã thứ nhất và thông tin âm thanh được giải mã thứ hai. Tín hiệu đầu ra của bộ lọc dự báo tuyến tính như vậy có thể được sử dụng để thích ứng với thông tin âm thanh được giải mã thứ hai (ví dụ, phần khởi tạo của thông tin âm thanh được giải mã thứ hai, mà ngay sau chuyển tiếp

giữa khung âm thanh thứ nhất và khung âm thanh thứ hai), sao cho có sự chuyển tiếp mượt giữa thông tin âm thanh được giải mã thứ nhất (được kết hợp với khung âm thanh được mã hóa trong miền dự báo tuyến tính) và thông tin âm thanh được giải mã thứ hai được biến đổi (được kết hợp với khung âm thanh được mã hóa trong miền tần số hoặc miền biến đổi) mà không cần biến đổi thông tin âm thanh được giải mã thứ nhất.

Các tác giả sáng chế đã tìm ra rằng đáp ứng đầu vào không của bộ lọc dự báo tuyến tính rất phù hợp cho việc cung cấp sự chuyển tiếp mượt vì trạng thái khởi tạo của bộ lọc dự báo tuyến tính dựa trên cơ sở của cả thông tin âm thanh được giải mã thứ nhất và thông tin âm thanh được giải mã thứ hai, trong đó rằng cưa có trong thông tin âm thanh được giải mã thứ hai được bù bởi rằng cưa giả, mà trở thành phiên bản được cải biên của thông tin âm thanh được giải mã thứ nhất.

Ngoài ra, các tác giả sáng chế đã tìm ra rằng không có độ trễ giải mã do việc cải biên thông tin âm thanh được giải mã thứ hai trên cơ sở đáp ứng đầu vào không thứ nhất và đáp ứng đầu vào không thứ hai, hoặc phụ thuộc vào đáp ứng đầu vào không được kết hợp, trong khi thông tin âm thanh được giải mã thứ nhất không thay đổi, vì đáp ứng đầu vào không thứ nhất và đáp ứng đầu vào không thứ hai hoặc đáp ứng đầu vào không được kết hợp thích ứng rất tốt để làm mượt sự chuyển tiếp giữa khung âm thanh được mã hóa trong miền dự báo tuyến tính và khung âm thanh tiếp theo được mã hóa trong miền tần số (hoặc miền biến đổi) mà không thay đổi thông tin âm thanh được giải mã thứ nhất, vì đáp ứng đầu vào không thứ nhất và đáp ứng đầu vào không thứ hai hoặc đáp ứng đầu vào không được kết hợp cải biên thông tin âm thanh được giải mã thứ hai sao cho thông tin âm thanh được giải mã thứ hai về cơ bản là giống với thông tin âm thanh được giải mã thứ nhất ít nhất là tại chuyển tiếp giữa khung âm thanh được mã hóa trong miền dự báo tuyến tính và khung âm thanh tiếp theo được mã hóa trong miền tần số.

Tóm lại, phương án được mô tả ở trên theo sáng chế cho phép cung cấp sự chuyển tiếp mượt giữa khung âm thanh được mã hóa trong miền mã hóa dự báo tuyến tính và khung âm thanh tiếp theo được mã hóa trong miền tần số (hoặc miền biến đổi), trong đó việc đưa ra độ trễ bổ sung được tránh vì chỉ thông tin âm thanh được giải mã

thứ hai (được kết hợp với khung âm thanh tiếp theo được mã hóa trong miền tần số) được cải biên, và trong đó chất lượng tốt của chuyển tiếp (không có các phần giả tiếp theo) có thể đạt được nhờ sử dụng đáp ứng đầu vào không thứ nhất và đáp ứng đầu vào không thứ hai, hoặc đáp ứng đầu vào không được kết hợp, mà thu được trong việc xét đến cả hai thông tin âm thanh được giải mã thứ nhất và thông tin âm thanh được giải mã thứ hai.

Trong phương án được ưu tiên, bộ giải mã miền tần số được tạo cấu hình để thực hiện biến đổi chồng nghịch đảo, sao cho thông tin âm thanh được giải mã thứ hai chứa răng cưa. Các tác giả sáng chế đã tìm ra rằng các khái niệm theo sáng chế ở trên hoạt động riêng biệt tốt ngay cả trong trường hợp mà bộ giải mã miền tần số (hoặc bộ giải mã miền biến đổi) tạo ra răng cưa. Các tác giả sáng chế đã tìm ra rằng răng cưa có thể được xóa bỏ với một nỗ lực vừa phải và có kết quả tốt bằng việc cung cấp răng cưa giả trong phiên bản được biến đổi của thông tin âm thanh được giải mã thứ nhất.

Trong phương án được ưu tiên, bộ giải mã miền tần số được tạo cấu hình để thực hiện biến đổi chồng nghịch đảo, sao cho thông tin âm thanh được giải mã thứ hai chứa răng cưa trong phần thời gian mà là chồng lấp theo thời gian với phần thời gian cho bộ giải mã miền dự báo tuyến tính mà cung cấp thông tin âm thanh được giải mã thứ nhất, và sao cho thông tin âm thanh được giải mã thứ hai không là răng cưa cho phần thời gian tiếp theo phần thời gian cho bộ giải mã miền dự báo tuyến tính mà cung cấp thông tin âm thanh được giải mã thứ nhất. Phương án này theo sáng chế dựa trên ý tưởng rằng sẽ có lợi khi sử dụng biến đổi chồng (hoặc biến đổi chồng nghịch đảo) và việc tạo cửa sổ mà giữ phần thời gian, mà không cho thông tin âm thanh được giải mã thứ nhất được cung cấp, không răng cưa. Các tác giả sáng chế đã tìm ra rằng đáp ứng đầu vào không thứ nhất và đáp ứng đầu vào không thứ hai hoặc đáp ứng đầu vào không được kết hợp có thể được cung cấp với nỗ lực tính toán nhỏ nếu không cần thiết cung cấp thông tin xóa bỏ răng cưa cho thời gian mà không có thông tin âm thanh được giải mã thứ nhất. Nói cách khác, tốt hơn là cung cấp đáp ứng đầu vào không thứ nhất và đáp ứng đầu vào không thứ hai hoặc đáp ứng đầu vào không được kết hợp, trên cơ sở trạng thái khởi tạo mà trong đó trạng thái khởi tạo răng cưa về cơ bản bị xóa bỏ (ví dụ, sử dụng răng cưa giả). Do đó, đáp ứng đầu vào không thứ nhất và đáp ứng đầu vào không thứ hai hoặc đáp ứng đầu vào không được kết hợp về cơ bản là không

có răng cưa, sao cho không có răng cưa trong thông tin âm thanh được giải mã thứ hai cho khoảng thời gian tiếp theo khoảng thời gian mà cho bộ giải mã miền dự báo tuyến tính cung cấp thông tin âm thanh được giải mã thứ nhất. Về vấn đề này, cần lưu ý rằng đáp ứng đầu vào không thứ nhất và đáp ứng đầu vào không thứ hai, hoặc đáp ứng đầu vào không được kết hợp được cung cấp diễn hình cho khoảng thời gian đã nêu tiếp theo khoảng thời gian mà cho bộ giải mã miền dự báo tuyến tính cung cấp thông tin âm thanh được giải mã thứ nhất (vì đáp ứng đầu vào không thứ nhất và đáp ứng đầu vào không thứ hai hoặc đáp ứng đầu vào không được kết hợp, về cơ bản là sự liên tục suy giảm của thông tin âm thanh được giải mã thứ nhất, xem xét đến thông tin âm thanh được giải mã thứ hai và, một cách điển hình, răng cưa giả mà bù cho răng cưa có trong thông tin âm thanh được giải mã thứ hai cho khoảng thời gian "chồng lấp".

Trong phương án được ưu tiên, phần thông tin âm thanh được giải mã thứ hai, mà được sử dụng để thu được phiên bản được cải biên của thông tin âm thanh được giải mã thứ nhất chứa răng cưa. Bằng cách cho phép một số răng cưa trong thông tin âm thanh được giải mã thứ hai, việc tạo cửa sổ có thể giữ lại một cách đơn giản và có thể tránh được việc tăng quá mức thông tin cần để mã hóa khung âm thanh được mã hóa trong miền tần số. Răng cưa mà có trong một phần thông tin âm thanh được giải mã thứ hai mà được sử dụng để thu được phiên bản được cải biên của thông tin âm thanh được giải mã thứ nhất có thể được bù bởi răng cưa giả nêu trên, sao cho không có sự giảm mạnh về chất lượng âm thanh.

Trong phương án được ưu tiên, răng cưa giả mà được sử dụng để thu được phiên bản cải biên của thông tin âm thanh được giải mã thứ nhất, ít nhất là bù một phần răng cưa mà có trong phần thông tin âm thanh được giải mã thứ hai mà được sử dụng để thu được phiên bản được cải biên của thông tin âm thanh được giải mã thứ nhất. Theo đó, có thể thu được chất lượng âm thanh tốt.

Trong phương án được ưu tiên, bộ xử lý chuyển tiếp được tạo cấu hình để áp dụng việc tạo cửa sổ thứ nhất cho thông tin âm thanh được giải mã thứ nhất, để thu được phiên bản được tạo cửa sổ của thông tin âm thanh được giải mã thứ nhất, và để áp dụng việc tạo cửa sổ thứ hai cho phiên bản thời gian được phản chiếu của thông tin âm thanh được giải mã thứ nhất, để thu được phiên bản được tạo cửa sổ của phiên bản

thời gian được phản chiếu của thông tin âm thanh được giải mã thứ nhất. Trong trường hợp này, bộ xử lý chuyển tiếp có thể được tạo cấu hình để kết hợp phiên bản được tạo cửa sổ của thông tin âm thanh được giải mã thứ nhất và phiên bản được tạo cửa sổ của phiên bản thời gian được phản chiếu của thông tin âm thanh được giải mã thứ nhất, để thu được phiên bản được cải biên của thông tin âm thanh được giải mã thứ nhất. Phương án này theo sáng chế dựa trên ý tưởng rằng một số phép tạo cửa sổ cần được áp dụng để xóa bỏ răng cưa thích hợp trong phiên bản cải biên của thông tin âm thanh được giải mã thứ nhất, mà được sử dụng như đầu vào cho sự cung cấp của đáp ứng đầu vào không. Theo đó, có thể đạt được đáp ứng đầu vào không (ví dụ, đáp ứng đầu vào không thứ hai hoặc đáp ứng đầu vào không được kết hợp) là rất phù hợp cho việc làm mượt sự chuyển tiếp giữa thông tin âm thanh được mã hóa trong miền mã hóa dự báo tuyến tính và khung âm thanh tiếp theo được mã hóa trong miền tần số.

Trong phương án được ưu tiên, bộ xử lý chuyển tiếp được tạo cấu hình để tổ hợp tuyến tính thông tin âm thanh được giải mã thứ hai với đáp ứng đầu vào không thứ nhất và đáp ứng đầu vào không thứ hai hoặc với đáp ứng đầu vào không được kết hợp, cho phần thời gian mà thông tin âm thanh được giải mã thứ nhất không được cung cấp bởi bộ giải mã miền dự báo tuyến tính, để thu được thông tin âm thanh được giải mã thứ hai được cải biên. Các tác giả sáng chế đã tìm ra rằng phép tổ hợp tuyến tính đơn giản (ví dụ, cộng và/hoặc trừ đơn giản, hoặc tổ hợp tuyến tính được gán trọng số hoặc tổ hợp tuyến tính nhiều xuyên kênh) rất phù hợp cho việc tạo ra sự chuyển tiếp mượt.

Trong phương án được ưu tiên, bộ xử lý chuyển tiếp được tạo cấu hình để thông tin âm thanh được giải mã thứ nhất không bị thay đổi bởi thông tin âm thanh được giải mã thứ hai khi cung cấp thông tin âm thanh được giải mã cho khung âm thanh được mã hóa trong miền dự báo tuyến tính, sao cho thông tin âm thanh được giải mã được cung cấp cho khung âm thanh được mã hóa trong miền dự báo tuyến tính được cung cấp độc lập từ thông tin âm thanh được giải mã được cung cấp cho khung âm thanh tiếp theo được mã hóa trong miền tần số. Các tác giả sáng chế đã tìm ra rằng khái niệm theo sáng chế không cần thay đổi thông tin âm thanh được giải mã thứ nhất trên cơ sở thông tin âm thanh được giải mã thứ hai để thu được sự chuyển tiếp đủ mượt. Do đó, bằng cách để thông tin âm thanh được giải mã thứ nhất không bị thay đổi bởi thông tin âm thanh được giải mã thứ hai, có thể tránh được độ trễ, vì thông tin âm thanh được

giải mã thứ nhất theo đó có thể được cung cấp để kết xuất (ví dụ, đến người nghe) ngay cả trước khi giải mã thông tin âm thanh được giải mã thứ hai (được kết hợp với khung âm thanh tiếp theo được mã hóa trong miền tần số) được hoàn thành. Ngược lại, đáp ứng đầu vào không (đáp ứng đầu vào không thứ nhất và thứ hai, hoặc đáp ứng đầu vào không được kết hợp) có thể được tính toán ngay khi thông tin âm thanh được giải mã thứ hai có sẵn. Do đó, độ trễ có thể tránh được.

Trong phương án được ưu tiên, bộ giải mã âm thanh được tạo cấu hình để cung cấp thông tin âm thanh được giải mã hoàn toàn cho khung âm thanh được mã hóa trong miền dự báo tuyến tính, mà theo sau là khung âm thanh được mã hóa trong miền tần số, trước khi giải mã (hoặc trước khi hoàn thành giải mã) khung âm thanh được mã hóa trong miền tần số. Khái niệm này là có khả năng do thực tế là thông tin âm thanh được giải mã thứ nhất không được cải biên trên cơ sở thông tin âm thanh được giải mã thứ hai và giúp cho tránh được độ trễ bất kỳ.

Trong phương án được ưu tiên, bộ xử lý chuyển tiếp được tạo cấu hình để tạo cửa sổ đáp ứng đầu vào không thứ nhất và đáp ứng đầu vào không thứ hai, hoặc đáp ứng đầu vào không được kết hợp, trước khi cải biên thông tin âm thanh được giải mã thứ hai phụ thuộc vào đáp ứng đầu vào không thứ nhất được tạo cửa sổ và đáp ứng đầu vào không thứ hai được tạo cửa sổ hoặc phụ thuộc vào đáp ứng đầu vào không được kết hợp được tạo cửa sổ. Theo đó, sự chuyển tiếp có thể làm nhẵn từng phần. Ngoài ra, có thể tránh được các vấn đề mà sẽ phát sinh từ đáp ứng đầu vào không rất dài.

Trong phương án được ưu tiên, bộ xử lý chuyển tiếp được tạo cấu hình để tạo cửa sổ đáp ứng đầu vào không thứ nhất và đáp ứng đầu vào không thứ hai hoặc đáp ứng đầu vào không được kết hợp, sử dụng cửa sổ tuyến tính. Các tác giả sáng chế đã tìm ra rằng việc sử dụng cửa sổ tuyến tính là khái niệm đơn giản tuy nhiên mang lại ấn tượng nghe tốt.

Phương án theo sáng chế tạo ra phương pháp cung cấp thông tin âm thanh được giải mã trên cơ sở thông tin âm thanh được mã hóa. Phương pháp bao gồm thực hiện giải mã miền dự báo tuyến tính để cung cấp thông tin âm thanh được giải mã thứ nhất trên cơ sở của khung âm thanh được mã hóa trong miền dự báo tuyến tính. Phương pháp bao gồm việc thực hiện giải mã miền tần số để cung cấp thông tin âm thanh được

giải mã thứ hai trên cơ sở của khung âm thanh được mã hóa trong miền tần số. Phương pháp cũng bao gồm bước thu đáp ứng đầu vào không thứ nhất của bộ lọc dự báo tuyến tính đáp ứng với trạng thái khởi tạo thứ nhất của phép lọc dự báo tuyến tính được xác định bởi thông tin âm thanh được giải mã thứ nhất, và để thu được đáp ứng đầu vào không thứ hai của phép lọc dự báo tuyến tính đáp ứng với trạng thái khởi tạo thứ hai của phép lọc dự báo tuyến tính được xác định bởi phiên bản được cải biên của thông tin âm thanh được giải mã thứ nhất, mà được cung cấp với rằng cưa giả, mà bao gồm sự đóng góp của một phần thông tin âm thanh được giải mã thứ hai. Theo cách thay thế, phương pháp bao gồm bước thu đáp ứng đầu vào không được kết hợp của phép lọc dự báo tuyến tính đáp ứng với trạng thái khởi tạo của phép lọc dự báo tuyến tính được xác định bởi phép kết hợp thông tin âm thanh được giải mã thứ nhất và phiên bản được biến đổi của thông tin âm thanh được giải mã thứ nhất, mà được cung cấp với rằng cưa giả, mà bao gồm sự đóng góp của một phần thông tin âm thanh được giải mã thứ hai. Phương pháp còn bao gồm bước cải biên thông tin âm thanh được giải mã thứ hai, mà được cung cấp trên cơ sở khung âm thanh được mã hóa trong miền tần số theo khung âm thanh được mã hóa trong miền dự báo tuyến tính, phụ thuộc vào đáp ứng đầu vào không thứ nhất và đáp ứng đầu vào không thứ hai, hoặc phụ thuộc vào đáp ứng đầu vào không được kết hợp để thu được sự chuyển tiếp mượt giữa thông tin âm thanh được giải mã thứ nhất và thông tin âm thanh được giải mã thứ hai được cải biên. Phương pháp này được dựa trên những xem xét như bộ giải mã âm thanh được mô tả ở trên và mang lại các ưu điểm tương tự.

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.

Phương án khác theo sáng chế tạo ra phương pháp cung cấp thông tin âm thanh được giải mã trên cơ sở thông tin âm thanh được mã hóa. Phương pháp bao gồm bước cung cấp thông tin âm thanh được giải mã thứ nhất trên cơ sở khung âm thanh được mã hóa trong miền dự báo tuyến tính. Phương pháp cũng bao gồm bước cung cấp thông tin âm thanh được giải mã thứ hai trên cơ sở của khung âm thanh được mã hóa trong miền tần số. Phương pháp cũng bao gồm bước thu đáp ứng đầu vào không của phép lọc dự báo tuyến tính, trong đó trạng thái khởi tạo của phép lọc dự báo tuyến tính được xác định phụ thuộc vào thông tin âm thanh được giải mã thứ nhất và thông tin

âm thanh được giải mã thứ hai. Phương pháp còn bao gồm bước cải biên thông tin âm thanh được giải mã thứ hai, mà được cung cấp trên cơ sở của khung âm thanh được mã hóa trong miền tần số theo khung âm thanh được mã hóa trong miền dự báo tuyến tính, phụ thuộc vào đáp ứng đầu vào không, để thu được sự chuyển tiếp mượt giữa thông tin âm thanh được giải mã thứ nhất và thông tin âm thanh được giải mã thứ hai được cải biên.

Phương pháp này dựa trên những cân nhắc tương tự như bộ giải mã âm thanh được mô tả ở trên.

Phương án khác theo sáng chế còn bao gồm chương trình máy tính để thực hiện phương pháp đã nêu.

