



(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)



1-0043094

(51)⁷ G10L 13/00

(13) B

(21) 1-2010-03252

(22) 22/05/2009

(86) PCT/CN2009/071931 22/05/2009

(87) WO2010/000161 A1 07/01/2010

(30) 200810068262.6 30/06/2008 CN

(45) 25/02/2025 443

(43) 25/05/2011 278A

(73) TENCENT TECHNOLOGY (SHENZHEN) COMPANY LIMITED (CN)

Room 403, East Block 2, SEG Park, Zhenxing Road, Futian District, Shenzhen city
518044, Guangdong Province, P. R. China

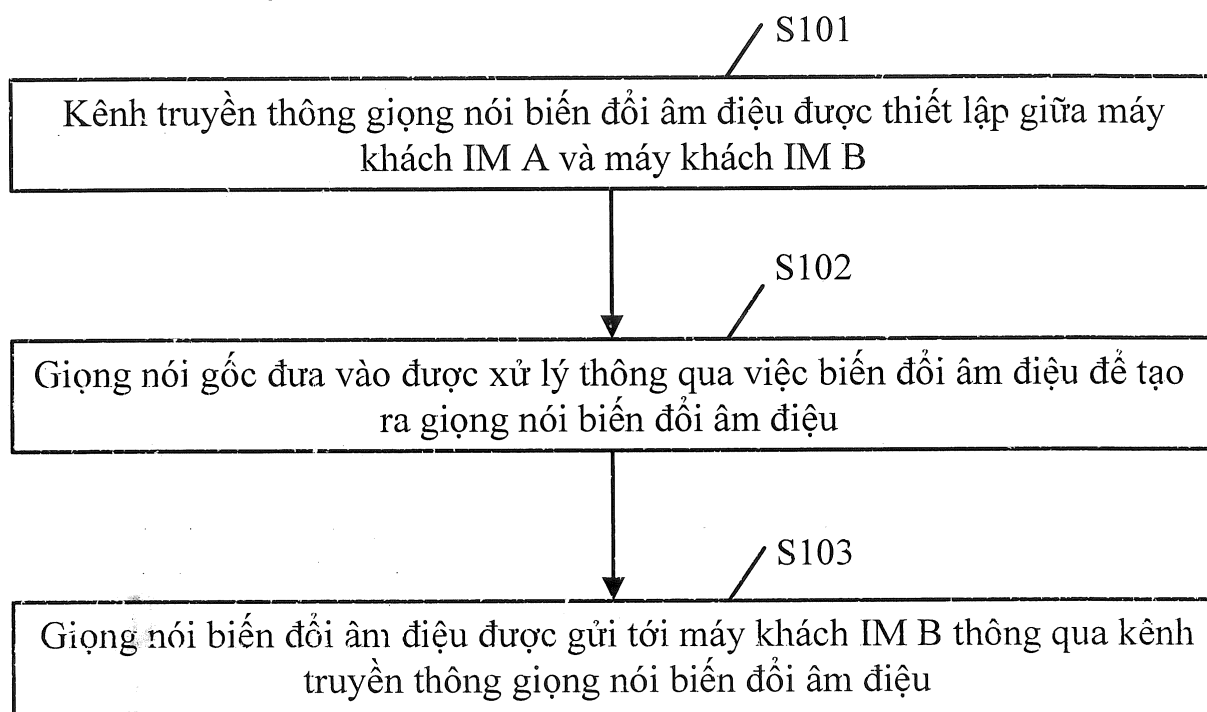
(72) LI, Dalong (CN); ZHENG, Quanzhan (CN); SHENG, Fuzhong (CN).

(74) Công ty TNHH Sở hữu trí tuệ VIPATCO (VIPATCO CO., LTD.)

(54) PHƯƠNG PHÁP VÀ THIẾT BỊ TRUYỀN THÔNG GIỌNG NÓI DỰA TRÊN HỆ
THỐNG TIN NHẮN NHANH

(21) 1-2010-03252

(57) Các phương án của sáng chế bộc lộ phương pháp và thiết bị để truyền thông giọng nói dựa trên hệ thống tin nhắn nhanh (IM - Instant Messaging). Phương pháp bao gồm các bước: a) thiết lập kênh truyền thông giọng nói biến đổi âm điệu giữa máy khách IM thứ hai và máy khách IM thứ nhất; b) xử lý thông tin giọng nói gốc được đưa vào thông qua việc biến đổi âm điệu để có được giọng nói biến đổi âm điệu; gửi giọng nói biến đổi âm điệu tới máy khách IM thứ nhất thông qua kênh truyền thông giọng nói biến đổi âm điệu. Theo các phương án của sáng chế, thông tin giọng nói thu thập được trong hệ thống IM trước tiên được xử lý thông qua việc biến đổi âm điệu, như vậy việc truyền thông giọng nói biến đổi âm điệu dựa trên hệ thống IM sẽ được thực hiện.



Hình 1

Lĩnh vực đề cập của sáng chế

Sáng chế đề cập tới kỹ thuật truyền thông, cụ thể là phương pháp và thiết bị dành cho truyền thông giọng nói dựa trên hệ thống tin nhắn nhanh (IM - Instant Messaging).

Tình trạng kỹ thuật của sáng chế

Cùng với sự phát triển của kỹ thuật tin nhắn nhanh (IM), hệ thống IM đã được kết hợp các chức năng bổ sung, như chức năng truyền thông giọng nói, bên cạnh các chức năng cơ bản của IM. Việc sử dụng hệ thống IM cho truyền thông giọng nói đã trở thành một trong các cách thức truyền thông phổ biến được dùng bởi con người. Tuy nhiên, cách thức truyền thông giọng nói hiện nay có các chức năng rất đơn giản, chẳng hạn như truyền thông giọng nói chỉ có thể dùng được giọng nói gốc của hai phía trong việc truyền thông giọng nói mà không thể thay đổi giọng nói của hai phía được. Kết quả là, danh tính hai bên không được ẩn đi. Như vậy cách thức truyền thông giọng nói hiện có bị thiếu tính mới lạ và không thu hút người dùng, đồng thời không làm thỏa mãn các yêu cầu của từng cá nhân người dùng.

Hiện nay mới chỉ có phương pháp truyền thông giọng nói không biến đổi âm điệu dựa trên nền hệ thống IM.

Bản chất của sáng chế

Các phương án của sáng chế đề cập đến phương pháp truyền thông giọng nói biến đổi âm điệu dựa trên hệ thống IM để giải quyết vấn đề hiện tại không có phương pháp nào dành cho truyền thông giọng nói dựa trên hệ thống IM với âm điệu được thay đổi.

Sáng chế đạt được thông qua giải pháp kỹ thuật sau đây.

Phương pháp truyền thông giọng nói dựa trên IM bao gồm:

- a) thiết lập kênh truyền thông giọng nói có biến đổi âm điệu giữa ít nhất hai máy khách IM;
- b) xử lý thông tin giọng nói gốc thông qua việc biến đổi âm điệu để lấy được giọng nói biến đổi âm điệu; và truyền dẫn giọng nói biến đổi âm điệu tới máy khách IM thứ nhất trong ít nhất hai máy khách IM thông qua kênh truyền thông giọng nói biến đổi âm điệu.