Mô tả vắn tắt các hình vẽ

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

Fig.1 là hình thể hiện sơ đồ khối dạng giản lược của bộ giải mã âm thanh, theo phương án của sáng chế;

Fig.2 là hình thể hiện sơ đồ khối dạng giản lược của bộ giải mã âm thanh, theo phương án khác của sáng chế;

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

Fig.4a là hình thể hiện sự biểu diễn dạng giản lược của các cửa sổ tại chuyển tiếp từ khung âm thanh được mã hóa MDCT sang khung âm thanh được mã hóa MDCT khác;

Fig.4b là hình thể hiện sự biểu diễn dạng giản lược của cửa sổ tại chuyển tiếp từ khung âm thanh được mã hóa CELP sang khung âm thanh được mã hóa MDCT;

Các hình vẽ Fig.5a, Fig.5b, và Fig.5c là các hình thể hiện sự biểu diễn dạng biểu đồ của các tín hiệu âm thanh trong bộ giải mã âm thanh truyền thống;

Các hình vẽ Fig.6a, Fig.6b, Fig.6c và Fig.6d là các hình thể hiện sự biểu diễn dạng biểu đồ của các tín hiệu âm thanh trong bộ giải mã âm thanh thông thường;

Fig.7a là hình thể hiện sự biểu diễn dạng biểu đồ của tín hiệu âm thanh thu được trên cơ sở của khung CELP trước và của đáp ứng đầu vào không thứ nhất;

Fig.7b là hình thể hiện sự biểu diễn dạng biểu đồ của tín hiệu âm thanh, mà là phiên bản thứ hai của khung CELP trước và của đáp ứng đầu vào không thứ hai;

Fig.7c là hình thể hiện sự biểu diễn dạng biểu đồ của tín hiệu âm thanh mà thu được nếu đáp ứng đầu vào không thứ hai được trừ từ tín hiệu âm thanh của khung MDCT hiện thời;

Fig.8a là hình thể hiện sự biểu diễn dạng biểu đồ của tín hiệu âm thanh thu được trên cơ sở của khung CELP trước;

Fig.8b là hình thể hiện sự biểu diễn dạng biểu đồ của tín hiệu âm thanh, mà thu được như phiên bản thứ hai của khung MDCT hiện thời; và

Fig.8c là hình thể hiện sự biểu diễn dạng biểu đồ của tín hiệu âm thanh, mà là sự kết hợp của tín hiệu âm thanh thu được trên cơ sở của khung CELP trước và của tín hiệu âm thanh mà là phiên bản thứ hai của khung MDCT;

Fig.9 là hình thể hiện lưu đồ của phương pháp cung cấp thông tin âm thanh được giải mã, theo phương án của sáng chế; và

Fig.10 là hình thể hiện lưu đồ của phương pháp cung cấp thông tin âm thanh được giải mã, theo phương án khác của sáng chế.

Mô tả chi tiết sáng chế

Bộ giải mã âm thanh theo Fig.1

Fig.1 là hình thể hiện sơ đồ khối dạng giản lược của bộ giải mã âm thanh 100, theo phương án của sáng chế. Bộ giải mã âm thanh 100 được tạo cấu hình để nhận thông tin âm thanh được mã hóa 110, mà có thể, ví dụ, bao gồm khung thứ nhất được mã hóa trong miền dự báo tuyến tính và khung thứ hai tiếp theo được mã hóa trong miền tần số. Bộ giải mã âm thanh 100 cũng được tạo cấu hình để cung cấp thông tin âm thanh được giải mã 112 trên cơ sở thông tin âm thanh được mã hóa 110.

Bộ giải mã âm thanh 100 bao gồm bộ giải mã miền dự báo tuyến tính 120, mà được tạo cấu hình để cung cấp thông tin âm thanh được giải mã thứ nhất 122 trên cơ

sở khung âm thanh được mã hóa trong miền dự báo tuyến tính. Bộ giải mã âm thanh 100 cũng bao gồm bộ giải mã miền tần số (hoặc bộ giải mã miền biến đổi 130), mà được tạo cấu hình để cung cấp thông tin âm thanh được giải mã thứ hai 132 trên cơ sở khung âm thanh được mã hóa trong miền tần số (hoặc trong miền biến đổi). Ví dụ, bộ giải mã miền dự báo tuyến tính 120 có thể là bộ giải mã CELP, bộ giải mã ACELP, hoặc bộ giải mã tương tự mà thực hiện lọc dự báo tuyến tính trên cơ sở tín hiệu kích thích và trên cơ sở phép biểu diễn được mã hóa của các đặc tính bộ lọc dự báo tuyến tính (hoặc các hệ số bộ lọc).

Bộ giải mã miền tần số 130 có thể, ví dụ, là bộ giải mã dạng AAC hoặc bộ giải mã bất kỳ mà dựa trên việc giải mã dạng AAC. Ví dụ, bộ giải mã miền tần số (hoặc bộ giải mã miền biến đổi) có thể nhận phép biểu diễn được mã hóa của các tham số miền tần số (hoặc các tham số miền biến đổi) và cung cấp thông tin âm thanh được giải mã thứ hai trên cơ sở của nó. Ví dụ, bộ giải mã miền tần số 130 có thể giải mã các hệ số miền tần số (hoặc các hệ số miền biến đổi), định tỉ lệ các hệ số miền tần số (hoặc các hệ số miền biến đổi) phụ thuộc vào các thừa số tỉ lệ (trong đó các thừa số tỉ lệ có thể được cung cấp các các băng tần số khác nhau, và có thể được biểu diễn ở các dạng khác nhau) và thực hiện chuyển đổi miền tần số sang miền thời gian (hoặc chuyển đổi miền biến đổi sang miền thời gian), ví dụ, như Biến đổi Fourier nhanh nghịch đảo hoặc biến đổi cosin rời rạc cải biên nghịch đảo (MDCT nghịch đảo).

Bộ giải mã âm thanh 100 cũng bao gồm bộ xử lý chuyển tiếp 140. Bộ xử lý chuyển tiếp 140 được tạo cấu hình để thu được đáp ứng đầu vào không của việc lọc dự báo tuyến tính, trong đó trạng thái khởi tạo của phép lọc dự báo tuyến tính được xác định phụ thuộc vào thông tin âm thanh được giải mã thứ nhất và thông tin âm thanh được giải mã thứ hai. Bộ xử lý chuyển tiếp 140 cũng được tạo cấu hình để cải biên thông tin âm thanh được giải mã thứ hai 132, mà được cung cấp trên cơ sở khung âm thanh được mã hóa trong miền tần số theo khung âm thanh được mã hóa trong miền dự báo tuyến tính, phụ thuộc vào đáp ứng đầu vào không, để thu được sự chuyển tiếp mượt giữa thông tin âm thanh được giải mã thứ nhất và thông tin âm thanh được giải mã thứ hai được cải biên.

Ví dụ, bộ xử lý chuyển tiếp 140 có thể bao gồm phép xác định trạng thái khởi tạo 144, mà nhận thông tin âm thanh được giải mã thứ nhất 122 và thông tin âm thanh được giải mã thứ hai 132 và trên cơ sở đó cung cấp thông tin trạng thái khởi tạo 146. Bộ xử lý chuyển tiếp 140 cũng bao gồm phép lọc dự báo tuyến tính 148, mà nhận thông tin trạng thái khởi tạo 146 và trên cơ sở đó cung cấp đáp ứng đầu vào không 150. Ví dụ, phép lọc dự báo tuyến tính có thể được thực hiện bởi bộ lọc dự báo tuyến tính, mà được khởi tạo trên cơ sở thông tin trạng thái khởi tạo 146 và được cung cấp với đầu vào không. Theo đó, phép lọc dự báo tuyến tính cung cấp đáp ứng đầu vào không 150. Bộ xử lý chuyển tiếp 140 cũng bao gồm phép cải biên 152, mà cải biên thông tin âm thanh được giải mã thứ hai 132 phụ thuộc vào đáp ứng đầu vào không 150, để từ đó thu được thông tin âm thanh được giải mã thứ hai được cải biên 142, mà góp phần vào thông tin đầu ra của bộ xử lý chuyển tiếp 140. Thông tin âm thanh được giải mã thứ hai được cải biên 142 được ghép nối một cách điển hình với thông tin âm thanh được giải mã thứ nhất 122 để thu được thông tin âm thanh được giải mã 112.

Về chức năng của bộ giải mã âm thanh 100, trường hợp cần được cân nhắc trong đó khung âm thanh được mã hóa trong miền dự báo tuyến tính (khung âm thanh thứ nhất) theo sau là khung âm thanh được mã hóa trong miền tần số (khung âm thanh thứ hai). Khung âm thanh thứ nhất mà được mã hóa trong miền dự báo tuyến tính sẽ được giải mã bởi bộ giải mã miền dự báo tuyến tính 120. Theo đó, thu được thông tin âm thanh được giải mã thứ nhất 122 mà được liên kết với khung âm thanh thứ nhất. Tuy nhiên, thông tin âm thanh được giải mã 122 được liên kết với khung âm thanh thứ nhất không bị tác động một cách đặc trưng bên trái bởi thông tin âm thanh được giải mã trên cơ sở khung âm thanh thứ hai, mà được mã hóa trên miền tần số. Tuy nhiên, thông tin âm thanh được giải mã thứ hai 132 được cung cấp bởi bộ giải mã miền tần số 130 trên cơ sở của khung âm thanh thứ hai mà được mã hóa trong miền tần số.

Không may là thông tin âm thanh được giải mã thứ hai 132 mà được liên kết với khung âm thanh thứ hai thường không bao gồm sự chuyển tiếp mượt với thông tin âm thanh được giải mã thứ nhất 122 mà được liên kết với thông tin âm thanh được giải mã thứ nhất.

Tuy nhiên, cần lưu ý rằng thông tin âm thanh được giải mã thứ hai được cung cấp cho khoảng thời gian mà cũng chồng lấp với khoảng thời gian được liên kết với khung âm thanh thứ nhất. Một phần thông tin âm thanh được giải mã thứ hai mà được cung cấp cho thời gian của khung âm thanh thứ nhất (tức là, phần khởi tạo của thông tin âm thanh được giải mã thứ hai 132) được đánh giá bởi phép xác định trạng thái khởi tạo 144. Ngoài ra, phép xác định trạng thái khởi tạo 144 cũng đánh giá ít nhất một phần thông tin âm thanh được giải mã thứ nhất. Theo đó, phép xác định trạng thái khởi tạo 144 thu được thông tin trạng thái khởi tạo 146 trên cơ sở một phần thông tin âm thanh được giải mã thứ nhất (mà là phần được liên kết với thời gian của khung âm thanh thứ nhất) và trên cơ sở một phần thông tin âm thanh được giải mã thứ hai (mà là một phần thông tin âm thanh được giải mã thứ hai 130 cũng được liên kết với thời gian của khung âm thanh thứ nhất). Theo đó, thông tin trạng thái khởi tạo 146 được cung cấp phụ thuộc vào thông tin được giải mã thứ nhất 132 và cũng phụ thuộc vào thông tin âm thanh được giải mã thứ hai.

Cần lưu ý rằng thông tin trạng thái khởi tạo 146 có thể cung cấp ngay khi thông tin âm thanh được giải mã thứ hai 132 (hoặc ít nhất là phần khởi tạo của nó được yêu cầu bởi phép xác định trạng thái khởi tạo 11) là có sẵn. Phép lọc dự báo tuyến tính 148 có thể cũng được thực hiện ngay sau khi thông tin trạng thái khởi tạo 146 có sẵn, vì phép lọc dự báo tuyến tính sử dụng các hệ số lọc mà đã biết từ việc giải mã khung âm thanh thứ nhất. Theo đó, đáp ứng đầu vào không 150 có thể được cung cấp ngay khi thông tin âm thanh được giải mã thứ hai 132 (hoặc ít nhất là phần khởi tạo của nó được yêu cầu bởi phép xác định giai đoạn khởi tạo 144) là có sẵn. Hơn nữa, đáp ứng đầu vào không 150 có thể được sử dụng để cải biên mà một phần thông tin âm thanh được giải mã thứ hai 132 mà được kết hợp với khung âm thanh thứ hai (hơn là với thời gian của khung âm thanh thứ nhất). Theo đó, phần thông tin âm thanh được giải mã thứ hai, mà thường nằm tại phần bắt đầu của thời gian được liên kết với khung âm thanh thứ hai, được cải biên. Do đó, sự chuyển tiếp mượt giữa thông tin âm thanh được giải mã thứ nhất 122 (mà thường kết thúc tại phần kết thúc của thời gian được liên kết với khung âm thanh thứ nhất) và thông tin âm thanh được giải mã thứ hai được cải biên 142 đạt được (trong đó phần thời gian của thông tin âm thanh được giải mã thứ hai 132 có thời gian mà được liên kết với khung âm thanh thứ nhất tốt hơn được loại

bỏ, và do đó tốt hơn là chỉ được sử dụng cho dự phòng thông tin trạng thái khởi tạo cho phép lọc dự báo tuyến tính). Theo đó, thông tin âm thanh được giải mã tổng thể 112 có thể được cung cấp mà không có độ trễ, vì dự phòng của thông tin âm thanh được giải mã thứ nhất 122 không bị trễ (vì thông tin âm thanh được giải mã thứ nhất 122 là độc lập với thông tin âm thanh được giải mã thứ hai 132), và do thông tin âm thanh được giải mã thứ hai được cải biên 142 có thể được cung cấp ngay khi thông tin âm thanh được giải mã thứ hai 132 có sẵn. Theo đó, các sự chuyển tiếp mượt giữa các khung âm thanh khác có thể đạt được giữa thông tin âm thanh được giải mã 112, ngay cả khi có phép chuyển từ khung âm thanh được mã hóa trong miền dự báo tuyến tính (khung âm thanh thứ nhất) về phía khung âm thanh được mã hóa trong miền tần số (khung âm thanh thứ hai).

Tuy nhiên, lưu ý rằng bộ giải mã âm thanh 100 có thể được bổ sung bởi bất kỳ đặc điểm và chức năng nào được mô tả ở đây.

Bộ giải mã âm thanh theo Fig.2

Fig.2 thể hiện sơ đồ khối dạng giản đồ của bộ giải mã âm thanh, theo phương án khác của sáng chế. Bộ giải mã âm thanh 200 được tạo cấu hình để nhận thông tin âm thanh được mã hóa 210, mà có thể, ví dụ, bao gồm một hoặc nhiều khung được mã hóa trong miền dự báo tuyến tính (hoặc một cách tương đương, trong phép biểu diễn miền dự báo tuyến tính), và một hoặc nhiều khung được mã hóa trong miền tần số (hoặc, một cách tương đương, trong miền biến đổi, hoặc một cách tương đương trong phép biểu diễn miền tần số, hoặc một cách tương đương trong phép biểu diễn miền biến đổi). Bộ giải mã âm thanh 200 được tạo cấu hình để cung cấp thông tin âm thanh được giải mã 212 trên cơ sở thông tin âm thanh được mã hóa 210, trong đó thông tin âm thanh được giải mã 212 có thể, ví dụ, nằm trong phép biểu diễn miền thời gian.

Bộ giải mã âm thanh 200 bao gồm bộ giải mã miền dự báo tuyến tính 220, mà về cơ bản giống với bộ giải mã miền dự báo tuyến tính 120, sao cho áp dụng những giải thích ở trên. Do đó, bộ giải mã miền dự báo tuyến tính 210 nhận các khung âm thanh được mã hóa trong phép biểu diễn miền dự báo tuyến tính mà có trong thông tin âm thanh được mã hóa 210, và cung cấp, trên cơ sở khung âm thanh được mã hóa trong phép biểu diễn miền dự báo tuyến tính, thông tin âm thanh được giải mã thứ nhất 222,

mà thường là ở dạng phép biểu diễn âm thanh miền thời gian (và thường đáp ứng với thông tin âm thanh được giải mã thứ nhất 122). Bộ giải mã âm thanh 200 cũng bao gồm bộ giải mã miền tần số 230, mà về cơ bản giống với bộ giải mã miền tần số 130, sao cho áp dụng những giải thích ở trên. Theo đó, bộ giải mã miền tần số 230 nhận khung âm thanh được mã hóa trong phép biểu diễn miền tần số (hoặc phép biểu diễn miền biến đổi) và trên cơ sở đó cung cấp thông tin âm thanh được giải mã thứ hai 232, mà thường ở dạng phép biểu diễn miền thời gian.

Bộ giải mã âm thanh 200 cũng bao gồm bộ xử lý chuyển tiếp 240, mà được tạo cấu hình để cải biên thông tin âm thanh được giải mã thứ hai 232, để từ đó suy ra thông tin âm thanh được giải mã thứ hai được cải biên 242.

Bộ xử lý chuyển tiếp 240 được tạo cấu hình để thu được đáp ứng đầu vào không của bộ lọc dự báo tuyến tính trong đáp ứng trạng thái khởi tạo của bộ lọc dự báo tuyến tính được xác định bởi thông tin âm thanh được giải mã thứ nhất 222. Bộ xử lý chuyển tiếp được tạo cấu hình để thu được đáp ứng đầu vào không thứ hai của bộ lọc dự báo tuyến tính đáp ứng với trạng thái khởi tạo thứ hai của bộ lọc dự báo tuyến tính được xác định bởi phiên bản được cải biên của thông tin âm thanh được giải mã thứ nhất, mà được cung cấp với răng cưa giả mà bao gồm sự đóng góp của một phần thông tin âm thanh được giải mã thứ hai 232. Ví dụ, bộ xử lý chuyển tiếp 240 bao gồm phép xác định trạng thái khởi tạo 242, mà nhận thông tin âm thanh được giải mã thứ nhất 222 và trên cơ sở đó cung cấp thông tin trạng thái khởi tạo thứ nhất 244. Ví dụ, thông tin trạng thái khởi tạo thứ nhất 244 có thể phản xạ đơn giản một phần thông tin âm thanh được giải mã thứ nhất 222, ví dụ phần mà liền kề với phần kết thúc của phần thời gian được liên kết với khung âm thanh thứ nhất. Bộ xử lý chuyển tiếp 240 có thể cũng bao gồm phép lọc dự báo tuyến tính (thứ nhất) 246, mà được tạo cấu hình để nhận thông tin trạng thái khởi tạo thứ nhất 244 làm trạng thái bộ lọc dự báo tuyến tính khởi tạo và để cung cấp trên cơ sở của thông tin trạng thái khởi tạo thứ nhất 244, đáp ứng đầu vào không thứ nhất 248. Bộ xử lý chuyển tiếp 240 cũng bao gồm việc cải biên/bổ sung răng cưa/kết hợp 250, mà được tạo cấu hình để nhận thông tin âm thanh được giải mã thứ nhất 222, hoặc ít nhất một phần của nó (ví dụ, phần nằm liền kề với phần kết thúc của phần thời gian được liên kết với khung âm thanh thứ nhất), và cũng nhận thông tin được giải mã thứ hai 232, hoặc ít nhất một phần của nó (ví dụ, phần thời gian của

thông tin âm thanh được giải mã thứ hai 232 mà được sắp xếp theo thời gian tại phần kết thúc của phần thời gian được liên kết với khung âm thanh thứ nhất, trong đó thông tin âm thanh được giải mã thứ hai được cung cấp, ví dụ, chủ yếu cho phần thời gian được liên kết với khung âm thanh thứ hai, nhưng cũng đến một số mức độ, cho phần kết thúc của phần thời gian được liên kết với khung âm thanh thứ nhất mà được mã hóa trong phép biểu diễn miền dự báo tuyến tính). Phép cải biên/bổ sung rằng cưa/kết hợp có thể, ví dụ, cải biên phần thời gian của thông tin âm thanh được giải mã thứ nhất, bổ sung rằng cưa giả trên cơ sở của phần thời gian của thông tin âm thanh được giải mã thứ nhất, và cũng bổ sung phần thời gian của thông tin âm thanh được giải mã thứ hai, để từ đó thu được thông tin trạng thái khởi tạo thứ hai 252. Nói cách khác, phép cải biên/bổ sung rằng cưa/kết hợp có thể là một phần của phép xác định trạng thái khởi tạo thứ hai. Thông tin trạng thái khởi tạo thứ hai xác định trạng thái khởi tạo của phép lọc dự báo tuyến tính thứ hai 254, mà được tạo cấu hình để cung cấp đáp ứng đầu vào không thứ hai 256 trên cơ sở thông tin trạng thái khởi tạo thứ hai.

Ví dụ, phép lọc dự báo tuyến tính thứ nhất và phép lọc dự báo tuyến tính thứ hai có thể sử dụng thiết lập bộ lọc (ví dụ, các hệ số bộ lọc) mà được cung cấp bởi bộ giải mã miền dự báo tuyến tính 220 cho khung âm thanh thứ nhất (mà được mã hóa trong phép biểu diễn miền dự báo tuyến tính). Nói cách khác, phép lọc dự báo tuyến tính thứ nhất và thứ hai 246, 254 có thể thực hiện tương tự phép lọc dự báo tuyến tính mà cũng được thực hiện bởi bộ giải mã miền dự báo tuyến tính 220 để thu được thông tin âm thanh được giải mã thứ nhất 222 được liên kết với khung âm thanh thứ nhất. Tuy nhiên, các trạng thái khởi tạo của phép lọc dự báo tuyến tính thứ nhất và thứ hai 246, 254 có thể thiết lập đến các trị số mà được xác định bởi phép xác định trạng thái khởi tạo thứ nhất 244 và bởi phép xác định trạng thái khởi tạo thứ hai 250 (mà chứa các cải biên/bổ sung rằng cưa/kết hợp). Tuy nhiên, tín hiệu đầu vào của các bộ lọc dự báo tuyến tính 246, 254 có thể được đặt thành không. Theo đó, đáp ứng đầu vào không thứ nhất 248 và đáp ứng đầu vào không thứ hai 256 thu được sao cho đáp ứng đầu vào không thứ nhất và đáp ứng đầu vào không thứ hai được dựa trên cơ sở của thông tin âm thanh được giải mã thứ nhất và thông tin âm thanh được giải mã thứ hai, và được tạo hình sử dụng cùng bộ lọc dự báo tuyến tính mà được sử dụng bởi bộ giải mã miền dự báo tuyến tính 220.

Bộ xử lý chuyển tiếp 240 cũng bao gồm phép cải biên 258, mà nhận thông tin âm thanh được mã hóa thứ hai 232 và cải biên thông tin âm thanh được giải mã thứ hai 232 phụ thuộc vào đáp ứng đầu vào không thứ nhất 248 và phụ thuộc vào đáp ứng đầu vào không thứ hai 256, để từ đó thu được thông tin âm thanh được giải mã thứ hai được cải biên 242. Ví dụ, phép cải biên 258 có thể cộng và/hoặc trừ đáp ứng đầu vào không thứ nhất 248 vào hoặc ra khỏi thông tin âm thanh được giải mã thứ hai 232, và có thể cộng hoặc trừ đáp ứng đầu vào không thứ hai 256 vào hoặc ra khỏi thông tin âm thanh được giải mã thứ hai để thu được thông tin âm thanh được giải mã thứ hai được cải biên 242.

Ví dụ, đáp ứng đầu vào không thứ nhất và đáp ứng đầu vào không thứ hai có thể được cung cấp cho khoảng thời gian mà được liên kết với khung âm thanh thứ hai, sao cho chỉ một phần thông tin âm thanh được giải mã thứ hai mà được liên kết với khoảng thời gian của khung âm thanh thứ hai được cải biên. Ngoài ra, các trị số của thông tin âm thanh được giải mã thứ hai 232 mà được liên kết với phần thời gian mà được liên kết với khung âm thanh thứ nhất có thể được loại bỏ trong sự dự phòng cuối cùng của thông tin âm thanh được giải mã thứ hai được cải biên (trên cơ sở của các đáp ứng đầu vào không).

Hơn nữa, bộ giải mã âm thanh 200 tốt hơn là được tạo cấu hình để ghép nối thông tin âm thanh được giải mã thứ nhất 222 và thông tin âm thanh được giải mã thứ hai được cải biên 242, để từ đó thu được thông tin âm thanh được giải mã tổng thể 212.

Về chức năng của bộ giải mã âm thanh 200, viện dẫn đến những giải thích ở trên của bộ giải mã âm thanh 100. Hơn nữa, các chi tiết bổ sung sẽ được mô tả sau đây, viện dẫn đến những hình vẽ khác.

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 tương tự bộ giải mã âm thanh 200, sao cho chỉ những khác biệt mới được mô tả chi tiết. Nếu không thì viện dẫn đến những giải thích đưa ra trước đó đối với bộ giải mã âm thanh 200.

Bộ giải mã âm thanh 300 được tạo cấu hình để nhận thông tin âm thanh được mã hóa 310, mà có thể tương ứng với thông tin âm thanh được mã hóa 210. Hơn nữa, bộ giải mã âm thanh 300 được tạo cấu hình để cung cấp thông tin âm thanh được giải mã 312, mà có thể tương ứng với thông tin âm thanh được giải mã 212.

Bộ giải mã âm thanh 300 bao gồm bộ giải mã miền dự báo tuyến tính 320, mà có thể tương ứng với bộ giải mã miền dự báo tuyến tính 220, và bộ giải mã miền tần số 330, mà tương ứng với bộ giải mã miền tần số 230. Bộ giải mã miền dự báo tuyến tính 320 cung cấp thông tin âm thanh được giải mã thứ nhất 322, ví dụ trên cơ sở khung âm thanh thứ nhất mà được mã hóa trong miền dự báo tuyến tính. Hơn nữa, bộ giải mã âm thanh miền tần số 330 cung cấp thông tin âm thanh được giải mã thứ hai 332, ví dụ trên cơ sở của khung âm thanh thứ hai (mà tiếp theo khung âm thanh thứ nhất) mà được mã hóa trong miền tần số (hoặc trong miền biến đổi). Thông tin âm thanh được giải mã thứ nhất 322 có thể tương ứng với thông tin âm thanh được giải mã thứ nhất 222, và thông tin âm thanh được giải mã thứ hai 332 có thể đáp ứng thông tin âm thanh được giải mã thứ hai 232.

Bộ giải mã âm thanh 300 cũng bao gồm bộ xử lý chuyển tiếp 340, mà có thể tương ứng, trong giới hạn của chức năng tổng thể của nó, với bộ xử lý chuyển tiếp 340, mà có thể cung cấp thông tin âm thanh được giải mã thứ hai được cải biên 342 trên cơ sở của thông tin âm thanh được giải mã thứ hai 332.

Bộ xử lý chuyển tiếp 340 được tạo cấu hình để thu được đáp ứng đầu vào không được kết hợp của bộ lọc dự báo tuyến tính đáp ứng với trạng thái khởi tạo (được kết hợp) của bộ lọc dự báo tuyến tính được xác định bởi tổ hợp của thông tin âm thanh được giải mã thứ nhất và của phiên bản được cải biên của thông tin âm thanh được giải mã thứ nhất, mà được cung cấp với răng cưa giả, mà bao gồm sự đóng góp của một phần thông tin âm thanh được giải mã thứ hai. Hơn nữa, bộ xử lý chuyển tiếp cũng được tạo cấu hình để cải biên thông tin âm thanh được giải mã thứ hai, mà được cung cấp trên cơ sở khung âm thanh được mã hóa trong miền tần số theo khung âm thanh được mã hóa trong miền dự báo tuyến tính, phụ thuộc vào đáp ứng đầu vào không được kết hợp, để thu được sự chuyển tiếp mượt giữa thông tin âm thanh được giải mã thứ nhất và thông tin âm thanh được giải mã thứ hai được cải biên.

Ví dụ, bộ xử lý chuyển tiếp 340 có thể bao gồm phép cải biên/bổ sung rằng cưa/kết hợp 342 mà nhận thông tin âm thanh được giải mã thứ nhất 322 và thông tin âm thanh được giải mã thứ hai 332 và trên cơ sở đó cung cấp thông tin trạng thái khởi tạo 344. Ví dụ, phép cải biên/bổ sung rằng cưa/kết hợp có thể là được xem là phép xác định trạng thái khởi tạo. Cần lưu ý rằng phép cải biên/bổ sung rằng cưa/kết hợp 342 có thể thực hiện chức năng của phép xác định trạng thái khởi tạo 242 và của phép xác định trạng thái khởi tạo 250. Thông tin trạng thái khởi tạo được kết hợp 344 có thể, ví dụ, bằng với (hoặc ít nhất tương ứng với) tổng của thông tin trạng thái khởi tạo thứ nhất 244 và của thông tin trạng thái khởi tạo thứ hai 252. Theo đó, phép cải biên/bổ sung rằng cưa/kết hợp 342 có thể, ví dụ, kết hợp một phần thông tin âm thanh được giải mã thứ nhất 322 với rằng cưa giả và cũng kết hợp với một phần thông tin âm thanh được giải mã thứ hai 332. Hơn nữa, phép cải biên/bổ sung rằng cưa/kết hợp 342 có thể cũng cải biên một phần thông tin âm thanh được giải mã thứ nhất và/hoặc cộng bản sao được tạo của bộ của thông tin âm thanh được giải mã thứ nhất 322, như sẽ được mô tả chi tiết hơn dưới đây. Theo đó, thu được thông tin trạng thái khởi tạo được kết hợp 344.

Bộ xử lý chuyển tiếp 340 cũng bao gồm phép lọc dự báo tuyến tính 346, mà nhận thông tin trạng thái khởi tạo được kết hợp 344 và cung cấp đáp ứng đầu vào không được kết hợp 348 cho phép cải biên 350, trên cơ sở của nó. Phép lọc dự báo tuyến tính 346 có thể, ví dụ, thực hiện phép lọc dự báo tuyến tính mà về cơ bản giống với phép lọc dự báo tuyến tính mà được thực hiện bởi bộ giải mã dự báo tuyến tính 320 để thu được thông tin âm thanh được giải mã thứ nhất 322. Tuy nhiên, trạng thái khởi tạo của phép lọc dự báo tuyến tính 346 có thể được xác định bởi thông tin trạng thái khởi tạo được kết hợp 344. Ngoài ra, tín hiệu đầu vào để cung cấp đáp ứng đầu vào không được kết hợp 348 có thể đặt thành không, sao cho phép lọc dự báo tuyến tính 344 cung cấp đáp ứng đầu vào không trên cơ sở thông tin trạng thái khởi tạo được kết hợp 344 (trong đó các tham số lọc hoặc các hệ số lọc, ví dụ, giống với các tham số lọc hoặc các hệ số lọc được sử dụng bởi bộ giải mã miền dự báo tuyến tính 320 để cung cấp thông tin âm thanh được giải mã thứ nhất 322 được liên kết với khung âm thanh thứ nhất. Hơn nữa, đáp ứng đầu vào không được kết hợp 348 được sử dụng để cải biên thông tin âm thanh được giải mã thứ hai 332, để từ đó suy ra thông tin âm thanh được giải mã

thứ hai được cải biên 342. Ví dụ, phép cải biên 350 có thể cộng đáp ứng đầu vào không được kết hợp 348 vào thông tin âm thanh được giải mã thứ hai 332 hoặc có thể trừ đáp ứng đầu vào không được kết hợp ra khỏi thông tin âm thanh được giải mã thứ hai.

Tuy nhiên, để chi tiết hơn, viện dẫn đến những giải thích của các bộ giải mã âm thanh 100, 200 và cũng viện dẫn đến những giải thích chi tiết sau đây.

Thảo luận về khái niệm chuyển tiếp

Sau đây, một số chi tiết về chuyển tiếp từ khung CELP sang khung MDCT sẽ được mô tả, mà có thể áp dụng được trong các bộ giải mã 100, 200, 300.

Ngoài ra, các sai khác khi so sánh với các khái niệm truyền thống sẽ được mô tả.

Tổng quan-MSCT và phép tạo cửa sổ

Trong các phương án theo sáng chế, vấn đề răng cưa sẽ được giải quyết bằng cách tăng chiều dài MDCT (ví dụ, khung âm thanh được mã hóa trong miền MDCT tiếp theo khung âm thanh được mã hóa trong miền dự báo tuyến tính) sao cho điểm gấp nếp bên trái (ví dụ, của tín hiệu âm thanh miền thời gian được khôi phục trên cơ sở của tập hợp của các hệ số MDCT sử dụng biến đổi MDCT nghịch đảo) được di chuyển tại bên trái của đường biên giữa các khung CELP và MDCT. Phần bên trái của cửa sổ MDCT (ví dụ, của cửa sổ mà được áp dụng cho tín hiệu âm thanh miền thời gian được khôi phục trên cơ sở tập hợp của các hệ số MDCT sử dụng biến đổi MDCT nghịch đảo cũng được thay đổi (ví dụ, khi so sánh với cửa sổ MDCT "thông thường", sao cho sự chồng lấp được giảm.

Như ví dụ, các hình vẽ Fig.4a và Fig.4b thể hiện phép biểu diễn dạng đồ thị của các cửa sổ khác nhau trong đó, Fig.4a thể hiện cửa sổ cho sự chuyển tiếp từ khung MDCT thứ nhất (tức là khung âm thanh thứ nhất được mã hóa trong miền tần số) sang khung MDCT khác (tức là khung âm thanh thứ hai được mã hóa trong miền tần số). Ngược lại, Fig.4b thể hiện cửa sổ mà được sử dụng cho sự chuyển tiếp từ khung CELP (tức là khung âm thanh thứ nhất được mã hóa trong miền dự báo tuyến tính) sang khung MDCT (tức là khung âm thanh thứ hai, tiếp theo, được mã hóa trong miền tần số).

Nói cách khác, Fig.4a thể hiện chuỗi khung âm thanh mà được xem xét như ví dụ so sánh. Ngược lại, Fig.4b thể hiện chuỗi mà khung âm thanh thứ nhất được mã hóa trong miền dự báo tuyến tính và được tiếp theo bởi khung âm thanh thứ hai được mã hóa trong miền tần số, trong đó trong trường hợp theo Fig.4b được xử lý theo cách có lợi riêng biệt bằng các phương án của sáng chế.

Ngay sau đây, viện dẫn đến Fig.4a, cần lưu ý rằng trục hoành 410 mô tả thời gian theo mili giây (milisecond -ms), và trục tung 412 mô tả biên độ của cửa sổ (ví dụ, biên độ được chuẩn hóa của cửa sổ) ở đơn vị bất kỳ. Có thể thấy rằng, chiều dài khung là bằng 20ms, sao cho khoảng thời gian được liên kết với khung âm thanh thứ nhất kéo dài giữa $t = -20\text{ms}$ và $t = 0$. Khoảng thời gian được liên kết với khung âm thanh thứ hai kéo dài từ thời gian $t = 0$ đến $t = 20\text{ms}$. Tuy nhiên, có thể thấy rằng cửa sổ thứ nhất mà tạo cửa sổ các mẫu âm thanh miền thời gian được cung cấp bởi phép biến đổi cosin rời rạc cải biên nghịch đảo trên cơ sở của các hệ số MDCT được giải mã, kéo dài giữa các thời gian $t = -20\text{ms}$ và $t = 8,75\text{ms}$. Do đó, chiều dài của khung thứ nhất 420 dài hơn chiều dài khung (20ms). Theo đó, ngay cả khi thời gian giữa $t = -20\text{ms}$ và $t = 0$ được liên kết với khung âm thanh thứ nhất, các mẫu âm thanh miền thời gian được cung cấp trên cơ sở của việc giải mã khung âm thanh thứ nhất, với thời gian nằm giữa $t = -20\text{ms}$ và $t = 8,75\text{ms}$. Do đó, có sự chồng lấp của xấp xỉ 8,75ms giữa các mẫu âm thanh miền thời gian được cung cấp trên cơ sở của khung âm thanh được mã hóa thứ nhất và các mẫu âm thanh miền thời gian được cung cấp trên cơ sở của khung âm thanh được giải mã thứ hai. Cần lưu ý rằng cửa sổ thứ hai được thiết kế với 422 và kéo dài giữa thời gian $t = 0\text{ms}$ và $t = 28,75\text{ms}$.

Hơn nữa, cần lưu ý rằng các tín hiệu âm thanh miền thời gian được tạo cửa sổ được cung cấp cho khung âm thanh thứ nhất và được cung cấp cho khung âm thanh thứ hai không phải không có răng cưa. Hơn nữa, thông tin âm thanh được giải mã (thứ hai) được tạo cửa sổ được cung cấp cho khung âm thanh thứ nhất chứa răng cưa giữa thời gian từ $t = -20\text{ms}$ và $t = -11,75\text{ms}$, và cũng giữa thời gian từ $t = 0\text{ms}$ và $t = 8,75\text{ms}$. Tương tự, thông tin âm thanh được giải mã được tạo cửa sổ được cung cấp cho khung âm thanh thứ hai chứa răng cưa giữa thời gian từ $t = 0\text{ms}$ và $t = 8,75\text{ms}$, và cũng giữa thời gian từ $t = 20\text{ms}$ và $t = 28,75\text{ms}$. Tuy nhiên, ví dụ, răng cưa có trong thông tin âm thanh được giải mã được cung cấp cho khung âm thanh thứ nhất loại bỏ

ra rằng cửa có trong thông tin âm thanh được giải mã được cung cấp cho khung âm thanh thứ hai tiếp theo trong phần thời gian giữa thời gian $t = 0\text{ms}$ và $t = 8,75\text{ms}$.

Hơn nữa, cần lưu ý rằng các cửa sổ 420 và 422, khoảng thời gian giữa các điểm gấp nếp MDCT là bằng 20ms, mà bằng với chiều dài khung.

Ngay sau đây, viện dẫn đến Fig.4b, trường hợp khác sẽ được mô tả, cụ thể cửa sổ cho sự chuyển tiếp từ khung CELP sang khung MDCT mà có thể được sử dụng trong các bộ giải mã âm thanh 100, 200, 300 để cung cấp thông tin âm thanh được giải mã thứ hai. Trên Fig.4b, trục hoành 430 mô tả thời gian theo mili giây và trục tung 432 mô tả biên độ của cửa sổ trong các đơn vị bất kỳ.

Có thể thấy trên Fig.4b, khung thứ nhất kéo dài giữa thời gian $t_1 = -20\text{ms}$ và thời gian $t_2 = 0\text{ms}$. Do đó, chiều dài khung của khung âm thanh thứ nhất, mà là các khung âm thanh CELP, là 20ms. Hơn nữa, khung âm thanh tiếp theo, thứ hai, kéo dài giữa thời gian t_2 và $t_3 = 20\text{ms}$. Do đó, chiều dài của khung âm thanh thứ hai, mà là khung âm thanh MDCT, cũng là 20ms.

Sau đây, một số chi tiết liên quan cửa sổ 440 sẽ được mô tả.

Cửa sổ 440 bao gồm độ dốc cửa sổ thứ nhất 442, mà kéo dài giữa thời gian $t_4 = -1,25\text{ms}$ và thời gian $t_2 = 0\text{ms}$. Độ dốc cửa sổ thứ hai 440 kéo dài giữa thời gian $t_3 = 20\text{ms}$ và thời gian $t_5 = 28,75\text{ms}$. Cần lưu ý rằng phép biến đổi cosin rời rạc cải biên, mà cung cấp thông tin âm thanh được giải mã (thứ hai) cho khung âm thanh thứ hai (hoặc được liên kết với) cung cấp các mẫu miền thời gian giữa thời gian t_4 và t_5 . Tuy nhiên, phép biến đổi cosin rời rạc cải biên (hoặc một cách chính xác hơn), phép biến đổi cosin rời rạc cải biên nghịch đảo) (mà có thể được sử dụng trong các bộ giải mã miền tần số) 130, 230, 330, nếu khung âm thanh được mã hóa trong miền tần số, ví dụ, miền MDCT, sau khung âm thanh được mã hóa trong miền dự báo tuyến tính) cung cấp các mẫu miền thời gian chứa răng cưa cho thời gian giữa t_4 và t_2 và cho thời gian giữa t_3 và t_5 trên cơ sở của phép biểu diễn miền tần số của khung âm thanh thứ hai. Ngược lại, phép biến đổi cosin rời rạc cải biên nghịch đảo cung cấp các mẫu miền thời gian không có răng cưa cho khoảng thời gian giữa thời gian t_2 và t_3 trên cơ sở của phép biểu diễn miền tần số của khung âm thanh thứ hai. Do đó, độ dốc cửa sổ thứ nhất 442 được liên kết với các mẫu âm thanh miền thời gian chứa một số răng cưa, và độ dốc

cửa sổ thứ hai 444 cũng được liên kết với các mẫu âm thanh miền thời gian chứa một số răng cưa.

Ngoài ra, cần lưu ý rằng thời gian giữa các điểm gấp nếp MDCT là bằng 25ms cho khung âm thanh thứ hai, mà hiểu rằng số hệ số MDCT được mã hóa cần lớn hơn về trạng thái được thể hiện trên Fig.4b so với trạng thái được thể hiện trên Fig.4a.