Các phương án của sáng chế cũng đề cập tới thiết bị truyền thông giọng nói dựa trên hệ thống tin nhắn nhanh (IM), và thiết bị này bao gồm:

bộ phận gửi yêu cầu, được làm tương thích để thiết lập kênh truyền thông giọng nói biến đổi âm điệu;

bộ phận thu thập giọng nói, được làm tương thích để tập hợp thông tin giọng nói gốc được đưa vào;

bộ phận biến đổi âm điệu, được làm tương thích để xử lý thông tin giọng nói gốc thu thập được bởi bộ phận thu thập giọng nói thông qua việc biến đổi âm điệu để lấy được giọng nói biến đổi âm điệu;

bộ phận gửi giọng nói, được làm tương thích để gửi giọng nói biến đổi âm điệu lấy được bởi bộ phận biến đổi âm điệu tới thiết bị thu thông qua kênh truyền thông giọng nói biến đổi âm điệu được thiết lập bởi bộ phận truyền dẫn yêu cầu.

Các phương án của sáng chế cũng đề cập đến phương pháp truyền thông giọng nói dựa trên hệ thống tin nhắn nhanh (IM), bao gồm các bước:

thiết lập kênh truyền thông giọng nói giữa ít nhất hai máy khách IM;

xử lý thông tin giọng nói gốc thông qua việc biến đổi âm điệu để lấy được giọng nói biến đổi âm điệu sau khi xác định để thực hiện truyền thông giọng nói biến đổi âm điệu; và truyền dẫn giọng nói biến đổi âm điệu tới máy khách IM thứ nhất trong ít nhất hai máy khách IM thông qua kênh truyền thông giọng nói.

Theo các phương án của sáng chế, thông tin giọng nói thu thập được trong hệ thống IM được xử lý trước thông qua việc biến đổi âm điệu, do vậy truyền thông giọng nói biến đổi âm điệu dựa trên hệ thống IM được thực hiện. Truyền thông giọng nói trong hệ thống IM sẽ tạo thêm tính giải trí, và có thể giới thiệu sản phẩm mới bổ sung cho các dịch vụ giá trị gia tăng trong các dịch vụ IM thông thường. Các dịch vụ IM sẽ trở nên thu hút hơn đối với người dùng và như vậy sẽ trở nên cạnh tranh hơn và mang đến trải nghiệm dịch vụ mới cho người dùng truyền thông giọng nói.

Mô tả vắn tắt các hình vẽ

Hình 1 là lưu đồ minh họa quy trình cơ bản của phương pháp theo phương án của sáng chế.

Hình 2 là lưu đồ minh họa quy trình chi tiết của phương pháp theo phương án của sáng chế.

Hình 3 là lưu đồ minh họa quy trình chi tiết của phương pháp theo phương án của sáng chế.

Hình 4 là lưu đồ minh họa quy trình sau khi máy khách IM B nhận được dữ liệu truyền thông giọng nói biến đổi âm điệu được gửi bởi máy khách IM A theo phương án của sáng chế.

Hình 5 là sơ đồ minh họa cấu trúc cơ bản của thiết bị theo phương án của sáng chế.

Hình 6 là sơ đồ minh họa cấu trúc chi tiết của thiết bị theo phương án của sáng chế.

Mô tả chi tiết sáng chế

Sáng chế được mô tả chi tiết dưới đây cùng với các hình vẽ đi kèm cũng như các phương án để làm sáng tỏ mục đích, giải pháp kỹ thuật và bản chất của sáng chế.

Theo phương án của sáng chế, kênh truyền thông giọng nói biến đổi âm điệu được thiết lập giữa ít nhất hai máy khách IM. Ví dụ, kênh truyền thông giọng nói biến đổi âm điệu được thiết lập giữa máy khách IM A, máy khách IM B và máy khách IM C. Để tiện cho việc mô tả, phần mô tả sau đây sẽ lấy việc thiết lập kênh truyền thông giọng nói biến đổi âm điệu giữa máy khách IM A và máy khách IM B làm ví dụ, và quy trình tương tự có thể được áp dụng cho các tình huống khác mà chưa được đề cập đến. Cụ thể, máy khách IM A sẽ gửi yêu cầu biến đổi âm điệu tới máy khách IM B, và thiết lập kênh truyền thông giọng nói biến đổi âm điệu với máy khách IM B. Sau đó, máy khách IM A xử lý giọng nói gốc thu thập được thông qua việc biến đổi âm điệu để lấy giọng nói biến đổi âm điệu của giọng nói gốc, và gửi giọng nói biến đổi âm điệu tới máy khách IM B thông qua kênh truyền thông giọng nói biến đổi âm điệu đã được thiết lập, qua đó thực hiện truyền thông giọng nói biến đổi âm điệu giữa các máy khách IM trong hệ thống IM.

Tham khảo Hình 1, đây là lưu đồ minh họa quy trình cơ bản của phương pháp theo phương án của sáng chế. Như được thể hiện trên Hình 1, phương án này lấy việc thiết lập kênh truyền thông giọng nói biến đổi âm điệu giữa máy khách IM A và máy khách IM B làm ví dụ. Quy trình bao gồm các bước sau.

Bước S101, kênh truyền thông giọng nói biến đổi âm điệu được thiết lập giữa máy khách IM A và máy khách IM B.

Bước S102, giọng nói gốc đưa vào được xử lý thông qua biến đổi âm điệu để tạo ra giọng nói biến đổi âm điệu.

Bước S103, giọng nói biến đổi âm điệu được gửi tới máy khách IM B thông qua kênh truyền thông giọng nói biến đổi âm điệu.

Cần lưu ý rằng máy khách IM A và máy khách IM B có thể được thực hiện theo các dạng khác nhau, như là máy khách dạng web hoặc máy khách mạng không dây, và không bị giới hạn trong các ví dụ dùng để mô tả sáng chế.

Cũng cần lưu ý rằng hoạt động của các bước S102 và S103 có thể được thực hiện bởi máy khách IM A, chẳng hạn máy khách IM A sẽ xử lý giọng nói gốc thông qua việc biến đổi âm điệu để có được giọng nói biến đổi âm điệu, và gửi giọng nói biến đổi âm điệu tới máy khách IM B thông qua kênh truyền thông giọng nói biến đổi âm điệu theo cách thức giao nhận máy chủ hoặc cách thức ngang hàng (P2P). Ngoài ra, các hoạt động được thực hiện bởi thiết bị biến đổi âm điệu được thiết kế sẵn, như là máy chủ, chẳng hạn máy chủ sẽ thu nhận giọng nói gốc được gửi bởi máy khách IM A, xử lý giọng nói gốc thông qua việc biến đổi âm điệu để có được giọng nói biến đổi âm điệu; và gửi giọng nói biến đổi âm điệu tới máy khách IM B thông qua kênh truyền thông giọng nói biến đổi âm điệu. Chi tiết thực hiện sẽ không bị giới hạn trong sáng chế này. Để tiện cho việc mô tả, truyền thông giọng nói giữa hai máy khách được lấy làm ví dụ trong mô tả dưới đây.

Như được đề cập ở trên, quy trình cơ bản của truyền thông giọng nói dựa trên hệ thống IM theo các phương án của sáng chế được thực hiện.

Phần trên mô tả quy trình của các phương án theo sáng chế một cách tổng quát, và quy trình sẽ được mô tả chi tiết cùng với các phương án.

Tham khảo Hình 2, Hình 2 là lưu đồ minh họa chi tiết quy trình của phương pháp theo phương án của sáng chế, chi tiết như sau.

1) Máy khách IM A gửi yêu cầu thực hiện truyền thông giọng nói biến đổi âm điệu tới máy khách IM B.

2) Máy khách IM B sẽ thu nhận yêu cầu thực hiện truyền thông giọng nói biến đổi âm điệu từ máy khách IM A, đáp ứng yêu cầu, và trả về thông tin tương ứng cho máy khách IM A. Khi thu nhận đáp ứng thực hiện truyền thông giọng nói biến đổi âm điệu từ máy khách IM B, máy khách IM A sẽ thiết lập kênh truyền thông giọng nói biến đổi âm điệu giữa máy khách IM A và máy khách IM B.

Để thiết lập kênh truyền thông giọng nói thành công, máy khách IM A và máy khách IM B sẽ thiết lập kênh truyền thông giọng nói biến đổi âm điệu kết hợp với máy chủ. Dĩ nhiên, máy khách IM A có thể gửi minh bạch hoặc không minh bạch yêu cầu

thực hiện truyền thông giọng nói biến đổi âm điệu tới máy khách IM B. Cụ thể, nếu máy khách IM A gửi minh bạch yêu cầu thực hiện truyền thông giọng nói biến đổi âm điệu tới máy khách IM B, thủ tục này không cần hiển thị trong giao diện của máy khách IM B.

3) Máy khách IM A sẽ xử lý giọng nói gốc thu thập được thông qua việc biến đổi âm điệu, và lấy giọng nói biến đổi âm điệu tương ứng với giọng nói gốc.

Các phương án của sáng chế đề cập đến một loạt các phương pháp biến đổi âm điệu, như là biến đổi âm điệu của giọng nói gốc, thay đổi giới tính của giọng nói gốc (thực chất là đổi giọng nói của người nam thành giọng nói của người nữ hoặc thay đổi giọng nói của người nữ thành giọng nói của người nam), thay đổi độ tuổi của giọng nói gốc (chẳng hạn thay đổi giọng nói của người trẻ thành giọng nói của người già), thay đổi giọng nói gốc của người dùng thành giọng nói của người nổi tiếng, thêm âm thanh nền vào giọng nói gốc (thực chất việc thêm âm thanh nền vào giọng nói của người dùng không phải là dạng biến đổi âm điệu mà là dạng trộn âm thanh; nhưng việc biến đổi âm điệu theo sáng chế này bao gồm cả việc trộn âm thanh).

Quy trình xử lý chi tiết giọng nói gốc thu thập được thông qua việc biến đổi âm điệu để có được giọng nói biến đổi âm điệu bao gồm các bước sau:

A) thu thập thông tin giọng nói đưa vào bởi người dùng và xử lý thông tin giọng nói thu thập được để tạo ra tín hiệu giọng nói kỹ thuật số có khả năng phân biệt và xử lý được bởi máy tính;

B) xử lý tín hiệu giọng nói kỹ thuật số thông qua việc biến đổi âm điệu và lấy được giọng nói biến đổi âm điệu tương ứng với tín hiệu giọng nói kỹ thuật số.

Trong phương án này, việc biến đổi âm điệu được thực hiện bằng cách: làm tan biến tín hiệu giọng nói kỹ thuật số sử dụng phân tích dự đoán tuyến tính (LP) và tổng hợp các mô hình thành phần đường bao phổ (biểu thị bởi Mã hóa dự đoán tuyến tính (LPC)) và phần kích thích (biểu thị bởi phần còn lại của LPC); lấy tần số Fooc-măng (Formant) và tham số độ dốc quang phổ từ hệ số LPC, và thực hiện chuyển đổi giọng nói sử dụng cách thức điện tín lượng tử hóa véctor. Với các hàm chuyển đổi, việc chuyển đổi đường bao tần số có thể tương thích với lượng tử hóa véctor, và chuyển đổi ngôn ngữ (chủ yếu đề cập đến giai đoạn cao độ) có thể tương thích với thuật toán cộng xếp chồng đồng bộ cao độ miền thời gian (TD-PSOLA).

Trong phương án này, cách thức biến đổi âm điệu cần được xác định trước khi thực hiện biến đổi âm điệu. Cụ thể, xác định cách biến đổi âm điệu để tương thích hiện

tại bao gồm: xác định thông tin biến đổi âm điệu hiện tại, và xác định cách biến đổi âm điệu để tương thích theo thông tin biến đổi âm điệu hiện tại. Thông tin biến đổi âm điệu hiện tại bao gồm: thông tin lựa chọn người dùng, và/hoặc thông tin biến đổi âm điệu được cho phép. Thông tin lựa chọn người dùng là sự lựa chọn bởi người dùng từ các cách biến đổi âm điệu cho trước; thông tin biến đổi âm điệu cho phép là thông tin biến đổi âm điệu cho phép bởi hệ thống IM đối với người dùng để thực hiện biến đổi âm điệu.

Tốt hơn là, để tạo ra sản phẩm bổ sung mới trong các dịch vụ gia tăng giá trị của dịch vụ IM thông thường, thì nhà cung cấp dịch vụ IM phải cung cấp một số cách biến đổi âm điệu như là một dịch vụ gia tăng giá trị. Theo các phương án của sáng chế, các cách biến đổi âm điệu được cung cấp có thể được xác định dựa trên các cách biến đổi âm điệu được phép của người dùng bắt đầu biến đổi âm điệu trong hệ thống IM. Trước khi người dùng ở máy khách IM A lựa chọn cách biến đổi âm điệu, thì người dùng sẽ gửi thông tin truy vấn cách thức biến đổi âm điệu cho phép tới máy chủ qua máy khách IM A, và theo định danh người dùng trong hệ thống IM, máy chủ sẽ trả về thông tin cách thức biến đổi âm điệu được phép, thực chất là các cách biến đổi âm điệu mà được dùng bởi người dùng. Tốt hơn là, người dùng máy khách IM A có thể nhập vào thông tin lựa chọn người dùng dựa trên thông tin biến đổi âm điệu được phép để xác định cách biến đổi âm điệu phù hợp dựa trên thông tin lựa chọn người dùng và thông tin biến đổi âm điệu được phép trả về bởi máy chủ. Logic lựa chọn dịch vụ khác cũng được sử dụng cho việc xác định cách biến đổi âm điệu dựa trên thông tin lựa chọn người dùng và thông tin cách biến đổi âm điệu được phép; khi người dùng chỉ có một cách biến đổi âm điệu có sẵn, thì cách biến đổi âm điệu có thể được xác định dựa trên thông tin cách biến đổi âm điệu được phép.