Tóm lại, các bộ giải mã âm thanh 100, 200, 300 có thể áp dụng các cửa sổ 420, 422 (ví dụ, cho việc tạo cửa sổ đầu ra của phép biến đổi cosin rời rạc cải biên nghịch đảo trong bộ giải mã miền tần số) trong trường hợp cả hai khung âm thanh thứ nhất và khung âm thanh thứ hai tiếp theo khung âm thanh thứ nhất được mã hóa trong miền tần số (ví dụ, trong miền MDCT). Ngược lại, các bộ giải mã âm thanh 100, 200, 300 có thể chuyển hoạt động của bộ giải mã miền tần số trong trường hợp mà khung âm thanh thứ hai, mà tiếp theo khung âm thanh thứ nhất được mã hóa trong miền dự báo tuyến tính, được mã hóa trong miền tần số (ví dụ, trong miền MDCT). Ví dụ, nếu khung âm thanh thứ hai được mã hóa trong miền MDCT và tiếp theo khung âm thanh thứ nhất trước mà được mã hóa trong miền CELP, phép biến đổi cosin rời rạc cải biên nghịch đảo sử dụng số lượng các hệ số MDCT đã tăng có thể được sử dụng (mà hiểu là chứa số lượng các hệ số MDCT đã tăng lên, ở dạng được mã hóa, trong phép biểu diễn miền tần số của khung âm thanh tiếp theo khung âm thanh trước được mã hóa trong miền dự báo tuyến tính, khi so sánh với phép biểu diễn miền tần số của khung âm thanh được mã hóa tiếp theo khung âm thanh trước được mã hóa cũng trong miền tần số). Hơn nữa, cửa sổ khác, cụ thể cửa sổ 440, được áp dụng cho cửa sổ đầu ra của phép biến đổi cosin rời rạc cải biên nghịch đảo (tức là, phép biểu diễn âm thanh miền thời gian được cung cấp bởi phép biến đổi cosin rời rạc cải biên nghịch đảo) để thu được thông tin âm thanh được giải mã thứ hai 132 trong trường hợp mà khung âm thanh (hiện thời) thứ hai trong miền tần số tiếp theo khung âm thanh được mã hóa trong miền dự báo tuyến tính (khi so sánh với trường hợp mà khung âm thanh (hiện thời) thứ hai tiếp theo khung âm thanh trước cũng được mã hóa trong miền tần số).

Để kết luận thêm nữa, phép biến đổi cosin rời rạc cải biên nghịch đảo có chiều dài được tăng lên (khi được so sánh với trường hợp thông thường) có thể được áp dụng bằng bộ giải mã miền tần số 130 trong trường hợp mà khung âm thanh được mã hóa

trong miền tần số sau khung âm thanh được mã hóa trong miền dự báo tuyến tính. Hơn nữa, cửa sổ 440 có thể được sử dụng trong trường hợp (trong khi các cửa sổ 420, 422 có thể được sử dụng trong trường hợp "thông thường" trong đó khung âm thanh được mã hóa trong miền tần số sau khung âm thanh trước được mã hóa trong miền tần số).

Về khái niệm theo sáng chế, cần lưu ý rằng tín hiệu CELP không bị cải biên để không đưa ra bất kỳ độ trễ bổ sung nào, như sẽ được thể hiện chi tiết sau đây. Thay vào đó, các phương án theo sáng chế tạo ra kỹ thuật loại bỏ sự gián đoạn bất kỳ mà có thể được đưa ra tại đường biên giữa các khung CELP và MDCT. Kỹ thuật này làm giảm sự gián đoạn sử dụng đáp ứng đầu vào không của bộ lọc tổng hợp CELP (mà được sử dụng, ví dụ, bởi bộ giải mã miền dự báo tuyến tính). Chi tiết sẽ được đưa ra sau đây.

Tổng quan - mô tả theo từng bước

Sau đây, sự mô tả theo từng bước sẽ được cung cấp. Sau đó, mô tả chi tiết hơn sẽ được đưa ra.

Phía bộ mã hóa

1. Khi khung trước (đôi khi cũng được ký hiệu là "khung thứ nhất") là CELP (hoặc, nói chung, được mã hóa trong miền dự báo tuyến tính), khung MDCT hiện thời (đôi khi cũng được ký hiệu là "khung thứ hai") mà có thể được xem xét như ví dụ của khung được mã hóa trong miền tần số hoặc trong miền biến đổi) được mã hóa với độ dài MDCT khác và cửa sổ MDCT khác. Ví dụ, cửa sổ 440 có thể được sử dụng trong trường hợp này (thay vì cửa sổ "thông thường" 422).

2. Độ dài MDCT tăng (ví dụ từ 20ms đến 25ms, tham khảo các hình vẽ Fig.4a và Fig.4b) sao cho điểm gấp bên trái được di chuyển tại bên trái của đường biên giữa các khung CELP và MDCT. Ví dụ, độ dài MDCT (mà có thể được xác định bởi số lượng các hệ số MDCT) có thể được chọn sao cho độ dài của (hoặc giữa) các điểm gấp MDCT bằng 25 mili giây (như được thể hiện trên Fig.4b) khi được so sánh với độ dài "thông thường" giữa các điểm gấp MDCT bằng 20ms (như được thể hiện trên Fig.4a). Cũng có thể được thấy rằng điểm gấp "bên trái" của biến đổi MDCT nằm giữa các thời gian t_4 và t_2 (hơn là nằm ở giữa giữa $t=0$ và $t = 8,75\text{ms}$), mà có thể được thấy trên

Fig.4b. Tuy nhiên, vị trí của điểm gấp nếp MDCT bên phải có thể không được thay đổi về bên trái (ví dụ, ở giữa giữa các thời gian t_3 và t_5), mà có thể được thấy từ sự so sánh các hình vẽ Fig.4a và 4b (hoặc, chính xác hơn là, của các cửa sổ 422 và 440).

3. Phần bên trái của cửa sổ MDCT được thay đổi sao cho độ dài chồng lấp giảm (ví dụ từ 8,75ms đến 1,25ms). Ví dụ, phần chứa răng cưa nằm giữa các thời gian $t_4 = -1,25\text{ms}$ và $t_2 = 0$ (tức là trước khi khoảng thời gian kết hợp với khung âm thanh thứ hai, mà bắt đầu tại $t=0$ và kết thúc tại $t=20\text{ms}$) trong trường hợp mà các khung âm thanh trước được mã hóa trong miền dự báo tuyến tính. Ngược lại, phần tín hiệu chứa răng cưa nằm giữa các thời gian $t = 0$ và $t = 8,75\text{ms}$ trong trường hợp mà các khung âm thanh trước được mã hóa trong miền tần số (ví dụ, trong miền MDCT).

Phía bộ giải mã

1. Khi khung trước (cũng được ký hiệu là khung âm thanh thứ nhất) là CELP (hoặc, nói chung, được mã hóa trong miền dự báo tuyến tính), khung MDCT hiện thời (cũng được ký hiệu là khung âm thanh thứ hai) (mà là ví dụ cho khung được mã hóa trong miền tần số hoặc miền biến đổi) được mã hóa với các độ dài MDCT giống nhau và cửa sổ MDCT giống nhau như được sử dụng trong phía bộ mã hóa. Nói cách khác, việc tạo cửa sổ được thể hiện trên Fig.4b được áp dụng trong sự cung cấp thông tin âm thanh được giải mã thứ hai, và các đặc tính được đề cập ở trên về biến đổi cosin rời rạc cải biên nghịch đảo (mà tương ứng với các đặc tính của biến đổi cosin rời rạc cải biên được sử dụng tại phía bộ mã hóa) cũng có thể áp dụng.

2. Để loại bỏ sự gián đoạn bất kỳ mà có thể xảy ra tại đường biên giữa các khung CELP và MDCT (ví dụ, tại đường biên giữa khung âm thanh thứ nhất và khung âm thanh thứ hai được đề cập ở trên, cơ chế sau đây được sử dụng:

a) Phần thứ nhất của tín hiệu được bằng cách đưa vào một cách giả răng cưa khuyết của phần chồng lấp của tín hiệu MDCT (ví dụ, của phần tín hiệu giữa các thời gian t_4 và t_2 của tín hiệu âm thanh miền thời gian được cung cấp bởi biến đổi cosin rời rạc cải biên nghịch đảo) sử dụng tín hiệu CELP (ví dụ, sử dụng thông tin âm thanh được giải mã thứ nhất) và thao tác chồng lấp và cộng. Độ dài của phần thứ nhất của tín hiệu là, ví dụ, bằng độ dài chồng lấp (ví dụ, 1,25ms).

b) Phần thứ hai của tín hiệu được xây dựng bằng cách trừ phần thứ nhất của tín hiệu cho tín hiệu CELP tương ứng (phần được bố trí chỉ trước đường biên khung, ví dụ, giữa khung âm thanh thứ nhất và khung âm thanh thứ hai).

c) Đáp ứng đầu vào không của bộ lọc tổng hợp CELP được tạo ra bằng cách lọc khung của các số không và sử dụng phần thứ hai của tín hiệu như các trạng thái bộ nhớ (hoặc như trạng thái khởi tạo).

d) Đáp ứng đầu vào không được, ví dụ, tạo cửa sổ sao cho nó làm giảm các số không sau một số mẫu (ví dụ 64).

e) Đáp ứng đầu vào không được tạo cửa sổ được cộng vào phần bắt đầu của tín hiệu MDCT (ví dụ, phần âm thanh bắt đầu tại thời gian $t_2 = 0$).

Mô tả từng bước - mô tả chi tiết chức năng bộ giải mã

Trong phần sau đây, chức năng của bộ giải mã sẽ được mô tả chi tiết hơn.

Các ký hiệu sau đây sẽ được sử dụng: độ dài khung được ký hiệu là N , tín hiệu CELP được giải mã được ký hiệu là $S_C(n)$, tín hiệu MDCT được giải mã (bao gồm tín hiệu chồng lấp được tạo cửa sổ) được ký hiệu là $S_M(n)$, cửa sổ được sử dụng để tạo cửa sổ phần bên trái của tín hiệu MDCT là $w(n)$ với L là độ dài cửa sổ, và bộ lọc tổng hợp CELP được ký hiệu $\frac{1}{A(z)}$ với $A(z) = \sum_{m=0}^M a_m z^{-m}$ và M là thứ tự bộ lọc.

Sự mô tả chi tiết của bước 1

Sau bước 1 phía bộ giải mã (giải mã khung MDCT hiện thời với cùng độ dài MDCT và cùng cửa sổ MDCT mà được sử dụng trong phía bộ mã hóa) thu được khung MDCT được giải mã hiện thời (ví dụ, phép biểu diễn miền thời gian của “khung âm thanh thứ hai” mà tạo thành thông tin âm thanh được giải mã thứ hai được đề cập ở trên). Khung này (ví dụ, khung thứ hai) không chứa bất kỳ răng cưa nào vì điểm gấp nếp bên trái không được di chuyển ở bên trái của đường biên giữa các khung CELP và MDCT (ví dụ, sử dụng khái niệm như được mô tả trong sự tham khảo chi tiết đến Fig.4b). Điều này có nghĩa là có thể có được sự khôi phục hoàn hảo trong khung hiện thời (ví dụ giữa các thời gian $t_2 = 0$ và $t_3 = 20\text{ms}$) tại tốc độ bit đủ cao. Tuy nhiên, tại

tốc độ bit thấp, tín hiệu không cần thiết phù hợp với tín hiệu đầu vào và do đó sự gián đoạn có thể được đưa vào tại đường biên giữa CELP và MDCT (ví dụ, tại thời gian $t=0$, như được thể hiện trên Fig.4).

Để tạo thuận lợi cho việc hiểu, vấn đề này sẽ được minh họa có sự tham khảo đến Fig.5. Đồ thị phía trên (Fig.5a) thể hiện tín hiệu CELP được giải mã $S_C(n)$, đồ thị ở giữa (Fig.5b) thể hiện tín hiệu MDCT được giải mã (bao gồm tín hiệu chồng lấp được tạo cửa sổ) $S_M(n)$ và đồ thị phía dưới (Fig.5c) thể hiện tín hiệu đầu ra thu được bằng cách loại bỏ tín hiệu chồng lấp được tạo cửa sổ và ghép nối khung CELP và khung MDCT. Có sự gián đoạn rõ ràng trong tín hiệu đầu ra (được thể hiện trên Fig.5c) tại đường biên giữa hai khung (ví dụ, tại thời gian $t=0\text{ms}$).

Ví dụ so sánh của quy trình xử lý thêm

Một giải pháp có khả năng cho vấn đề này là phương thức tiếp cận trong tài liệu tham khảo 1 được đề cập ở trên (“Các cửa sổ nhiễu xuyên kênh hiệu quả cho sự chuyển tiếp giữa mã hóa âm thanh trên cơ sở LPC” bởi J. Lecomte và cộng sự), mà mô tả khái niệm được sử dụng trong MPEG USAC. Trong phần sau đây, sự mô tả vắn tắt của phương thức tiếp cận tham chiếu đã nêu sẽ được cung cấp.

Phiên bản thứ hai, tín hiệu CELP được giải mã $\bar{S}_C(n)$ được khởi tạo thứ nhất bằng với tín hiệu CELP được giải mã

$$\bar{S}_C(n) = S_C(n), n = -N, \dots, -1$$

sau đó rằng cửa khuyết được đưa vào một cách giả trong vùng chồng lấp

$$\bar{S}_C(n) = S_C(n)w(-n-1)w(-n-1) + S_C(-n-L-1)w(n+L)w(-n-1),$$

$$n = -L, \dots, -1$$

cuối cùng, phiên bản thứ hai của tín hiệu CELP được giải mã thu được sử dụng thao tác chồng lấp và cộng.

$$\bar{S}_C(n) = \bar{S}_C(n) + S_M(n), n = -L, \dots, -1$$

Như có thể thấy trên các hình vẽ từ Fig.6a đến Fig.6d, phương thức tiếp cận so sánh này loại bỏ sự gián đoạn (cụ thể xem Fig.6d). Vấn đề với phương thức tiếp cận này là ở chỗ nó đưa vào độ trễ bổ sung (bằng với chiều dài chồng lấp), vì khung đã

qua được cải biên sau khi khung hiện thời đã được giải mã. Trong một số ứng dụng, giống như mã hóa âm thanh độ trễ thấp, được mong muốn (hoặc thậm chí là được yêu cầu) để có độ trễ càng nhỏ càng tốt.

Mô tả chi tiết các bước xử lý

Trái với phương thức tiếp cận thông thường được đề cập ở trên, phương thức tiếp cận được đề xuất ở đây để loại bỏ sự gián đoạn không có bất kỳ độ trễ bổ sung nào. Nó không cải biên khung CELP đã qua (cũng được ký hiệu là khung âm thanh thứ nhất) nhưng thay vì cải biên khung MDCT hiện thời (cũng được ký hiệu là khung âm thanh thứ hai được mã hóa trong miền tần số sau khung âm thanh thứ nhất được mã hóa trong miền dự báo tuyến tính).

Bước a)

Trong bước thứ nhất, "phiên bản thứ hai" của khung ACELP đã qua $\bar{S}_C(n)$ được tính toán giống như được mô tả ở trên. Ví dụ, sự tính toán sau đây có thể được sử dụng:

Phiên bản thứ hai, tín hiệu CELP được giải mã $\bar{S}_C(n)$ được khởi tạo thứ nhất bằng với tín hiệu CELP được giải mã

$$\bar{S}_C(n) = S_C(n), n = -N, \dots, -1$$

sau đó răng cưa khuyết được đưa vào một cách giả trong vùng chồng lấp

$$\bar{S}_C(n) = S_C(n)w(-n-1)w(-n-1) + S_C(-n-L-1)w(n+L)w(-n-1),$$

$$n = -L, \dots, -1$$

cuối cùng, phiên bản thứ hai của tín hiệu CELP được giải mã thu được sử dụng thao tác chồng lấp và cộng

$$\bar{S}_C(n) = \bar{S}_C(n) + S_M(n), n = -L, \dots, -1$$

Tuy nhiên, trái lại với tài liệu tham khảo 1 (“các cửa sổ nhiều xuyên kênh có hiệu quả cho sự chuyển tiếp giữa mã hóa âm thanh dựa trên LPC và không dựa trên LPC” bởi J. Lecomte và cộng sự), tín hiệu ACELP được giải mã sau không được thay thế bởi phiên bản này của khung ACELP đã qua, để không đưa vào bất kỳ độ trễ bổ sung nào.

Nó chỉ được sử dụng làm tín hiệu trung gian cho sự cải biên khung MDCT hiện thời như được mô tả trong các bước tiếp theo.

Nói cách khác, sự xác định trạng thái khởi tạo 144, phép cải biên/bổ sung răng cưa/kết hợp 250 của phép cải biên/bổ sung răng cưa/kết hợp 342 có thể, ví dụ, cung cấp tín hiệu $\widetilde{S}_C(n)$ như sự đóng góp vào thông tin trạng thái khởi tạo 146 hoặc vào thông tin trạng thái khởi tạo được kết hợp 344, hoặc như thông tin trạng thái khởi tạo thứ hai 252. Do đó, sự xác định trạng thái khởi tạo 144, phép cải biên/bổ sung răng cưa/kết hợp 250 hoặc phép cải biên/bổ sung răng cưa/kết hợp 342 có thể, ví dụ, áp dụng việc tạo cửa sổ vào tín hiệu CELP được giải mã S_C (phép nhân với các trị số cửa sổ $w(-n-1)w(-n-1)$), cộng phiên bản thời gian-được phản chiếu của tín hiệu CELP được giải mã ($S_C(-n-L-1)$) được định tỉ lệ với việc tạo cửa sổ ($w(n+L)w(-n-1)$) và cộng tín hiệu MDCT được giả mã $S_M(n)$, để từ đó thu được sự đóng góp vào thông tin trạng thái khởi tạo 146, 344, hoặc thậm chí để thu được thông tin trạng thái khởi tạo thứ hai 252.

Bước b)

Khái niệm cũng bao gồm tạo ra hai tín hiệu bằng cách tính toán đáp ứng đầu vào không (ZIR) của bộ lọc tổng hợp CELP (mà thường có thể được xem xét như bộ lọc dự báo tuyến tính) sử dụng hai bộ nhớ khác nhau (cũng được ký hiệu là các trạng thái khởi tạo) của các bộ lọc tổng hợp CELP.

ZIR thứ nhất $s_Z^1(n)$ được tạo ra bằng cách sử dụng tín hiệu CELP được giải mã trước $S_C(n)$ như các bộ nhớ cho bộ lọc tổng hợp CELP.

$$s_Z^1(n) = S_C(n), n = -L, \dots, -1$$

$$s_Z^1(n) = - \sum_{m=1}^M a_m s_Z^1(n-m), n = 0, \dots, N-1$$

với $M \leq L$

ZIR thứ hai $s_z^2(n)$ được tạo ra bằng cách sử dụng phiên bản thứ hai của tín hiệu CELP được giải mã trước $\bar{s}_c(n)$ như các bộ nhớ cho bộ lọc tổng hợp CELP.

$$s_z^2(n) = \bar{s}_c(n), n = -L, \dots, -1$$

$$s_z^2(n) = - \sum_{m=1}^M a_m s_z^2(n-m), n = 0, \dots, N-1$$

với $M \leq L$

Cần lưu ý là bộ đáp ứng đầu vào không thứ nhất và đáp ứng đầu vào không thứ hai có thể được tính toán một cách riêng rẽ, trong đó đáp ứng đầu vào không thứ nhất có thể thu được trên cơ sở thông tin âm thanh được giải mã thứ nhất (ví dụ, sử dụng việc xác định trạng thái khởi tạo 242 và lọc dự báo tuyến tính 246) và trong đó đáp ứng đầu vào không thứ hai có thể được tính toán, ví dụ, sử dụng phép cải biên/bổ sung răng cưa/kết hợp 250, mà có thể cung cấp “phiên bản thứ hai của khung CELP đã qua $\bar{s}_c(n)$ ” phụ thuộc vào thông tin âm thanh được giải mã thứ nhất 222 và thông tin âm thanh được giải mã thứ hai 232, và cũng sử dụng phép lọc dự báo tuyến tính thứ hai 254. Tuy nhiên, theo cách khác, lọc tổng hợp CELP đơn có thể được áp dụng. Ví dụ, phép lọc dự báo tuyến tính 148, 346 có thể được áp dụng, trong đó tổng của $s_c(n)$ và $\bar{s}_c(n)$ được sử dụng như đầu vào cho lọc dự báo tuyến tính (được kết hợp) đã nêu.

Điều này là do thực tế là lọc dự báo tuyến tính là thao tác tuyến tính, sao cho tổ hợp có thể được thực hiện trước khi lọc hoặc sau khi lọc mà không thay đổi kết quả. Tuy nhiên, phụ thuộc vào các dấu hiệu, chênh lệch giữa $s_c(n)$ và $\bar{s}_c(n)$ cũng có thể được sử dụng làm trạng thái khởi tạo (đối với $n = -L, \dots, -1$) của phép lọc dự báo tuyến tính (được kết hợp).

Tóm lại, thông tin trạng thái khởi tạo thứ nhất $s_z^1(n)$, $n = -L, \dots, -1$ và thông tin trạng thái khởi tạo thứ hai $s_z^2(n)$, $n = -L, \dots, -1$ có thể thu được một cách riêng rẽ hoặc theo cách được kết hợp. Ngoài ra, các đáp ứng đầu vào không có thể thu được bằng cách lọc dự báo tuyến tính riêng rẽ của thông tin trạng thái khởi tạo riêng rẽ,

hoặc sử dụng lọc dự báo tuyến tính (được kết hợp) trên cơ sở thông tin trạng thái khởi tạo được kết hợp.

Như được thể hiện trên đồ thị của Fig.7, mà sẽ được giải thích chi tiết trong phần sau đây, $S_C(n)$ và $s_Z^1(n)$ là liên tục, $\overline{S_C}(n)$ và $s_Z^2(n)$ là liên tục. Hơn nữa, vì $\overline{S_C}(n)$ và $S_M(n)$ cũng liên tục, $S_M(n) - s_Z^2(n)$ là tín hiệu mà bắt đầu từ trị số rất gần với 0.

Tham chiếu đến Fig.7, một số chi tiết sẽ được giải thích.

Fig.7a thể hiện sự biểu diễn dạng đồ thị của khung CELP trước và của đáp ứng đầu vào không thứ nhất. Trục hoành 710 mô tả thời gian theo mili giây và trục tung 712 mô tả biên độ theo các đơn vị tùy ý.

Ví dụ, tín hiệu âm thanh được cung cấp cho khung CELP trước (cũng được ký hiệu là khung âm thanh thứ nhất) được thể hiện giữa các thời gian t_{71} và t_{72} . Ví dụ, tín hiệu $S_C(n)$ với $n < 0$ có thể được thể hiện giữa các thời gian t_{71} và t_{72} . Hơn nữa, đáp ứng đầu vào không có thể được thể hiện giữa các thời gian t_{72} và t_{73} . Ví dụ, đáp ứng đầu vào không thứ nhất $s_Z^1(n)$ có thể được thể hiện giữa các thời gian t_{72} và t_{73} .

Fig.7b thể hiện sự biểu diễn dạng đồ thị của phiên bản thứ hai của khung CELP trước và đáp ứng đầu vào không thứ hai. Trục hoành được ký hiệu là 720, và thể hiện thời gian theo mili giây. Trục tung được ký hiệu là 722 và thể hiện biên độ theo các đơn vị tùy ý. Phiên bản thứ hai của khung CELP trước được thể hiện giữa các thời gian t_{71} (-20ms) và t_{72} (0ms), và đáp ứng đầu vào không thứ hai được thể hiện giữa các thời gian t_{72} và t_{73} (+20ms). Ví dụ, tín hiệu $\overline{S_C}(n)$, $n < 0$, được thể hiện giữa các thời gian t_{71} và t_{72} . Hơn nữa, tín hiệu $s_Z^2(n)$ với $n \geq 0$ được thể hiện giữa các thời gian t_{72} và t_{73} .

Hơn nữa, chênh lệch giữa $S_M(n)$ và $s_Z^2(n)$ được thể hiện trên Fig.7c, trong đó trục hoành 730 ký hiệu thời gian theo mili giây và trong đó trục tung 732 ký hiệu biên độ theo các đơn vị tùy ý.

Hơn nữa, cần được lưu ý rằng đáp ứng đầu vào không $s_Z^1(n)$ với $n \geq 0$ là sự tiếp tục ổn định (một cách cơ bản) của tín hiệu $S_C(n)$ với $n < 0$. Một cách tương tự, đáp

ứng đầu vào không thứ hai $s_z^2(n)$ với $n \geq 0$ là sự tiếp tục ổn định (một cách cơ bản) của tín hiệu $\bar{s}_c(n)$ với $n < 0$.

Bước c)

Tín hiệu MDCT hiện thời (ví dụ, thông tin âm thanh được giải mã thứ hai 132, 232, 332) được thay thế bằng phiên bản thứ hai 142, 242, 342 của MDCT hiện thời (tức là của tín hiệu MDCT kết hợp với khung âm thanh thứ hai, hiện thời).

$$\bar{s}_M(n) = s_M(n) - s_z^2(n) + s_z^1(n)$$

Sau đó, là đơn giản để thể hiện rằng $s_c(n)$ và $\bar{s}_M(n)$ là liên tục: $s_c(n)$ và $s_z^1(n)$ là liên tục, $s_M(n) - s_z^2(n)$ bắt đầu từ trị số rất gần với 0.

Ví dụ, $\bar{s}_M(n)$ có thể được xác định bằng sự cải biên 152, 258, 350 phụ thuộc vào thông tin âm thanh được giải mã thứ hai 132, 232, 323 và phụ thuộc vào đáp ứng đầu vào không thứ nhất $s_z^1(n)$ và đáp ứng đầu vào không thứ hai $s_z^2(n)$ (ví dụ như được thể hiện trên Fig.2), hoặc phụ thuộc vào đáp ứng đầu vào không được kết hợp (ví dụ, đáp ứng đầu vào không được kết hợp $s_z^1(n) - s_z^2(n)$, 150, 348). Như có thể thấy trên đồ thị của Fig.8, phương thức tiếp cận được đề xuất loại bỏ sự gián đoạn.

Ví dụ, Fig.8a thể hiện sự biểu diễn dạng đồ thị của các tín hiệu cho khung CELP trước (ví dụ, của thông tin âm thanh được giải mã thứ nhất), trong đó trục hoành 810 mô tả thời gian theo mili giây và trong đó trục tung 812 mô tả biên độ theo các đơn vị tùy ý.. Như có thể thấy, thông tin âm thanh được giải mã thứ nhất được cung cấp (ví dụ bằng cách giải mã miền dự báo tuyến tính) giữa các thời gian t_{81} (-20ms) và t_{82} (0ms).

Hơn nữa, như có thể thấy trên Fig.8b, phiên bản thứ hai của khung MDCT hiện thời (ví dụ, thông tin âm thanh được giải mã thứ hai được cải biên 142, 242, 342) được cung cấp bắt đầu chỉ từ thời gian t_{82} (0ms), mặc dù thông tin âm thanh được giải mã thứ hai 132, 232, 332 thường được cung cấp bắt đầu từ thời gian t_4 (như được thể hiện trên Fig.4b). Cần lưu ý rằng thông tin âm thanh được giải mã thứ hai 132, 232, 332 được cung cấp giữa các thời gian t_4 và t_2 (như được thể hiện trên Fig.4b) không được

sử dụng trực tiếp cho sự cung cấp của phiên bản thứ hai của khung MDCT hiện thời (tín hiệu $\widetilde{S}_M(n)$) mà chỉ được sử dụng cho sự cung cấp của các thành phần tín hiệu $s_z^2(n)$. Với mục đích làm rõ, cần lưu ý rằng trục hoành 820 ký hiệu thời gian theo mili giây, và trục tung 822 ký hiệu biên độ theo các đơn vị tùy ý.

Fig.8c thể hiện sự ghép nối khung CELP trước (như được thể hiện trên Fig.8a) và phiên bản thứ hai của khung MDCT hiện thời (như được thể hiện trên Fig.8b). Trục hoành 830 mô tả thời gian theo mili giây và trục tung 832 mô tả biên độ theo các đơn vị tùy ý. Như có thể thấy, có sự chuyển tiếp liên tục cơ bản giữa khung CELP trước (giữa các thời gian t_{81} và t_{82} và phiên bản thứ hai của khung MDCT hiện thời (bắt đầu tại thời gian t_{82} và kết thúc, ví dụ, tại thời gian t_5 , như được thể hiện trên Fig.4b). Do đó, các biến dạng có thể nghe được tại sự chuyển tiếp từ khung thứ nhất (mà được mã hóa trong miền dự báo tuyến tính) tới khung thứ hai (mà được mã hóa trong miền tần số) được tránh khỏi.

Cũng là đơn giản để thể hiện rằng sự khôi phục hoàn hảo đạt được tại tốc độ cao: tại tốc độ cao $S_C(n)$ và $\widetilde{S}_C(n)$ là rất giống nhau và cả hai rất giống với tín hiệu đầu vào, sau đó hai ZIR rất giống nhau, do đó chênh lệch của hai ZIR rất gần với 0 và cuối cùng $\widetilde{S}_M(n)$ rất giống với $S_M(n)$ và cả hai rất giống với tín hiệu đầu vào.

Bước d)

Một cách tùy ý, cửa sổ có thể được áp dụng cho hai ZIR, để không ảnh hưởng toàn bộ khung MDCT hiện thời. Điều này là hữu ích ví dụ để làm giảm độ phức tạp, hoặc nếu ZIR không gần với 0 tại phần kết thúc của khung MDCT.

Một ví dụ của cửa sổ là cửa sổ tuyến tính đơn $v(n)$ có độ dài P

$$v(n) = \frac{P-n}{P}, n = 0, \dots, P-1$$

ví dụ với $P = 64$.

Ví dụ, cửa sổ có thể xử lý đáp ứng đầu vào không 150, các đáp ứng đầu vào không 248, 256 hoặc đáp ứng đầu vào không được kết hợp 348.

Phương pháp theo Fig.9

Fig.9 thể hiện lưu đồ của phương pháp cung cấp thông tin âm thanh được giải mã trên cơ sở thông tin âm thanh được mã hóa. Phương pháp 900 bao gồm bước cung cấp thông tin âm thanh được giải mã thứ nhất trên cơ sở khung âm thanh được mã hóa trong miền dự báo tuyến tính. Phương pháp 900 bao gồm bước cung cấp 920 thông tin âm thanh được giải mã thứ hai trên cơ sở khung âm thanh được mã hóa trong miền tần số. Phương pháp 900 cũng bao gồm bước thu 930 đáp ứng đầu vào không của phép lọc dự báo tuyến tính, trong đó trạng thái khởi tạo của phép lọc dự báo tuyến tính được xác định phụ thuộc vào thông tin âm thanh được giải mã thứ nhất và thông tin âm thanh được giải mã thứ hai.

Phương pháp 900 còn bao gồm bước cải biên 940 thông tin âm thanh được giải mã thứ hai, mà được cung cấp trên cơ sở khung âm thanh được mã hóa trong miền tần số sau khung âm thanh được mã hóa trong miền dự báo tuyến tính, phụ thuộc vào đáp ứng đầu vào không, để thu được sự chuyển tiếp mượt giữa thông tin âm thanh được giải mã thứ nhất và thông tin âm thanh được giải mã thứ hai được biến đổi.

Phương pháp 900 có thể được bổ sung bởi dấu hiệu và chức năng bất kỳ trong số các dấu hiệu và các chức năng được mô tả ở đây, cũng đối với các bộ giải mã âm thanh.

Phương pháp theo Fig.10

Fig.10 thể hiện lưu đồ của phương pháp 1000 để cung cấp thông tin âm thanh được giải mã trên cơ sở thông tin âm thanh được mã hóa. T

Phương pháp 1000 bao gồm bước thực hiện 1010 việc giải mã miền dự báo tuyến tính để cung cấp thông tin âm thanh được giải mã thứ nhất trên cơ sở khung âm thanh được mã hóa trong miền dự báo tuyến tính.

Phương pháp 1000 cũng bao gồm bước thực hiện 1020 việc giải mã miền tần số để cung cấp thông tin âm thanh được giải mã thứ hai trên cơ sở khung âm thanh được mã hóa trong miền tần số

Phương pháp 1000 còn bao gồm bước thu 1030 đáp ứng đầu vào không thứ nhất của phép lọc dự báo tuyến tính đáp ứng với trạng thái khởi tạo thứ nhất của phép lọc dự báo tuyến tính được xác định bởi thông tin âm thanh được giải mã thứ nhất và thu được 1040 đáp ứng đầu vào không thứ hai của phép lọc dự báo tuyến tính đáp ứng với trạng thái khởi tạo thứ hai của phép lọc dự báo tuyến tính được xác định bởi phiên bản được cải biên của thông tin âm thanh được giải mã thứ nhất, mà được cung cấp với răng cưa giả, và bao gồm tổ hợp của một phần thông tin âm thanh được giải mã thứ hai.

Nói cách khác, phương pháp 1000 bao gồm bước thu 1050 đáp ứng đầu vào không được kết hợp của phép lọc dự báo tuyến tính đáp ứng với trạng thái khởi tạo của phép lọc dự báo tuyến tính được xác định bởi tổ hợp của thông tin âm thanh được giải mã thứ nhất và của phiên bản được cải biên của thông tin âm thanh được giải mã thứ nhất, mà được cung cấp với răng cưa giả, và bao gồm sự đóng góp của một phần thông tin âm thanh được giải mã thứ hai.

Phương pháp 1000 còn bao gồm bước cải biên 1060 thông tin âm thanh được giải mã thứ hai, mà được cung cấp trên cơ sở khung âm thanh được mã hóa trong miền tần số sau khung âm thanh được mã hóa trong miền dự báo tuyến tính, phụ thuộc vào đáp ứng đầu vào không thứ nhất và đáp ứng đầu vào không thứ hai, hoặc phụ thuộc vào đáp ứng đầu vào không được kết hợp, để thu được sự chuyển tiếp mượt giữa thông tin âm thanh được giải mã thứ nhất và thông tin âm thanh được giải mã thứ hai được biến đổi.

Cần lưu ý rằng phương pháp 1000 có thể được bổ sung bởi dấu hiệu và chức năng bất kỳ trong số các dấu hiệu và các chức năng được mô tả ở đây, cũng đối với các bộ giải mã âm thanh.

6. Kết luận

Để kết luận, các phương án theo sáng chế sẽ đề cập đến sự chuyển tiếp CELP thành MDCT. Các chuyển tiếp này thường có hai vấn đề:

1. Răng cưa do khung MDCT trước khuyết; và

2. Sự gián đoạn tại đường biên giữa khung CELP và khung MDCT, do bản chất mã hóa dạng sóng không hoàn hảo của hai sơ đồ mã hóa hoạt động tại các tốc độ bit trung bình/thấp.

Trong các phương án theo sáng chế, vấn đề răng cưa được giải quyết bằng cách tăng độ dài MDCT sao cho điểm gấp bên trái được di chuyển ở phía bên trái của đường biên giữa các khung CELP và MDCT. Phần bên trái của cửa sổ MDCT cũng được thay đổi sao cho sự chồng lấp được giảm. Trái lại với các giải pháp thông thường, tín hiệu CELP không được cải biên để không đưa vào bất kỳ độ trễ bổ sung nào. Thay vào đó, cơ chế được tạo ra để loại bỏ bất kỳ sự gián đoạn nào mà có thể được đưa vào tại đường biên giữa các khung CELP và MDCT. Cơ chế này làm giảm sự gián đoạn sử dụng đáp ứng đầu vào không của các bộ lọc tổng hợp CELP. Các chi tiết bổ sung được mô tả ở đây.

7. Các lựa chọn thực hiện

Mặc dù một số khía cạnh đã được mô tả trong phạm vi 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 dấu hiệu của bước phương pháp. Tương tự, các khía cạnh được mô tả trong phạm vi 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 số hoặc tất cả các bước phương pháp có thể được thực hiện bởi (hoặc sử dụng) thiết bị phần cứng, như 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 phương tiện 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 phương tiện 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, phương tiện lưu trữ dạng số có thể đọc được bằng máy tính.

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.

Nói chung, các phương án của sáng chế có thể được triển khai như sản phẩm chương trình máy tính với mã chương trình, mã chương trình có hiệu lực để triển khai một trong các phương pháp khi sản phẩm chương trình máy tính chạy trên máy tính. Mã chương trình có thể ví dụ được lưu trữ trên vật mang có thể đọc được bằng máy.

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

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

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

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 Internet.

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

Phương án nữa bao gồm máy tính có chương trình máy tính được cài đặt trên đó để thực hiện một trong số 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 tải (ví dụ, bằng điện tử hoặc quang học) chương trình máy tính để thực hiện một trong số các phương pháp được mô tả ở đây đến bộ nhận. Bộ nhận có thể là, ví dụ, máy tính, thiết bị di động, thiết bị nhớ hoặc tương tự. Thiết bị hoặc hệ thống có thể, ví dụ, bao gồm máy chủ tệp tin để chuyển chương trình máy tính đến bộ nhận.

Theo một số phương án, thiết bị logic lập trình được (ví dụ, mảng cổng 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. Trong một số phương án, mảng cổng lập trình được dạng trường có thể kết hợp với bộ vi xử lý để thực hiện một trong số các phương pháp được mô tả ở đây. Thông thường, các phương pháp 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 bên trên chỉ mang tính minh họa cho các nguyên lý của sáng chế. Cần hiểu rằng các cải biên 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 trung bình trong lĩnh vực kỹ thuật tương ứng. Do đó, mục đích của sáng chế 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 bản mô tả và giải thích của các phương án ở đây.

YÊU CẦU BẢO HỘ

1. Bộ giải mã âm thanh (100;200;300) cung cấp thông tin âm thanh được giải mã (112;212;312) trên cơ sở thông tin âm thanh được mã hóa (110;210;310), bộ giải mã âm thanh bao gồm:

bộ giải mã miền dự báo tuyến tính (120;220;320) được tạo cấu hình để cung cấp thông tin âm thanh được giải mã thứ nhất (122;222;322; $S_C(n)$) trên cơ sở khung âm thanh được mã hóa trong miền dự báo tuyến tính;

bộ giải mã miền tần số (130;230;330) được tạo cấu hình để cung cấp thông tin âm thanh được giải mã thứ hai (132;232;332; $S_M(n)$) trên cơ sở khung âm thanh được mã hóa trong miền tần số; và

bộ xử lý chuyển tiếp (140;240;340),

trong đó bộ xử lý chuyển tiếp được tạo cấu hình để thu được đáp ứng đầu vào không (150; 256;348) của phép lọc dự báo tuyến tính (148; 254; 346), trong đó trạng thái khởi tạo (146;252;344) của phép lọc dự báo tuyến tính được xác định phụ thuộc vào thông tin âm thanh được giải mã thứ nhất và thông tin âm thanh được giải mã thứ hai, và

trong đó bộ xử lý chuyển tiếp được tạo cấu hình để cải biên thông tin âm thanh được giải mã thứ hai (132; 232;332; $S_M(n)$), mà được cung cấp trên cơ sở khung âm thanh được mã hóa trong miền tần số tiếp theo khung âm thanh được mã hóa trong miền dự báo tuyến tính, phụ thuộc vào đáp ứng đầu vào không, để thu được sự chuyển tiếp mượt giữa thông tin âm thanh được giải mã thứ nhất ($S_C(n)$) và thông tin âm thanh được giải mã thứ hai được cải biên ($\tilde{S}_M(n)$).