Biến đổi âm điệu được thực hiện dựa trên các tín hiệu giọng nói gốc của người dùng. Do đó, khi xác định cách biến đổi âm điệu để biến đổi giọng nói gốc, phương án cũng lấy thông tin đặc tính người dùng để xem xét, như là tính năng phân đoạn giọng nói gốc của người dùng, để cung cấp thêm cách biến đổi âm điệu phù hợp cho người dùng sao cho giọng nói biến đổi âm điệu có thể được nhận ra bởi người mà người dùng đang giao tiếp với họ. Cách biến đổi âm điệu có thể được xác định bởi logic lựa chọn dịch vụ dựa trên thông tin lựa chọn người dùng và thông tin đặc tính người dùng, hoặc dựa trên thông tin lựa chọn người dùng, thông tin biến đổi âm điệu được phép và thông tin đặc tính người dùng. Logic lựa chọn dịch vụ được xác định bởi nhà cung cấp dịch vụ IM, chỉ rõ có bao nhiêu dịch vụ biến đổi âm điệu (chẳng hạn “chuyển giọng nói nam

sang giọng nói nữ” là một dạng dịch vụ biến đổi âm điệu) có tác dụng đối với thông tin biến đổi âm điệu được phép nhất định và môi trường truyền thông giọng nói nhất định, sau đó logic lựa chọn dịch vụ được dùng để xác định cách biến đổi âm điệu.

Sau khi nhận thông tin lựa chọn người dùng, máy khách IM A sẽ phân tích tín hiệu giọng nói gốc của người dùng để lấy thông tin đặc tính người dùng. Khi thông tin đặc tính người dùng không phù hợp với yêu cầu biến đổi âm điệu, thì cách biến đổi âm điệu được yêu cầu bởi người dùng sẽ được thay đổi. Ví dụ, khi giọng nói gốc của người dùng trầm và khàn và người dùng chọn lựa cách biến đổi âm điệu là “giọng trẻ con”, thì hiệu quả biến đổi âm điệu sẽ rất kém (không thể nhận ra được là “giọng nói trẻ con”). Do đó, hệ thống sẽ đề xuất người dùng chọn cách biến đổi âm điệu khác.

Để cải thiện chất lượng giọng nói nghe được bởi người nghe trong truyền thông và cung cấp đúng cách biến đổi âm điệu đối với người dùng, phương án khác còn tính đến thông tin môi trường giọng nói của người nghe. Cách biến đổi âm điệu có thể được xác định bởi logic lựa chọn dịch vụ dựa trên thông tin lựa chọn người dùng và thông tin môi trường giọng nói của người tiếp nhận, hoặc dựa trên thông tin lựa chọn người dùng, thông tin biến đổi âm điệu được phép và thông tin môi trường giọng nói của người tiếp nhận. Thông tin môi trường giọng nói của người tiếp nhận được gửi bởi máy khách IM B tới máy khách IM A khi máy khách IM B trả về đáp ứng yêu cầu truyền thông giọng nói biến đổi âm điệu cho máy khách IM A. Thông tin môi trường giọng nói có thể được chọn bởi người dùng máy khách IM B, hoặc được lấy bởi máy khách IM B dựa trên phân tích tín hiệu giọng nói thu được bằng mic (microphone).

Theo các phương án của sáng chế, cách biến đổi âm điệu của máy khách IM A có thể được xác định bằng logic lựa chọn dịch vụ dựa trên thông tin lựa chọn người dùng và bất kỳ hoặc kết hợp với thông tin cách biến đổi âm điệu được phép, thông tin đặc tính người dùng và thông tin môi trường giọng nói của người tiếp nhận.

Trong các phương án của sáng chế, thông tin giọng nói thu được có thể chứa các tín hiệu như là tiếng vang và nhiễu sẽ gây bất lợi trong quá trình xử lý, truyền tải và nhận dạng thông tin giọng nói. Như vậy, trước khi thông tin giọng nói kỹ thuật số được xử lý thông qua việc biến đổi âm điệu, thì thông tin giọng nói kỹ thuật số cần được xử lý để loại bỏ nhiễu, thực chất là bất kỳ hoặc bất kỳ sự kết hợp nào của việc loại bỏ tiếng vang, giảm nhiễu và điều khiển khuếch đại tín hiệu và tương tự như vậy, để đạt được hiệu quả tốt hơn trong truyền thông giọng nói biến đổi âm điệu và cải thiện chất lượng nghe đối với người tiếp nhận.

4) Máy khách IM A sẽ gửi giọng nói biến đổi âm điệu lấy được tới máy khách IM B bằng kênh truyền thông giọng nói biến đổi âm điệu đã được thiết lập.

Theo các phương án của sáng chế, để thuận tiện cho việc truyền tải giọng nói biến đổi âm điệu, máy khách IM A có thể nhóm và đóng gói giọng nói biến đổi âm điệu trước khi gửi giọng nói biến đổi âm điệu để có được các gói giọng nói biến đổi âm điệu, sau đó gửi các gói giọng nói biến đổi âm điệu tới máy khách IM B.

Theo các phương án của sáng chế, sau khi âm điệu của giọng nói gốc đã thu được biến đổi, thì giọng nói biến đổi âm điệu tương ứng với giọng nói gốc thu được sẽ được nén và mã hóa theo quy tắc mã hóa cho trước, chẳng hạn như là G.729, G.729, G.723.1, sao cho băng thông cần thiết để truyền tải dữ liệu giọng nói biến đổi âm điệu được giảm xuống và truyền thông giọng nói biến đổi âm điệu thời gian thực được thuận tiện.

Để tránh méo tín hiệu do mất gói giọng nói và lỗi khi truyền tải trong mạng, sau khi giọng nói biến đổi âm điệu được nén và mã hóa, luồng bit (bit streams) lấy được sau khi nén và mã hóa sẽ được xử lý thông qua việc tăng cường lặp lại (redundancy enhancing) bằng việc sử dụng kỹ thuật mã hóa kênh.

Quy trình máy khách IM B gửi yêu cầu truyền thông giọng nói biến đổi âm điệu tới máy khách IM A là tương tự như quy trình được mô tả ở trên, và không được mô tả ở đây nữa. Cần hiểu rằng máy khách IM A và máy khách IM B có thể thực hiện truyền thông giọng nói biến đổi một chiều hoặc hai chiều. Truyền thông giọng nói như đề cập ở trên có thể được thực hiện trong hệ thống IM dựa trên mạng có dây hoặc mạng không dây.

Khi bất kỳ máy khách IM A và IM B yêu cầu ngắt kết nối hoặc khi mạng bị lỗi, việc truyền thông sẽ bị kết thúc và kênh truyền thông giọng nói biến đổi âm điệu sẽ bị loại bỏ.

Hình 3 là lưu đồ minh họa phương pháp theo phương án của sáng chế. Theo phương án này, kênh truyền thông giọng nói được thiết lập giữa máy khách IM A và máy khách IM B, máy khách IM A và máy khách IM B sẽ thực hiện truyền thông giọng nói. Phương pháp bao gồm các bước sau:

1) Máy khách IM A sẽ gửi yêu cầu truyền thông giọng nói tới máy khách IM B.

2) Máy khách IM B đáp ứng sau khi nhận được yêu cầu truyền thông giọng nói từ máy khách IM A, và trả về thông tin cho máy khách IM A. Khi nhận được thông tin đáp ứng để thực hiện truyền thông giọng nói từ máy khách IM B, máy khách IM A sẽ thiết lập kênh truyền thông giọng nói giữa máy khách IM A và máy khách IM B.