2. Bộ giải mã âm thanh theo điểm 1,

trong đó bộ xử lý chuyển tiếp được tạo cấu hình để thu được đáp ứng đầu vào không thứ nhất (248; $s_{\frac{1}{2}}(n)$) của bộ lọc dự báo tuyến tính (246) đáp ứng với trạng thái khởi tạo thứ nhất (244; $S_C(n)$) của bộ lọc dự báo tuyến tính được xác định bởi thông tin âm thanh được giải mã thứ nhất (222; $S_C(n)$), và

trong đó bộ xử lý chuyển tiếp được tạo cấu hình để thu được đáp ứng đầu vào không thứ hai $(256; s_z^2(n))$ của bộ lọc dự báo tuyến tính (254) đáp ứng với trạng thái khởi tạo thứ hai (252) của bộ lọc dự báo tuyến tính được xác định bởi phiên bản được cải biên $(\bar{S}_C(n))$ của thông tin âm thanh được giải mã thứ nhất $(222, S_C(n))$, mà được cung cấp với răng cưa giả, và bao gồm sự đóng góp của một phần thông tin âm thanh được giải mã thứ hai $(232, S_M(n))$, hoặc

trong đó bộ xử lý chuyển tiếp được tạo cấu hình để thu được đáp ứng đầu vào không được kết hợp $(150; 348)$ của bộ lọc dự báo tuyến tính $(148; 346)$ đáp ứng với trạng thái khởi tạo $(146; 344)$ của bộ lọc dự báo tuyến tính được xác định bởi tổ hợp của thông tin âm thanh được giải mã thứ nhất $(122; 322; S_C(n))$ và của phiên bản được cải biên $(\bar{S}_C(n))$ của thông tin âm thanh được giải mã thứ nhất $(122; 322; S_C(n))$, mà được cung cấp với răng cưa giả, và bao gồm sự đóng góp của một phần thông tin âm thanh được giải mã thứ hai $(132; 332; S_M(n))$;

trong đó bộ xử lý chuyển tiếp được tạo cấu hình để cải biên thông tin âm thanh được giải mã thứ hai $(132; 232; 332; S_M(n))$, mà được cung cấp trên cơ sở khung âm thanh được mã hóa trong miền tần số tiếp theo khung âm thanh được mã hóa trong miền dự báo tuyến tính, phụ thuộc vào đáp ứng đầu vào không thứ nhất $(248; s_z^1(n))$ và đáp ứng đầu vào không thứ hai $(256; s_z^2(n))$, hoặc phụ thuộc vào đáp ứng đầu vào không được kết hợp $(150; s_z^1(n) - s_z^2(n))$, để thu được sự chuyển tiếp mượt giữa thông tin âm thanh được giải mã thứ nhất $(122; 222; 322; S_C(n))$ và thông tin âm thanh được giải mã thứ hai được cải biên $(142; 242; 342; \bar{S}_M(n))$.

3. Bộ giải mã âm thanh $(100; 200; 300)$ theo điểm 1 hoặc điểm 2, trong đó bộ giải mã miền tần số $(130; 230; 330)$ được tạo cấu hình để thực hiện biến đổi chồng nghịch đảo, sao cho thông tin âm thanh được giải mã thứ hai $(132; 232; 332)$ chứa răng cưa.

4. Bộ giải mã âm thanh $(100; 200; 300)$ theo điểm 1 hoặc điểm 2 hoặc điểm 3, trong đó bộ giải mã miền tần số $(130; 230; 330)$ được tạo cấu hình để thực hiện biến đổi chồng nghịch đảo, sao cho thông tin âm thanh được giải mã thứ hai $(132; 232; 332)$ chứa răng cưa trong phần thời gian mà chồng lấp theo thời gian với phần thời gian mà

bộ giải mã miền dự báo tuyến tính (120;220;320) cung cấp thông tin âm thanh được giải mã thứ nhất (122;222;322), và sao cho thông tin âm thanh được giải mã thứ hai là không có răng cưa cho phần thời gian tiếp theo phần thời gian mà bộ giải mã miền dự báo tuyến tính cung cấp cho thông tin âm thanh được giải mã thứ nhất.

5. Bộ giải mã âm thanh (100;200;300) theo một trong số các điểm từ 1 đến 4, trong đó phần thông tin âm thanh được giải mã thứ hai (132;232;332), mà được sử dụng để thu được phiên bản được cải biên ($\bar{S}_c(n)$) của thông tin âm thanh được giải mã thứ nhất, chứa răng cưa.

6. Bộ giải mã âm thanh (100;200;300) theo điểm 5, trong đó răng cưa giả, mà được sử dụng để thu được phiên bản được cải biên ($\bar{S}_c(n)$) của thông tin âm thanh được giải mã thứ nhất, bù ít nhất một phần răng cưa mà có trong phần thông tin âm thanh được giải mã thứ hai (132;232;332), mà được sử dụng để thu được phiên bản được cải biên của thông tin âm thanh được giải mã thứ nhất.

7. Bộ giải mã âm thanh (100;200;300) theo một trong số các điểm từ 1 đến 6, trong đó bộ xử lý chuyển tiếp (140;240;340) được tạo cấu hình để thu được đáp ứng đầu vào không thứ nhất $s_z^1(n)$, hoặc thành phần thứ nhất $s_z^1(n)$ của đáp ứng đầu vào không được kết hợp, theo

$$s_z^1(n) = - \sum_{m=1}^M a_m s_z^1(n-m), n = 0, \dots, N-1$$

hoặc theo

$$s_z^1(n) = + \sum_{m=1}^M a_m s_z^1(n-m), n = 0, \dots, N-1$$

với

$$s_z^1(n) = S_c(n), n = -L, \dots, -1$$

$$M \leq L$$

trong đó n ký hiệu chỉ số thời gian,

trong đó $s_{\frac{1}{2}}(n)$ với $n=0,\dots,N-1$ ký hiệu đáp ứng đầu vào không thứ nhất (248) cho chỉ số thời gian n , hoặc thành phần thứ nhất của đáp ứng đầu vào không được kết hợp (150;348) cho chỉ số thời gian n ;

trong đó $s_{\frac{1}{2}}(n)$ với $n=-L,\dots,-1$ ký hiệu trạng thái khởi tạo thứ nhất (244) cho chỉ số thời gian n , hoặc thành phần thứ nhất của trạng thái khởi tạo (146;344) cho chỉ số thời gian n ;

trong đó m ký hiệu biến số chạy;

trong đó M ký hiệu độ dài bộ lọc của bộ lọc dự báo tuyến tính;

trong đó a_m ký hiệu các hệ số bộ lọc của bộ lọc dự báo tuyến tính;

trong đó $S_c(n)$ ký hiệu trị số được giải mã trước của thông tin âm thanh được giải mã thứ nhất (122;222;322) đối với chỉ số thời gian n ;

trong đó N ký hiệu độ dài xử lý.

8. Bộ giải mã âm thanh (100;200;300) theo một trong số các điểm từ 1 đến 7, trong đó bộ xử lý chuyển tiếp (140;240;340) được tạo cấu hình để áp dụng việc tạo cửa sổ thứ nhất $((w(-n-1)w(-n-1)))$ cho thông tin âm thanh được giải mã thứ nhất (122;222;322; $S_c(n)$), để thu được phiên bản được tạo cửa sổ của thông tin âm thanh được giải mã thứ nhất, và để áp dụng việc tạo cửa sổ thứ hai $(w(n+L)w(-n-1))$ cho phiên bản thời gian-được phản chiếu ($S_c(-n-L-1)$) của thông tin âm thanh được giải mã thứ nhất (122;222;322; $S_c(n)$), để thu được phiên bản được tạo cửa sổ của phiên bản thời gian-được phản chiếu của thông tin âm thanh được giải mã thứ nhất, và

trong đó bộ xử lý chuyển tiếp được tạo cấu hình để tổ hợp phiên bản được tạo cửa sổ của thông tin âm thanh được giải mã thứ nhất và phiên bản được tạo cửa sổ của phiên bản thời gian-được phản chiếu của thông tin âm thanh được giải mã thứ nhất, để thu được phiên bản được cải biên ($\tilde{S}_c(n)$) của thông tin âm thanh được giải mã thứ nhất.

9. Bộ giải mã âm thanh theo một trong số các điểm từ 1 đến 8, trong đó bộ xử lý chuyển tiếp (140;240;340) được tạo cấu hình để thu được phiên bản được cải biên $\overline{S_c}(n)$ của thông tin âm thanh được giải mã thứ nhất $S_c(n)$ theo

$$\overline{S_c}(n) = S_c(n)w(-n-1)w(-n-1) + S_c(-n-L-1)w(n+L)w(-n-1) + S_M(n),$$

$$n = -L, \dots, -1,$$

trong đó n ký hiệu chỉ số thời gian,

trong đó $w(-n-1)$ ký hiệu trị số cho hàm cửa sổ cho chỉ số thời gian $(-n-1)$;

trong đó $w(n+L)$ ký hiệu trị số của hàm cửa sổ cho chỉ số thời gian $(n+L)$;

trong đó $S_c(n)$ ký hiệu trị số được giải mã trước của thông tin âm thanh được giải mã thứ nhất (122;222;322) cho chỉ số thời gian (n) ;

trong đó $S_c(-n-L-1)$ ký hiệu trị số được giải mã trước của thông tin âm thanh được giải mã thứ nhất cho chỉ số thời gian $(-n-L-1)$;

trong đó $S_M(n)$ ký hiệu trị số được giải mã của thông tin âm thanh được giải mã thứ hai (132;232;332) cho chỉ số thời gian n ; và

trong đó L mô tả độ dài của cửa sổ.

10. Bộ giải mã âm thanh (100;200;300) theo một trong số các điểm từ 1 đến 9, trong đó bộ xử lý chuyển tiếp (140;240;340) được tạo cấu hình để thu được đáp ứng đầu vào không thứ hai ($256; s_Z^2(n)$), hoặc thành phần thứ hai $s_Z^2(n)$ của đáp ứng đầu vào không được kết hợp (150;348) theo

$$s_Z^2(n) = - \sum_{m=1}^M a_m s_Z^2(n-m), n = 0, \dots, N-1$$

hoặc theo

$$s_Z^2(n) = + \sum_{m=1}^M a_m s_Z^2(n-m), n = 0, \dots, N-1$$

với

$$s_z^2(n) = \overline{s_c}(n), n = -L, \dots, -1$$

$$M \leq L$$

trong đó n ký hiệu chỉ số thời gian,

trong đó $s_z^2(n)$ với $n=0, \dots, N-1$ ký hiệu đáp ứng đầu vào không thứ hai cho chỉ số thời gian n , hoặc thành phần thứ hai của đáp ứng đầu vào không được kết hợp cho chỉ số thời gian n ;

trong đó $s_z^2(n)$ với $n=-L, \dots, -1$ ký hiệu trạng thái khởi tạo thứ hai (252) cho chỉ số thời gian n , hoặc thành phần thứ hai của trạng thái khởi tạo (146;344) cho chỉ số thời gian n ;

trong đó m ký hiệu biên số chạy,

trong đó M ký hiệu độ dài bộ lọc của bộ lọc dự báo tuyến tính (148;254;346);

trong đó a_m ký hiệu các hệ số bộ lọc của bộ lọc dự báo tuyến tính;

trong đó $\overline{s_c}(n)$ ký hiệu các trị số của phiên bản được cải biên của thông tin âm thanh được giải mã thứ nhất cho chỉ số thời gian n ;

trong đó N ký hiệu độ dài xử lý.

11. Bộ giải mã âm thanh (100;200;300) theo một trong số các điểm từ 1 đến 10, trong đó bộ xử lý chuyển tiếp (140;240;340) được tạo cấu hình để tổ hợp tuyến tính thông tin âm thanh được giải mã thứ hai (132;232;332) với đáp ứng đầu vào không thứ nhất (248) và đáp ứng đầu vào không thứ hai (256), hoặc đáp ứng đầu vào không được kết hợp (150;348), cho phần thời gian mà không có thông tin âm thanh được giải mã thứ nhất (122;222;322) được cung cấp bởi bộ giải mã miền dự báo tuyến tính (120;220;320), để thu được thông tin âm thanh được giải mã thứ hai được cải biên.

12. Bộ giải mã âm thanh (100;200;300) theo một trong số các điểm từ 1 đến 11, trong đó bộ xử lý chuyển tiếp (140;240;340) được tạo cấu hình để thu được thông tin âm thanh được giải mã thứ hai được cải biên $\bar{S}_M(n)$ theo

$$\bar{S}_M(n) = S_M(n) - s_Z^2(n) + s_Z^1(n), \text{ với } n=0, \dots, N-1,$$

hoặc theo

$$\bar{S}_M(n) = S_M(n) - v(n)s_Z^2(n) + v(n)s_Z^1(n), \text{ với } n=0, \dots, N-1,$$

trong đó n ký hiệu chỉ số thời gian;

trong đó $S_M(n)$ ký hiệu các trị số của thông tin âm thanh được giải mã thứ hai cho chỉ số thời gian n ;

trong đó $s_Z^1(n)$ với $n=0, \dots, N-1$ ký hiệu đáp ứng đầu vào không thứ nhất cho chỉ số thời gian n , hoặc thành phần thứ nhất của đáp ứng đầu vào không được kết hợp cho trị số thời gian n ; và

trong đó $s_Z^2(n)$ với $n=0, \dots, N-1$ ký hiệu đáp ứng đầu vào không thứ hai cho chỉ số thời gian n , hoặc thành phần thứ hai của đáp ứng đầu vào không được kết hợp cho chỉ số thời gian n ;

trong đó $v(n)$ ký hiệu các trị số của hàm cửa sổ;

trong đó N ký hiệu độ dài xử lý.

13. Bộ giải mã âm thanh (100;200;300) theo một trong số các điểm từ 1 đến 12, trong đó bộ xử lý chuyển tiếp (140;240;340) được tạo cấu hình để để lại thông tin âm thanh được giải mã thứ nhất (122;222;322) không bị thay đổi bởi thông tin âm thanh được giải mã thứ hai (132;232;332) khi cung cấp thông tin âm thanh được giải mã cho khung âm thanh được mã hóa trong miền dự báo tuyến tính, sao cho thông tin âm thanh được giải mã được cung cấp cho khung âm thanh được mã hóa trong miền dự báo tuyến tính được cung cấp độc lập với thông tin âm thanh được giải mã được cung cấp cho khung âm thanh đứng sau được mã hóa trong miền tần số.

14. Bộ giải mã âm thanh (100;200;300) theo một trong số các điểm từ 1 đến 13, trong đó bộ giải mã âm thanh được tạo cấu hình để cung cấp thông tin âm thanh được giải mã đầy đủ (122;222;322) cho khung âm thanh được mã hóa trong miền dự báo tuyến tính, mà được tiếp theo bởi khung âm thanh được mã hóa trong miền tần số, trước khi giải mã khung âm thanh được mã hóa trong miền tần số.

15. Bộ giải mã âm thanh (100;200;300) theo một trong số các điểm 1 đến 14, trong đó bộ xử lý chuyển tiếp (140;240;340) được tạo cấu hình để tạo cửa sổ đáp ứng đầu vào không thứ nhất (248) và đáp ứng đầu vào không thứ hai (256), hoặc đáp ứng đầu vào không được kết hợp (150;348), trước khi cải biên thông tin âm thanh được giải mã thứ hai (132;232;332) phụ thuộc vào đáp ứng đầu vào không thứ nhất được tạo cửa sổ và đáp ứng đầu vào không thứ hai được tạo cửa sổ, hoặc phụ thuộc vào đáp ứng đầu vào không được kết hợp được tạo cửa sổ.

16. Bộ giải mã âm thanh (100;200;300) theo điểm 15, trong đó bộ xử lý chuyển tiếp được tạo cấu hình để tạo cửa sổ đáp ứng đầu vào không thứ nhất và đáp ứng đầu vào không thứ hai, hoặc đáp ứng đầu vào không được kết hợp, sử dụng cửa sổ tuyến tính.

17. Phương pháp cung cấp thông tin âm thanh được giải mã (900) trên cơ sở thông tin âm thanh được mã hóa, phương pháp bao gồm các bước:

cung cấp (910) thông tin âm thanh được giải mã thứ nhất ($S_C(n)$) trên cơ sở khung âm thanh được mã hóa trong miền dự báo tuyến tính;

cung cấp (920) thông tin âm thanh được giải mã thứ hai ($S_M(n)$) trên cơ sở khung âm thanh được mã hóa trong miền tần số; và

thu được (930) đáp ứng đầu vào không của phép lọc dự báo tuyến tính, trong đó trạng thái khởi tạo của phép lọc dự báo tuyến tính được xác định phụ thuộc vào thông tin âm thanh được giải mã thứ nhất và thông tin âm thanh được giải mã thứ hai, và

cải biên (940) thông tin âm thanh được giải mã thứ hai ($S_M(n)$), mà được cung cấp trên cơ sở khung âm thanh được mã hóa trong miền tần số tiếp theo khung âm thanh được mã hóa trong miền dự báo tuyến tính, phụ thuộc vào đáp ứng đầu vào

không, để thu được sự chuyển tiếp mượt giữa thông tin âm thanh được giải mã thứ nhất ($S_C(n)$) và thông tin âm thanh được giải mã thứ hai được cải biên ($\bar{S}_M(n)$).

18. Vật ghi đọc được bằng máy tính chứa chương trình máy tính để thực hiện phương pháp theo điểm 17 khi chương trình máy tính chạy trên máy tính.

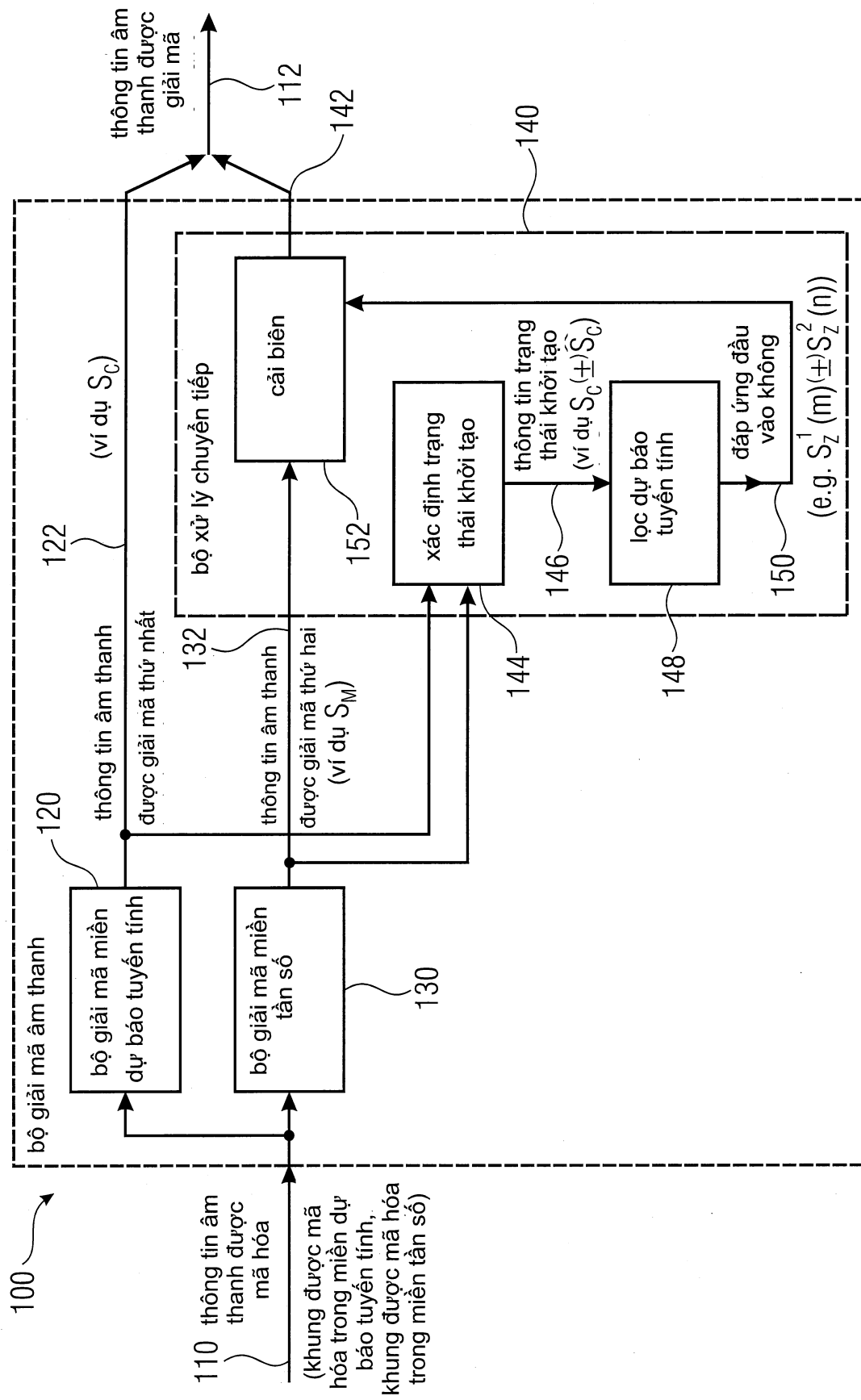


FIG 1

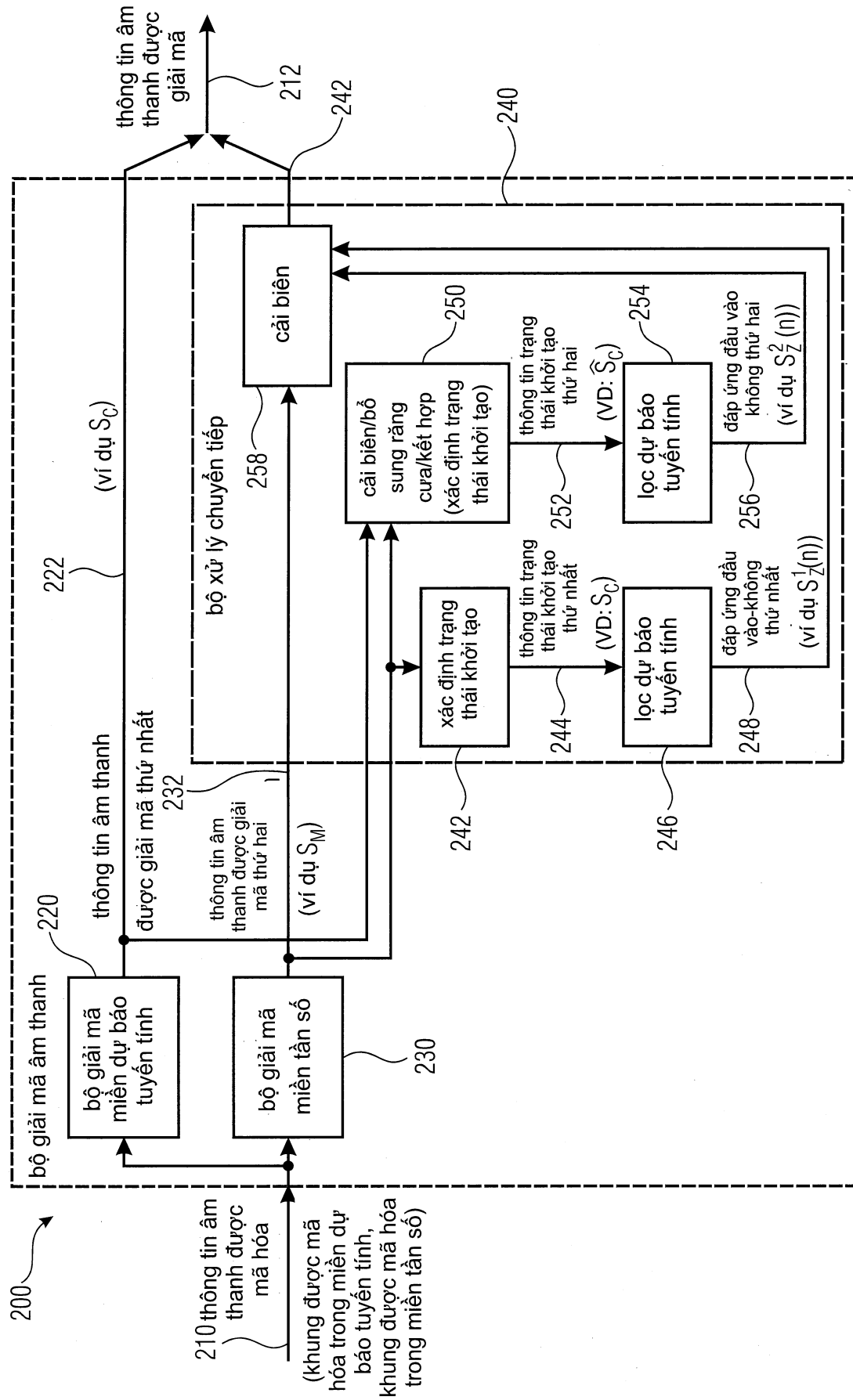


FIG 2

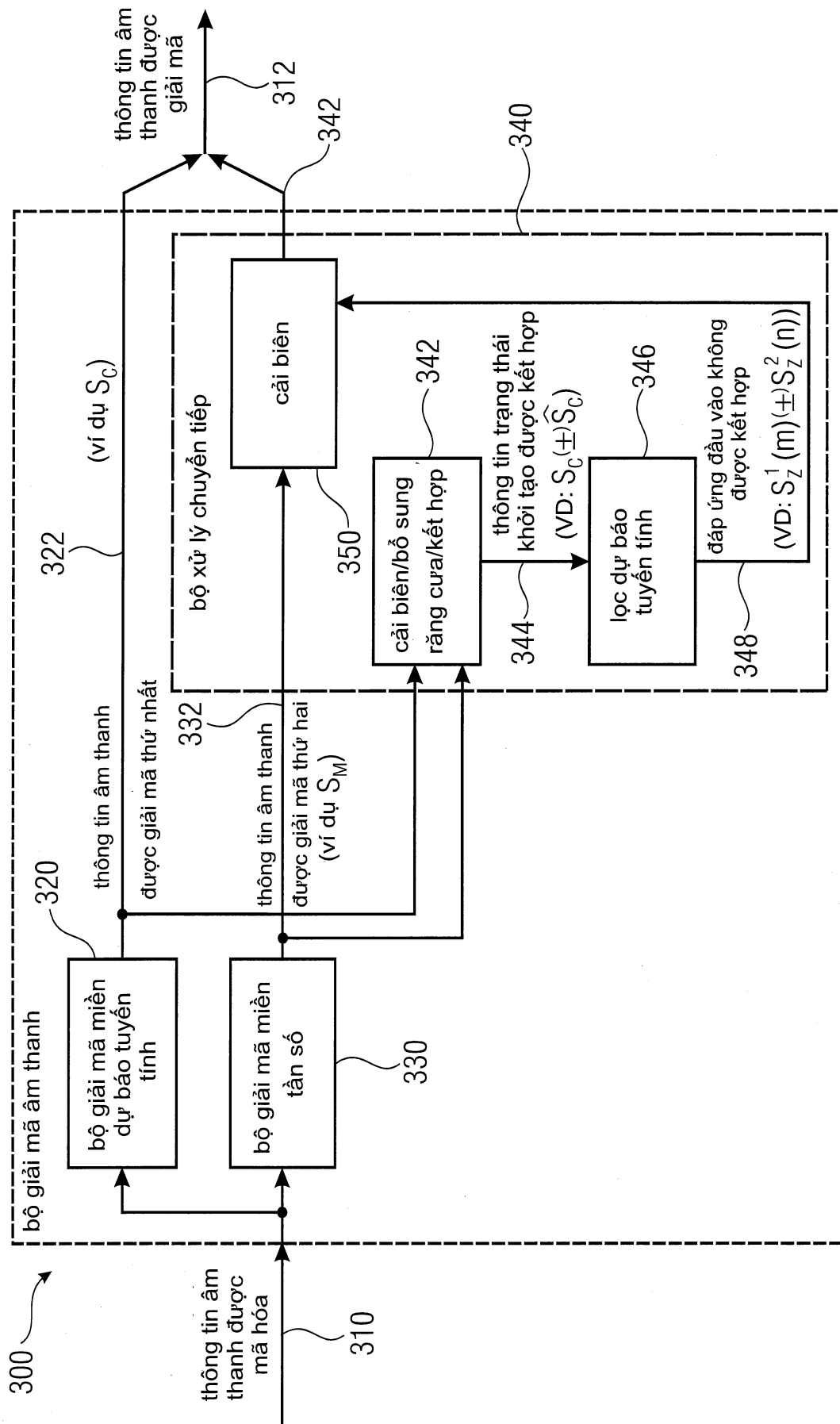
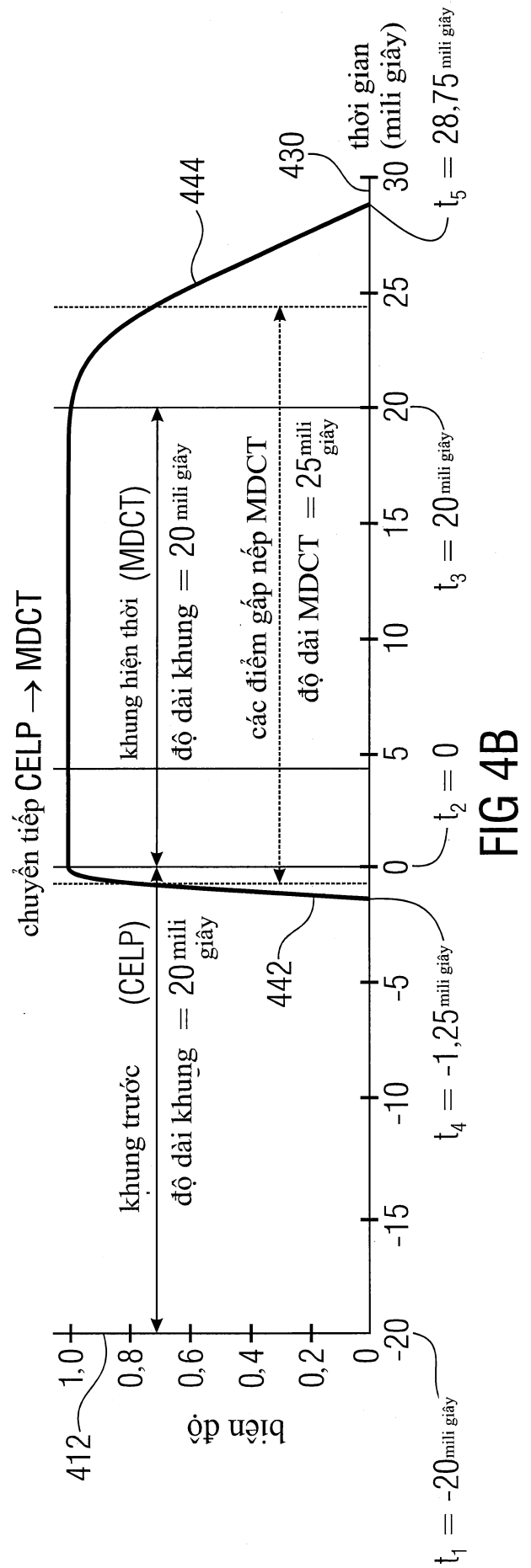
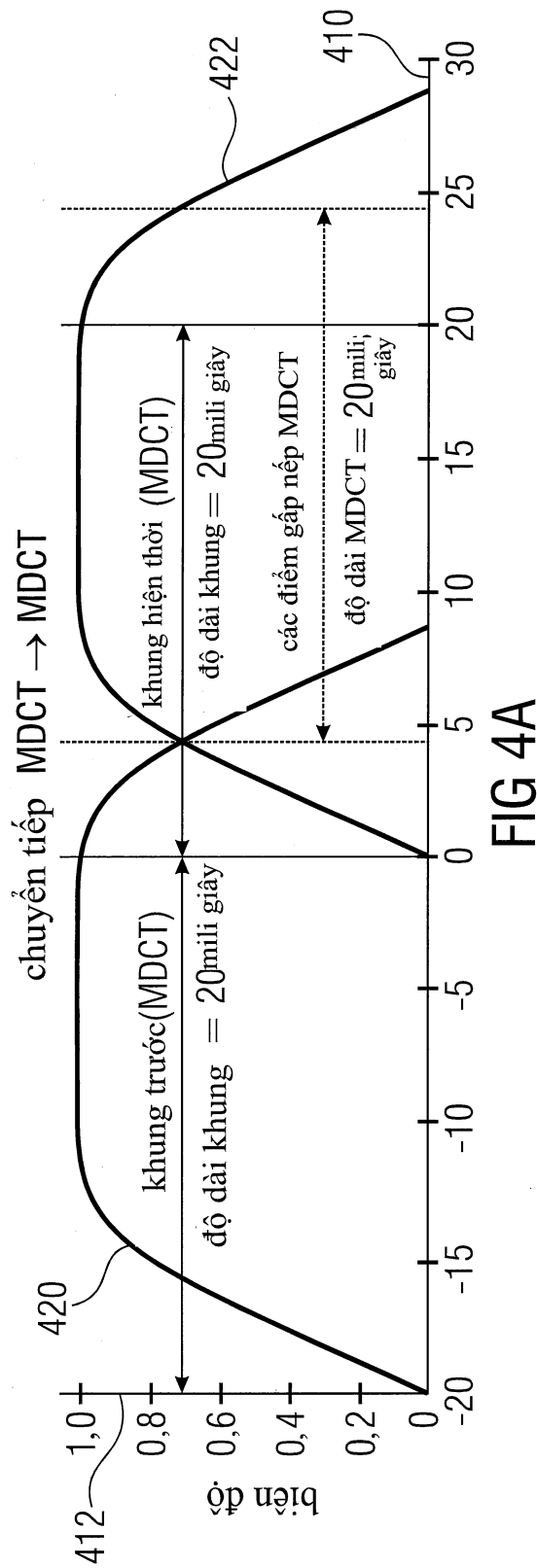