Sau khi thiết lập kênh truyền thông giọng nói, máy khách IM A và máy khách IM B sẽ thực hiện truyền thông giọng nói với nhau bằng kênh truyền thông giọng nói.

3) Máy khách IM A sẽ gửi yêu cầu truyền thông giọng nói biến đổi âm điệu tới máy khách IM B.

4) Máy khách IM B sẽ đáp ứng sau khi nhận được yêu cầu truyền thông giọng nói biến đổi âm điệu từ máy khách IM A, và trả về thông tin đáp ứng tới máy khách IM A. Khi nhận được thông tin đáp ứng để thực hiện truyền thông giọng nói biến đổi âm điệu từ máy khách IM B, thì máy khách IM A sẽ thiết lập kênh truyền thông giọng nói biến đổi âm điệu giữa máy khách IM A và máy khách IM B.

Sau khi kênh truyền thông giọng nói biến đổi âm điệu được thiết lập, kênh truyền thông giọng nói được thiết lập trước đó sẽ bị hủy. Máy khách IM A sẽ gửi yêu cầu truyền thông giọng nói biến đổi âm điệu một cách minh bạch hoặc không minh bạch tới máy khách IM B. Nếu máy khách IM A gửi minh bạch yêu cầu truyền thông giọng nói biến đổi âm điệu tới máy khách IM B, thủ tục này sẽ không được hiển thị trên giao diện của máy khách IM B.

5) Máy khách IM A sẽ xử lý giọng nói gốc thu thập được thông qua việc biến đổi âm điệu, và lấy được giọng nói biến đổi âm điệu tương ứng với giọng nói gốc.

6) Máy khách IM A sẽ gửi giọng nói biến đổi âm điệu tới máy khách IM B thông qua kênh truyền thông giọng nói biến đổi âm điệu đã được thiết lập.

Cần lưu ý rằng phương án này lấy việc thiết lập kênh truyền thông giọng nói biến đổi âm điệu giữa máy khách IM A và máy khách IM B sau khi xác lập được kênh truyền thông giọng nói giữa máy khách IM A và máy khách IM B làm ví dụ. Để làm cho phương án này đơn giản và dễ thực hiện hơn, thì máy khách IM A không thiết lập kênh truyền thông giọng nói biến đổi âm điệu với máy khách IM B sau khi nhận được thông tin đáp ứng để thực hiện truyền thông giọng nói biến đổi âm điệu từ máy khách IM B, mà chỉ sử dụng kênh truyền thông giọng nói đã được thiết lập tại bước 2) để gửi giọng nói biến đổi âm điệu tới máy khách IM B. Như vậy, thao tác thiết lập kênh truyền thông giọng nói biến đổi âm điệu tại bước 4) được bỏ qua. Tốt hơn là, một trong các tiêu chí xác định liệu có nên thiết lập kênh truyền thông giọng nói biến đổi âm điệu không có thể xác định rằng liệu băng thông của kênh truyền thông giọng nói có đủ cho việc truyền tải giọng nói biến đổi âm điệu lấy được ở bước 5) không.

7) Kênh truyền thông giọng nói biến đổi âm điệu được hủy khi việc truyền thông được kết thúc.

Khi bất kỳ máy khách IM A và IM B yêu cầu ngắt kết nối hoặc khi mạng bị lỗi, việc truyền thông được kết thúc và kênh truyền thông giọng nói biến đổi âm điệu sẽ bị hủy.

Sau khi máy khách IM B nhận dữ liệu truyền thông giọng nói biến đổi âm điệu được gửi bởi máy khách IM A, việc xử lý dữ liệu truyền thông được thực hiện bởi máy khách IM B tương tự như quy trình xử lý trong truyền thông giọng nói gốc. Quá trình xử lý được thể hiện trên Hình 4, và bao gồm các bước như sau.

Bước S401, dữ liệu truyền thông được tiếp nhận và bung ra.

Các gói dữ liệu truyền thông được tiếp nhận thông qua kênh truyền thông giọng nói biến đổi âm điệu đã thiết lập, được bung ra theo cùng giao thức truyền tải mạng tương thích với máy khách IM A, và ghép nối để có được luồng mã nén.

Bước S402, dữ liệu bung ra được giải mã thành tín hiệu giọng nói.

Luồng mã nén đã bung ra được giải mã bằng việc sử dụng thao tác ngược đối với thao tác mã hóa của máy khách IM A để lấy được tín hiệu giọng nói mà có thể được nhận biết bởi tai con người.

Bước S403, các tín hiệu giọng nói được khuếch đại.

Tín hiệu giọng nói có thể bị méo do truyền tải trên mạng, việc nén tín hiệu giọng nói, biến đổi âm giọng nói, v.v. Như vậy, việc khuếch đại tín hiệu là cần thiết đối với tín hiệu giọng nói lấy được bằng việc giải mã. Cường độ tín hiệu có thể tương thích với phép lọc Kalman, ước lượng biên độ quang phổ thời gian ngắn sai lệch bình phương trung bình tối thiểu (MMSE), hoặc lọc thích ứng, v.v.

Bước S404, tín hiệu giọng nói được khuếch đại sẽ được xuất ra.

Tín hiệu giọng nói khuếch đại được xuất ra thông qua thiết bị đầu ra, như là tai nghe, hộp tăng âm và các âm thanh.

Để lấy được luồng bit (bit streams) giọng nói có thể được giải mã đúng, dữ liệu sau khi được nhận và bung ra sẽ được xử lý thông qua dung sai sai lệch/loại bỏ phần lặp lại, để loại bỏ các tín hiệu lặp lại bị chèn bởi máy khách IM A vào luồng mã nén và thay đổi hoặc loại bỏ các dữ liệu sai lệch.

Phương pháp mô tả trên đây được đề cập bởi các phương án của sáng chế một cách chi tiết, và phần mô tả chi tiết sau đây đề cập đến thiết bị theo các phương án của sáng chế.

Hình 5 là sơ đồ minh họa cấu trúc cơ bản của thiết bị theo phương án của sáng chế. Như được thể hiện trên Hình 5, thiết bị bao gồm bộ phận gửi yêu cầu 501, bộ phận thu thập giọng nói 502, bộ phận biến đổi âm điệu 503 và bộ phận gửi giọng nói 504.

Bộ phận gửi yêu cầu 501 được làm tương thích để thiết lập kênh truyền thông giọng nói biến đổi âm điệu.

Bộ phận thu thập giọng nói 502 được làm tương thích để thu thập thông tin giọng nói gốc được đưa vào.

Bộ phận biến đổi âm điệu 503 được làm tương thích để xử lý thông tin giọng nói gốc thu thập được bởi bộ phận thu thập giọng nói 502 thông qua việc biến đổi âm điệu để lấy được giọng nói biến đổi âm điệu.

Bộ phận gửi giọng nói 504 được làm tương thích để gửi giọng nói biến đổi âm điệu có được bởi bộ phận biến đổi âm điệu 503 thông qua kênh truyền thông giọng nói biến đổi âm điệu được thiết lập bởi bộ phận gửi yêu cầu 501.

Các bộ phận nói trên tạo thành một thiết bị cơ bản dành cho truyền thông giọng nói dựa trên hệ thống IM.

Để làm rõ hơn thiết bị truyền thông giọng nói dựa trên hệ thống IM, cấu trúc thiết bị theo phương án của sáng chế sẽ được mô tả chi tiết dưới đây.

Hình 6 là sơ đồ khối minh họa cấu trúc chi tiết của thiết bị theo phương án của sáng chế. Tham khảo Hình 6, chỉ có những phần liên quan đến phương án của sáng chế mới được thể hiện trên Hình 6.

Thiết bị có thể được áp dụng cho bất kỳ thiết bị máy khách IM nào, như là máy vi tính, máy tính xách tay, thiết bị trợ giúp cá nhân kỹ thuật số (PDA) và điện thoại thông minh, và có thể là bộ phận phần mềm, phần cứng, hoặc kết hợp giữa phần cứng và phần mềm trong các thiết bị IM nói trên, hoặc là một plug-in độc lập được tích hợp trong các thiết bị máy khách IM hoặc hoạt động trong hệ thống ứng dụng của thiết bị máy khách IM. Cụ thể, thiết bị bao gồm: bộ phận gửi yêu cầu 601, bộ phận thu thập giọng nói 602, bộ phận biến đổi âm điệu 603 và bộ phận gửi giọng nói 604.

Bộ phận gửi yêu cầu 601 được làm tương thích để thiết lập kênh truyền thông giọng nói biến đổi âm điệu.

Bộ phận thu thập giọng nói 602 được làm tương thích để thu thập thông tin giọng nói gốc đưa vào.

Bộ phận biến đổi âm điệu 603 được làm tương thích để xử lý thông tin giọng nói gốc thu thập được bởi bộ phận thu thập giọng nói thông qua việc biến đổi âm điệu để có được giọng nói biến đổi âm điệu.

Bộ phận gửi giọng nói 604 được làm tương thích để gửi giọng nói biến đổi âm điệu có được bởi bộ phận biến đổi âm điệu 603 thông qua kênh truyền thông giọng nói biến đổi âm điệu được thiết lập bởi bộ phận gửi yêu cầu 601.

Cần lưu ý rằng bộ phận gửi yêu cầu 601, bộ phận thu thập giọng nói 602, bộ phận biến đổi âm điệu 603 và bộ phận gửi giọng nói 604 có thể nằm trong cùng một thực thể, chẳng hạn máy khách IM A, hoặc nằm trong các thực thể khác, chẳng hạn như bộ phận gửi yêu cầu 601 và bộ phận thu thập giọng nói 602 ở trong cùng một thực thể như là máy khách IM A, trong khi đó bộ phận biến đổi âm điệu 603 và bộ phận gửi giọng nói 604 lại ở trong một thiết bị biến đổi âm điệu như là máy chủ. Cách thức thực hiện chi tiết phụ thuộc vào tình huống cụ thể, và không bị giới hạn trong sáng chế này.

Cụ thể, bộ phận gửi yêu cầu 601 sẽ thiết lập kênh truyền thông giọng nói biến đổi âm điệu sau khi nhận được đáp ứng để thực hiện truyền thông giọng nói biến đổi âm điệu. Đáp ứng thực hiện truyền thông giọng nói biến đổi âm điệu là đáp ứng với yêu cầu truyền thông giọng nói biến đổi âm điệu được gửi bởi bộ phận gửi yêu cầu 601. Trong phương án này, bộ phận gửi yêu cầu 601 cũng được làm tương thích để nhận thông tin của yêu cầu truyền thông giọng nói biến đổi âm điệu được đưa vào bởi người dùng.

Bộ phận thu thập giọng nói 603 còn được làm tương thích để biến đổi thông tin giọng nói thu được thành thông tin giọng nói kỹ thuật số. Thông tin giọng nói kỹ thuật số có thể được nhận dạng và xử lý bởi máy vi tính.

Trong phương án này, bộ phận biến đổi âm điệu 603 bao gồm: môđun xác định thông tin biến đổi âm điệu 6031, môđun logic dịch vụ 6032 và môđun biến đổi âm điệu 6033.

Môđun xác định thông tin biến đổi âm điệu 6031 được làm tương thích để xác định và xuất ra thông tin biến đổi âm điệu hiện hành. Thông tin biến đổi âm điệu hiện hành bao gồm thông tin lựa chọn người dùng và/hoặc thông tin biến đổi âm điệu được phép.

Môđun logic dịch vụ 6032 được làm tương thích để tạo logic lựa chọn dịch vụ, được làm tương thích để thực hiện biến đổi âm điệu và xuất ra giọng nói biến đổi âm điệu tới môđun biến đổi âm điệu 6033. Logic lựa chọn dịch vụ được xác định bởi nhà cung cấp dịch vụ IM, và xác định có bao nhiêu đối tượng dịch vụ biến đổi âm điệu

(chẳng hạn “đổi giọng nói nam thành giọng nói nữ” có thể là một đối tượng dịch vụ biến đổi âm điệu) có sẵn đối với thông tin biến đổi âm điệu được phép nhất định và môi trường truyền thông giọng nói nhất định.

Môđun biến đổi âm điệu 6033 được làm tương thích để xác định cách biến đổi âm điệu dựa trên thông tin biến đổi âm điệu thu được xuất ra bởi môđun xác định thông tin biến đổi 6031 và logic lựa chọn dịch vụ được đưa ra bởi môđun logic dịch vụ 6032, thực hiện biến đổi âm điệu thành thông tin giọng nói kỹ thuật số có được bởi bộ phận thu thập giọng nói 602 theo cách biến đổi âm điệu, và đưa ra giọng nói biến đổi âm điệu tương ứng với thông tin giọng nói kỹ thuật số. Cụ thể, môđun biến đổi âm điệu 6033 sử dụng logic lựa chọn dịch vụ để xác định cách biến đổi âm điệu dựa trên thông tin lựa chọn người dùng và/hoặc thông tin biến đổi âm điệu được phép bao gồm thông tin biến đổi âm điệu. Việc thực hiện chi tiết tương tự như đã trình bày ở phần trên, và không được mô tả lại nữa.

Để cung cấp cách biến đổi âm điệu đúng đối với người dùng nhằm đảm bảo rằng giọng nói biến đổi âm điệu có thể được nhận dạng bởi người tiếp nhận là người truyền thông cùng, bộ phận biến đổi âm điệu 603 còn bao gồm môđun lấy đặc tính người dùng 6034 theo phương án điển hình của sáng chế.

Môđun lấy đặc tính người dùng 6034 được làm tương thích để lấy thông tin đặc tính từ thông tin giọng nói lấy được bởi bộ phận thu thập giọng nói 602, tạo và đưa ra thông tin đặc tính.

Như vậy, môđun biến đổi âm điệu 6033 sử dụng logic lựa chọn dịch vụ để quyết định cách biến đổi âm điệu dựa trên thông tin lựa chọn người dùng và/hoặc thông tin biến đổi âm điệu cho phép được trích xuất ra từ thông tin biến đổi âm điệu hiện tại nhận được và còn dựa trên thông tin đặc tính người dùng nhận được.

Để cải thiện chất lượng giọng nói nghe được bởi người nhận và cung cấp đúng cách biến đổi âm điệu cho người dùng, bộ phận biến đổi âm điệu 603 còn bao gồm môđun lấy môi trường phía đối tác 6035 theo phương án điển hình khác.