FIG 3



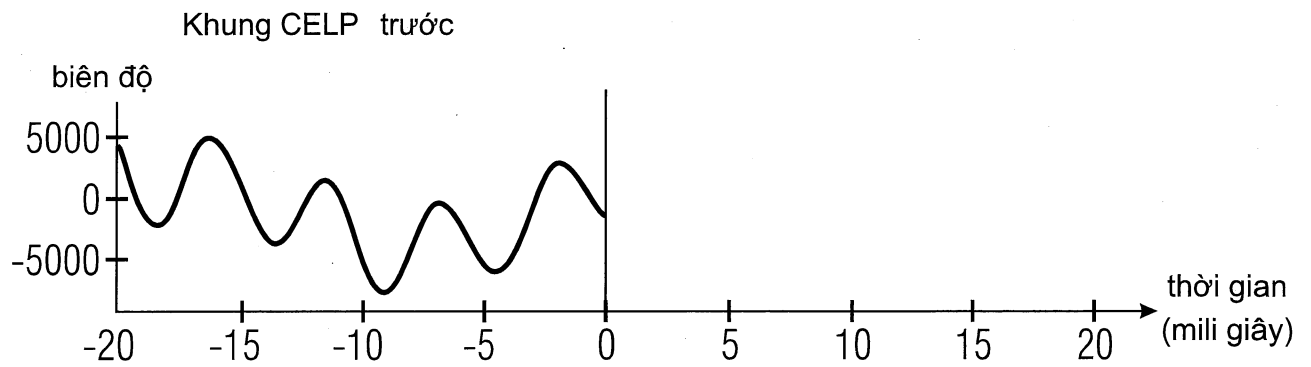


FIG 5A

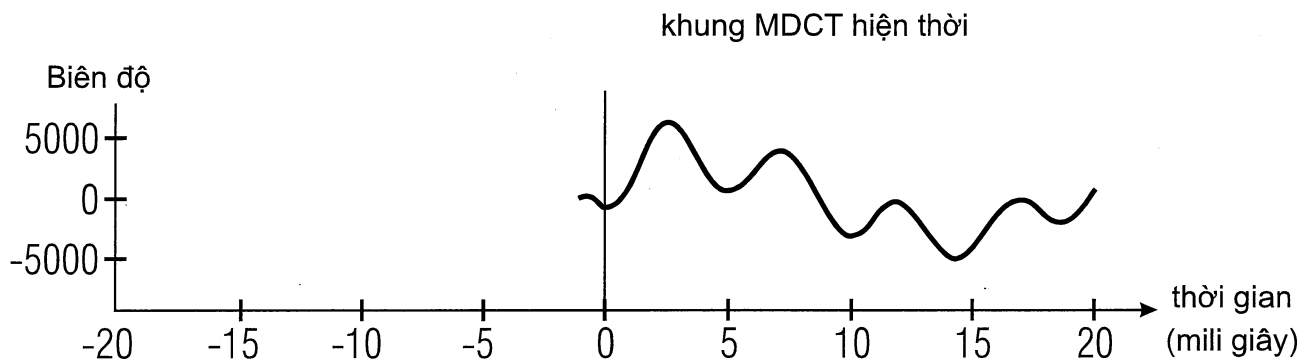


FIG 5B

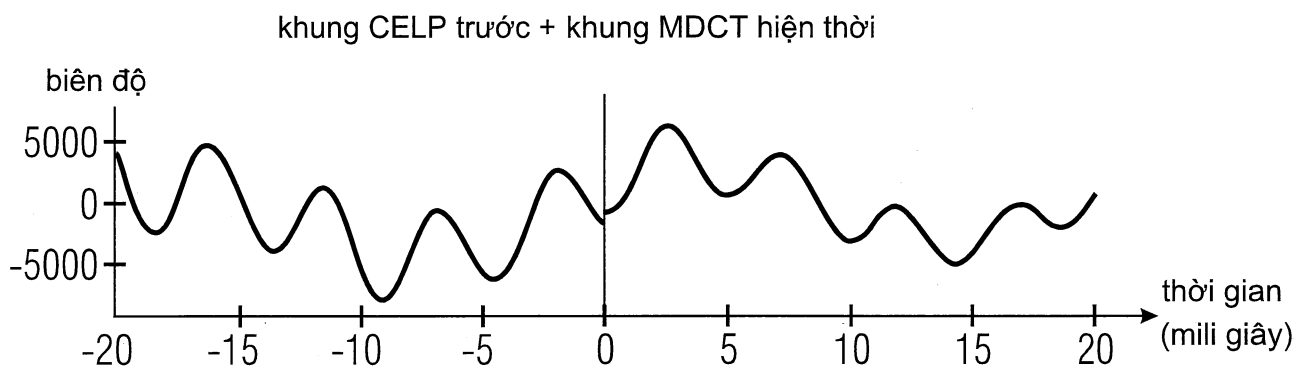


FIG 5C

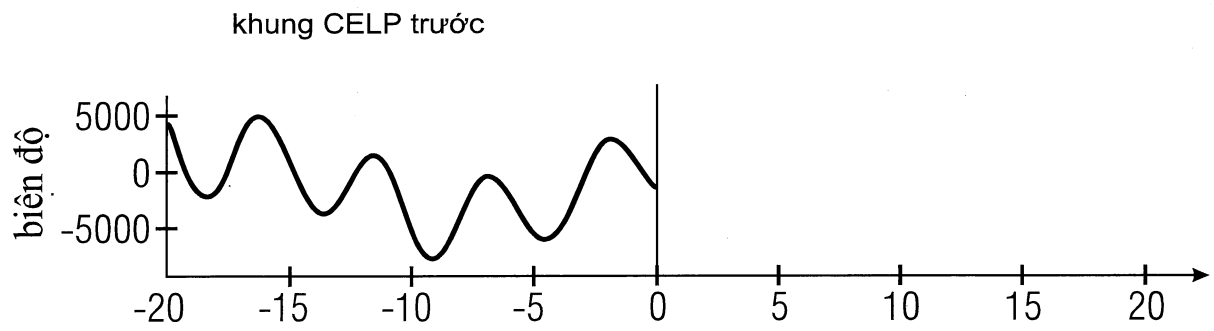


FIG 6A

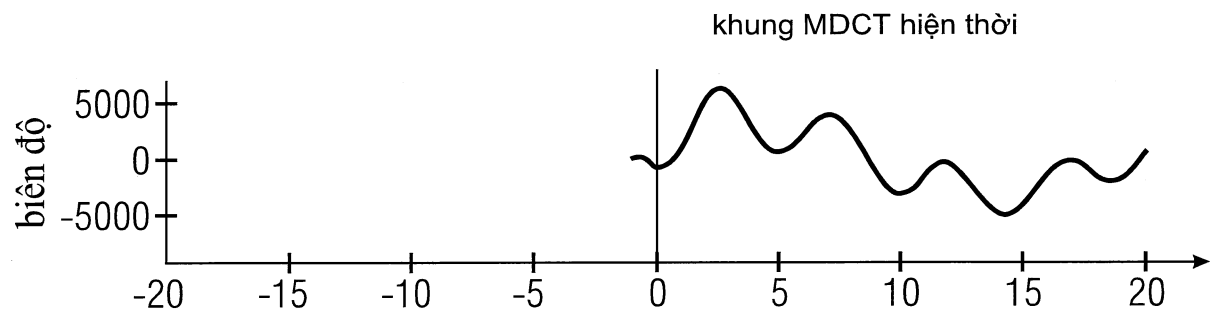


FIG 6B

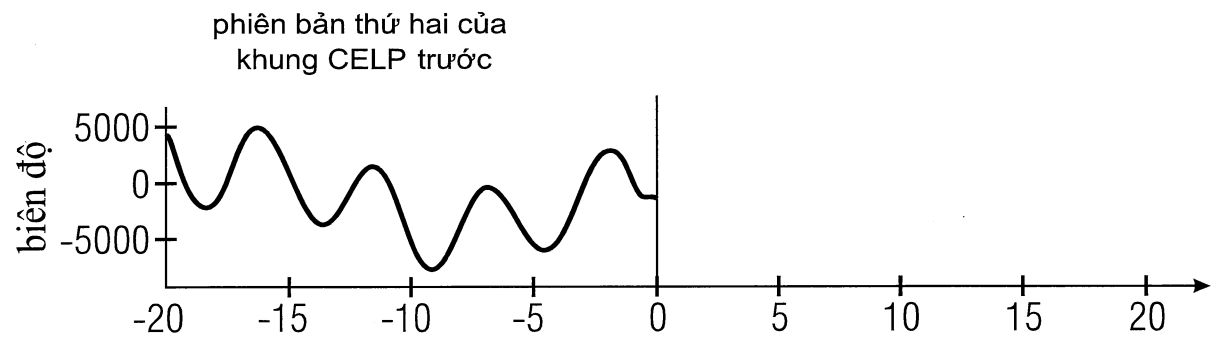


FIG 6C

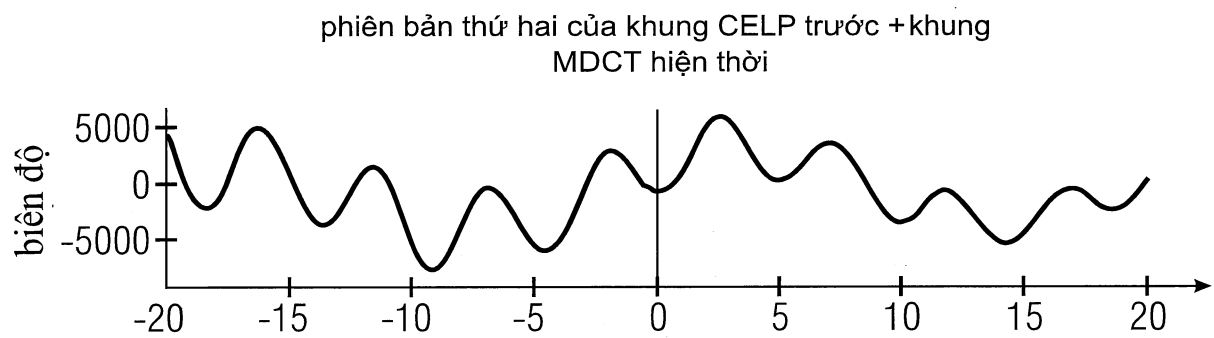


FIG 6D

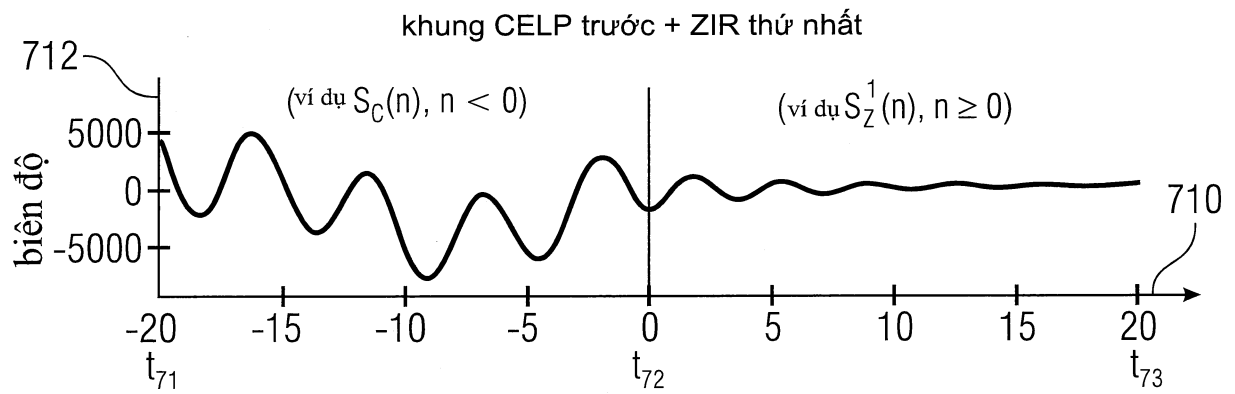


FIG 7A

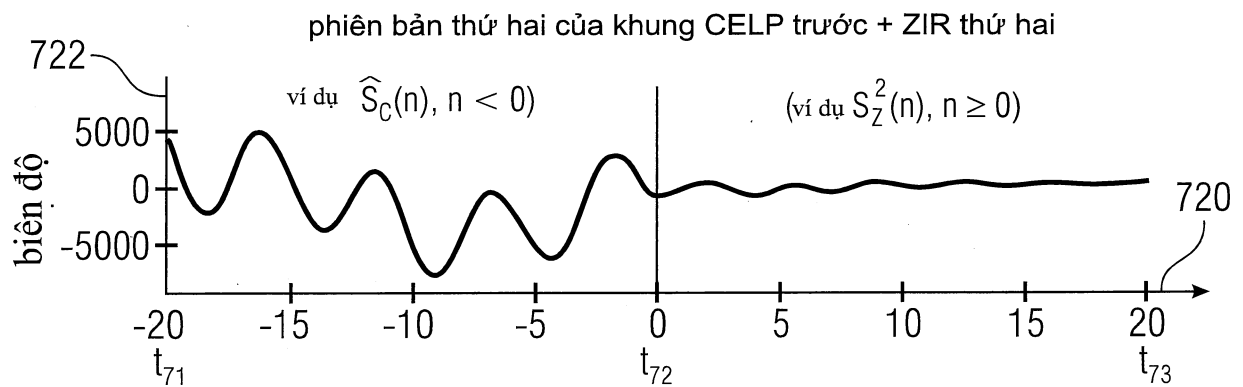


FIG 7B

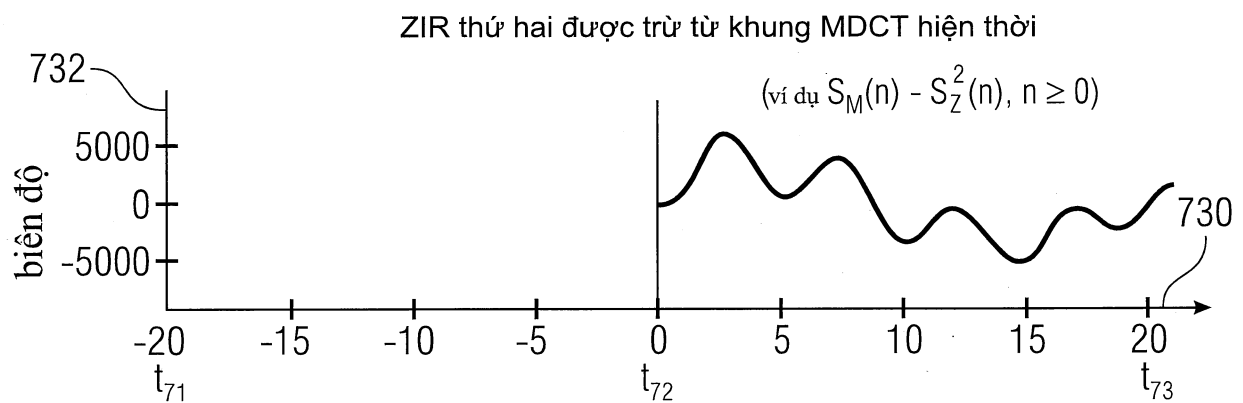


FIG 7C

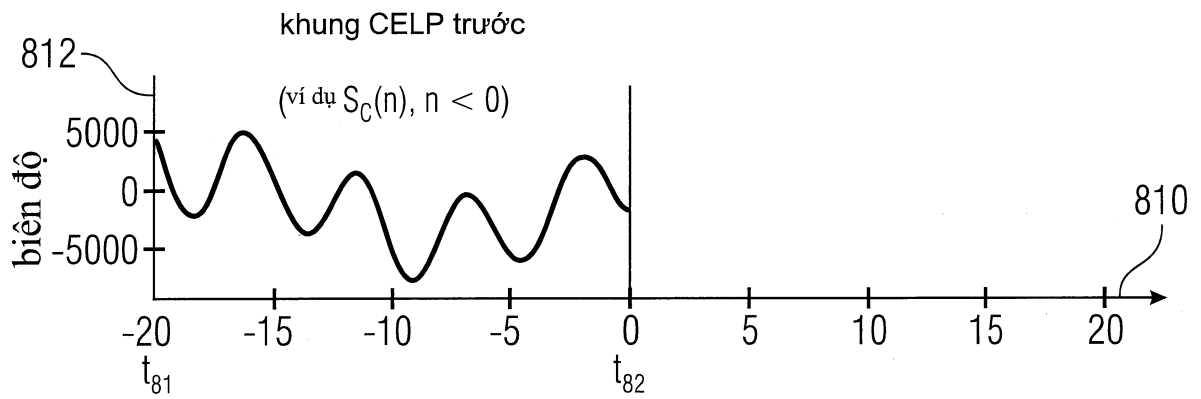


FIG 8A

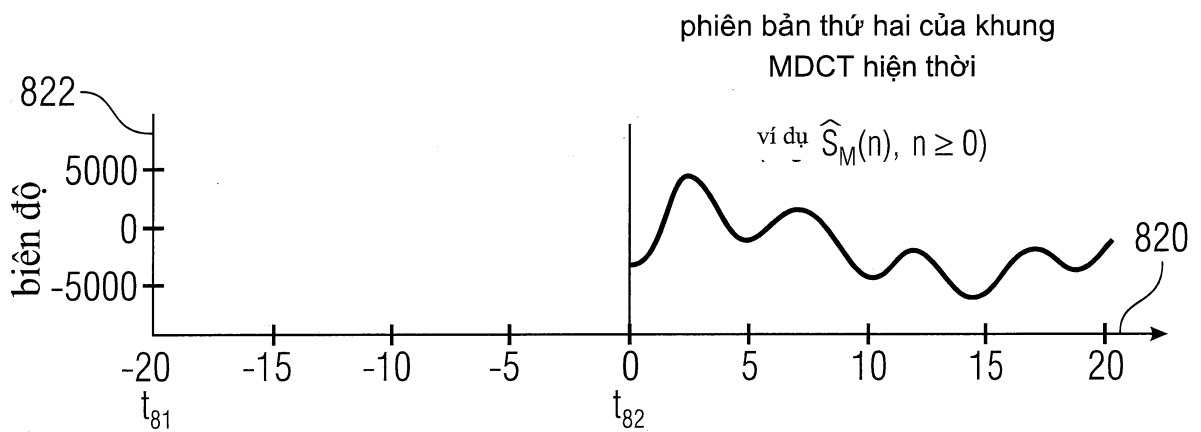


FIG 8B

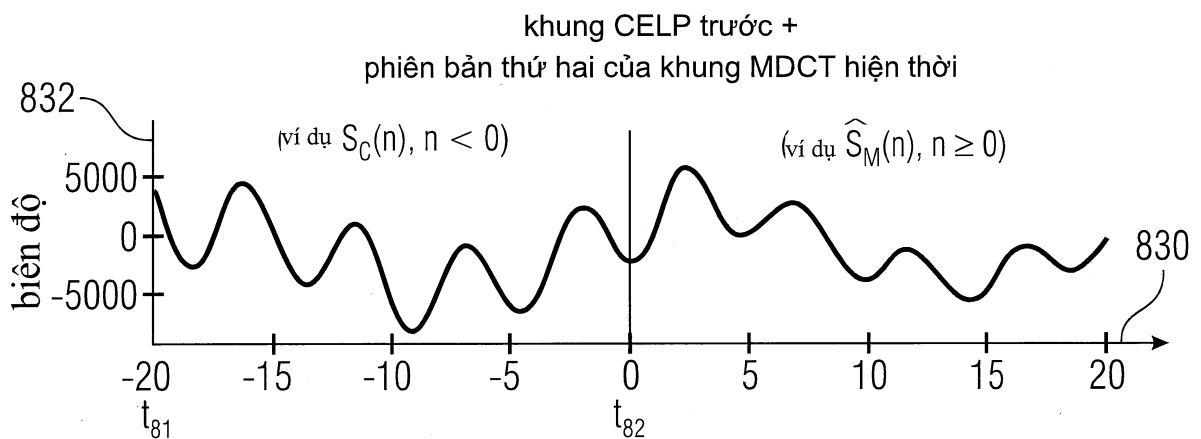


FIG 8C

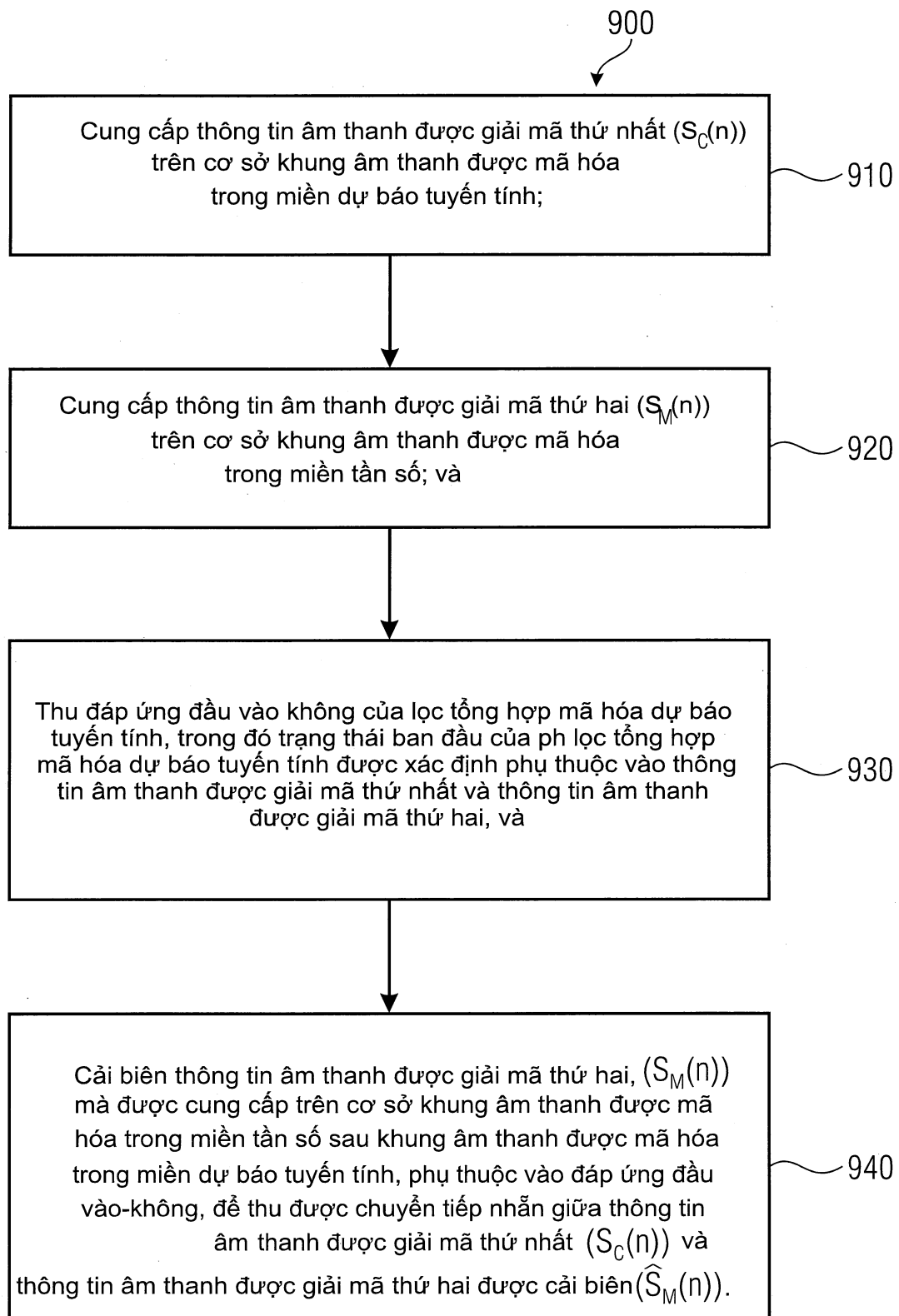


FIG 9

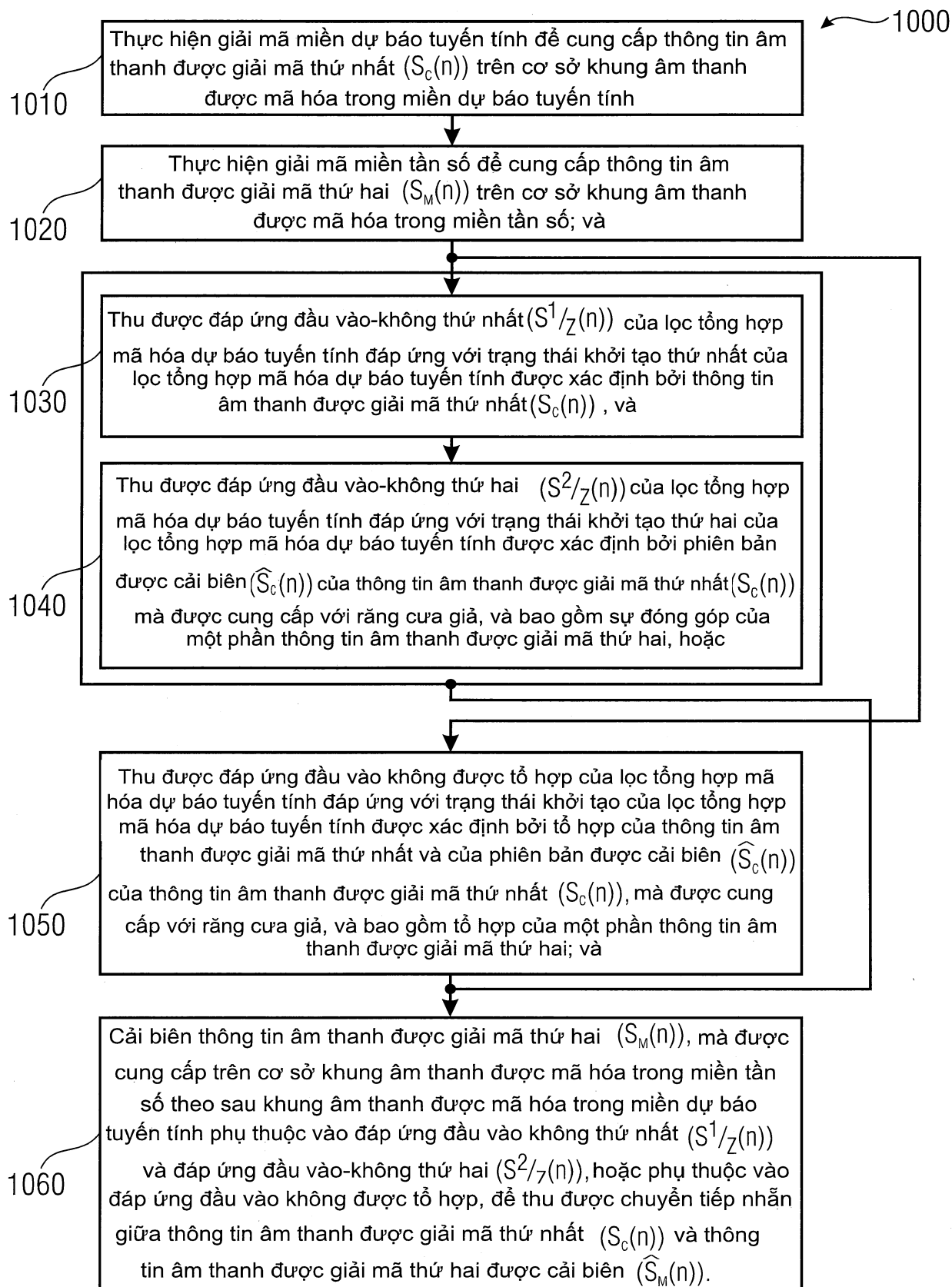


FIG 10