Môđun lấy môi trường phía đối tác 6035 được làm tương thích để lấy thông tin môi trường giọng nói phía đối tác được chứa trong đáp ứng truyền thông giọng nói biến đổi âm điệu nhận được bởi bộ phận gửi yêu cầu 601. Trong phương án này, đáp ứng truyền thông giọng nói biến đổi âm điệu được trả về bởi bên nhận bao gồm thông tin môi trường đối tác, và bộ phận gửi yêu cầu 601 sẽ tạo ra thông tin môi trường phía đối tác dựa trên thông tin môi trường giọng nói nhận được. Sau đó môđun lấy môi trường

phía đối tác 6035 sẽ lấy thông tin môi trường phía đối tác được tạo ra bởi bộ phận gửi yêu cầu 601.

Tuy nhiên, môđun lấy đặc tính người dùng 6034 và môđun lấy môi trường phía đối tác 6035 không phải lúc nào cũng bao gồm trong thiết bị. Tốt hơn là, thiết bị trong phương án của sáng chế có thể có một hoặc cả hai môđun lấy đặc tính người dùng 6034 và môđun lấy môi trường phía đối tác 6035. Hình 6 là hình vẽ minh họa ví dụ về bộ phận biến đổi âm điệu 603 bao gồm môđun lấy đặc tính người dùng 6034 và môđun lấy môi trường phía đối tác 6035.

Như vậy, môđun biến đổi âm điệu 6033 sẽ xác định cách biến đổi âm điệu dựa trên logic lựa chọn dịch vụ được gửi bởi môđun logic dịch vụ 6032, thông tin biến đổi âm điệu hiện tại được gửi bởi môđun xác định thông tin biến đổi âm điệu 6031, và thông tin đặc tính được gửi bởi môđun lấy đặc tính người dùng 6034; hoặc dựa vào logic lựa chọn dịch vụ được gửi bởi môđun logic dịch vụ 6032, thông tin biến đổi âm điệu hiện tại được gửi bởi môđun xác định thông tin biến đổi âm điệu 6031, và thông tin môi trường phía đối tác được gửi bởi môđun lấy môi trường từ xa 6035; hoặc dựa vào logic lựa chọn dịch vụ được gửi bởi môđun logic dịch vụ 6032, thông tin biến đổi âm điệu hiện tại được gửi bởi môđun xác định thông tin biến đổi âm điệu 6031, thông tin đặc tính được gửi bởi môđun lấy đặc tính người dùng 6034, và thông tin môi trường giọng nói phía đối tác được gửi bởi môđun lấy môi trường từ xa 6035.

Để có được hiệu ứng tốt hơn trong việc truyền thông giọng nói biến đổi âm điệu và cải thiện chất lượng giọng nói được nghe bởi người nhận, thiết bị còn bao gồm bộ phận khử nhiễu 605 theo phương án khác của sáng chế.

Bộ phận khử nhiễu 605 sẽ nhận thông tin giọng nói kỹ thuật số có được bởi bộ phận thu thập giọng nói 602, thực hiện khử nhiễu, và lấy thông tin giọng nói kỹ thuật số từ phần đã khử nhiễu.

Để giảm băng thông cần thiết cho việc truyền tải dữ liệu truyền thông giọng nói biến đổi âm điệu để thực hiện truyền thông giọng nói biến đổi âm điệu theo thời gian thực, thiết bị còn bao gồm: bộ phận mã hóa 606 và/hoặc bộ phận tối ưu hóa 607 theo phương án chưa nhắc đến trong sáng chế. Hình 6 là hình vẽ minh họa ví dụ về thiết bị bao gồm bộ phận mã hóa 606 và bộ phận tối ưu hóa 607.

Bộ phận mã hóa 606 được làm tương thích để nén và mã hóa giọng nói biến đổi âm điệu có được bởi bộ phận biến đổi âm điệu 603, và lấy được luồng bit (bit streams) giọng nói biến đổi âm điệu.

Bộ phận tối ưu hóa 607 được làm tương thích để thực hiện việc tăng cường lặp lại và/hoặc nhóm lại và đóng gói các luồng bit (bit streams) giọng nói biến đổi âm điệu có được bởi bộ phận mã hóa 606, và xuất ra dữ liệu giọng nói biến đổi âm điệu sau khi được xử lý tới bộ phận gửi giọng nói 604. Bộ phận tối ưu hóa 607 được sử dụng chính để hạn chế giọng nói biến đổi âm điệu bị méo do mất các gói dữ liệu và sai lệch trong khi truyền tải trên mạng, hoặc được dùng để làm cho việc truyền tải giọng nói biến đổi âm điệu một cách thuận tiện hơn. Khi thiết bị không có bộ phận mã hóa 606, bộ phận tối ưu hóa 607 có thể thực hiện việc tăng cường lặp lại và/hoặc nhóm lại và đóng gói giọng nói biến đổi âm điệu có được bởi bộ phận biến đổi âm điệu 603, và xuất ra dữ liệu giọng nói biến đổi âm điệu đã được xử lý tới bộ phận gửi giọng nói 604.

Như được thể hiện trên Hình 6, bộ phận tối ưu hóa 607 trong phương án này bao gồm:

môđun tăng cường lặp lại 6071, được làm tương thích để thực hiện việc tăng cường lặp lại đối với luồng bit (bit streams) giọng nói biến đổi âm điệu lấy được bởi bộ phận mã hóa 606 hoặc đối với giọng nói biến đổi âm điệu lấy được bởi bộ phận biến đổi âm điệu 603, và xuất ra luồng bit (bit streams) giọng nói biến đổi âm điệu sau khi xử lý;

môđun nhóm lại và đóng gói 6072, được làm tương thích để nhóm lại và đóng gói dữ liệu giọng nói biến đổi âm điệu nhận được để có các gói dữ liệu giọng nói biến đổi âm điệu. Môđun nhóm lại và đóng gói 6072 sẽ tiếp nhận giọng nói biến đổi âm điệu hoặc luồng bit (bit streams) giọng nói biến đổi âm điệu được xuất ra tương ứng bởi bộ phận biến đổi âm điệu 603, bộ phận mã hóa 606 hoặc môđun tăng cường lặp lại 6071.

Cần lưu ý rằng bộ phận tối ưu hóa 607 chỉ bao gồm môđun tăng cường lặp lại 6071 hoặc môđun nhóm lại và đóng gói 6072.

Như được thể hiện trên Hình 6, để tiếp nhận và xử lý thông tin giọng nói, thiết bị còn bao gồm các bộ phận sau.

Bộ phận đáp ứng yêu cầu 608 được làm tương thích để thu nhận yêu cầu truyền thông giọng nói biến đổi âm điệu được gửi bởi bộ phận gửi yêu cầu 601, trả về đáp ứng truyền thông giọng nói biến đổi âm điệu, tạo và xuất ra thông tin kích hoạt thu nhận giọng nói tới bộ phận thu nhận giọng nói 609.

Bộ phận thu nhận giọng nói 609 được làm tương thích để thu nhận thông tin kích hoạt thu nhận giọng nói được xuất ra bởi bộ phận đáp ứng yêu cầu 608, nếu các gói dữ liệu thu nhận hiện thời được xử lý thông qua việc nhóm lại và đóng gói, bung các gói dữ

liệu ra theo cùng một giao thức mạng được dùng bởi phía đối tác truyền thông giọng nói, và ghép nối dữ liệu đã được nhóm để lấy và đưa ra luồng mã nén.

Bộ phận giải mã 610 được làm tương thích để giải mã dữ liệu lấy được bởi bộ phận thu nhận 609, thực chất là luồng mã nén, để tạo ra tín hiệu giọng nói.

Bộ phận khuếch đại tín hiệu giọng nói 611 được làm tương thích để giải mã dữ liệu nhận được bởi bộ phận giải mã 610, thực chất là giải mã tín hiệu giọng nói, để lấy tín hiệu giọng nói sau khi đã giải mã, và thực hiện khuếch đại tín hiệu giọng nói lấy được bằng việc giải mã để có tín hiệu giọng nói đã khuếch đại.

Bộ phận xuất ra giọng nói 612 được làm tương thích để xuất ra tín hiệu giọng nói đã khuếch đại, và có thể là tai nghe, hộp tăng âm hoặc cạc âm thanh.

Nếu các gói dữ liệu hiện nhận được bởi bộ phận thu nhận giọng nói 609 bao gồm tín hiệu lặp lại được chèn vào luồng mã nén, thiết bị còn bao gồm: bộ phận nghịch đảo lặp lại/bộ phận cho phép sai lệch 613.

Bộ phận nghịch đảo lặp lại/bộ phận cho phép sai lệch 613 được làm tương thích để loại bỏ tín hiệu lặp lại được chèn vào bởi bên đối tác truyền thông giọng nói từ luồng mã nén nhận được bởi bộ phận thu nhận giọng nói 609, và biến đổi hoặc loại bỏ dữ liệu sai lệch. Do đó, chất lượng giọng nói được cải thiện đáng kể.

Tốt hơn là, bộ phận đáp ứng yêu cầu 608, bộ phận thu nhận giọng nói 609, bộ phận giải mã 610, bộ phận khuếch đại tín hiệu giọng nói 611, bộ phận xuất ra giọng nói 612 và bộ phận nghịch đảo lặp lại/bộ phận cho phép sai lệch 613 có thể có trong thực thể truyền thông khác với thực thể mà bao gồm bộ phận gửi yêu cầu 601, bộ phận thu thập giọng nói 602, bộ phận biến đổi âm điệu 603, bộ phận gửi giọng nói 604, bộ phận khử nhiễu 605, bộ phận mã hóa 606 và bộ phận tối ưu hóa 607. Ví dụ, nếu như bộ phận gửi yêu cầu 601, bộ phận thu thập giọng nói 602, bộ phận biến đổi âm điệu 603, bộ phận gửi giọng nói 604, bộ phận khử nhiễu 605, bộ phận mã hóa 606 và bộ phận tối ưu hóa 607 nằm trong một thực thể, chẳng hạn như máy khách IM A, thì bộ phận thu nhận giọng nói 609, bộ phận giải mã 610, bộ phận khuếch đại tín hiệu giọng nói 611, bộ phận xuất ra giọng nói 612 và bộ phận nghịch đảo lặp lại/bộ phận cho phép sai lệch 613 có trong thực thể phía đối tác của máy khách IM A, chẳng hạn là máy khách IM B. Rõ ràng là, nếu bộ phận gửi yêu cầu 601, bộ phận thu thập giọng nói 602 có trong một thực thể, chẳng hạn máy khách IM A, và nếu bộ phận biến đổi âm điệu 603 và bộ phận gửi giọng nói 604 có trong thiết bị biến đổi âm điệu có sẵn, chẳng hạn máy chủ 1, thì bộ phận đáp ứng yêu cầu 608, bộ phận thu nhận giọng nói 609, bộ phận giải mã 610, bộ phận

phần khuếch đại tín hiệu giọng nói 611, bộ phận xuất ra giọng nói 612 và bộ phận nghịch đảo lặp lại/bộ phận cho phép sai lệch 613 sẽ có trong máy chủ 1 của phía đối tác, chẳng hạn là máy khách B. Phần trên chỉ là ví dụ, và không được dùng để giới hạn sáng chế này.

Theo các phương án của sáng chế, tín hiệu giọng nói thu thập được trong hệ thống IM trước tiên được xử lý thông qua việc biến đổi âm điệu, qua đó truyền thông giọng nói biến đổi âm điệu dựa trên hệ thống IM được thực hiện. Truyền thông giọng nói trong hệ thống IM được thực hiện có tính giải trí hơn, và có thể trở thành dịch vụ gia tăng giá trị mới của dịch vụ IM thông thường. Dịch vụ IM sẽ trở nên thu hút hơn đối với người dùng và như vậy có tính cạnh tranh hơn. Đồng thời cũng cung cấp những trải nghiệm dịch vụ mới đối với người dùng truyền thông giọng nói, như việc bảo vệ danh danh người dùng trong việc sử dụng truyền thông giọng nói biến đổi âm điệu.

Trên đây chỉ là phần mô tả các phương án của sáng chế và không dùng để hạn chế phạm vi bảo hộ của sáng chế này. Bất kỳ sự thay đổi hoặc thay thế tương đương hoặc cải tiến theo bản chất và phạm vi của sáng chế cũng nằm trong phạm vi bảo hộ của sáng chế này.

YÊU CẦU BẢO HỘ

1. Phương pháp truyền thông giọng nói dựa trên hệ thống tin nhắn nhanh (IM - Instant Messaging), bao gồm các bước:

a) thiết lập kênh truyền thông giọng nói biến đổi âm điệu giữa ít nhất hai máy khách IM;

b) xử lý thông tin giọng nói gốc thông qua việc biến đổi âm điệu để lấy được giọng nói biến đổi âm điệu; và truyền dẫn giọng nói biến đổi âm điệu tới máy khách IM thứ nhất trong ít nhất hai máy khách IM thông qua kênh truyền thông giọng nói biến đổi âm điệu;

trong đó phương pháp này còn bao gồm các bước:

xác định cách biến đổi âm điệu, và

thực hiện biến đổi âm điệu theo cách biến đổi âm điệu đã được xác định;

trong đó việc xác định cách biến đổi âm điệu bao gồm ít nhất một trong cách như sau:

lấy thông tin môi trường giọng nói của máy khách IM thứ nhất bằng cách phân tích các tín hiệu giọng nói thu thập được bởi mic (microphone) của máy khách IM thứ nhất;

cung cấp ít nhất hai cách biến đổi âm điệu cho người dùng và xác định cách biến đổi âm điệu theo sự lựa chọn của người dùng từ ít nhất hai cách biến đổi âm điệu và thông tin môi trường giọng nói.

2. Phương pháp theo điểm 1, trong đó bước b) được thực hiện bởi máy khách IM thứ hai trong ít nhất hai máy khách IM nằm giữa kênh truyền thông giọng nói biến đổi âm điệu đã được thiết lập, hoặc được thực hiện bởi thiết bị biến đổi âm điệu có sẵn.

3. Phương pháp theo điểm 2, trong đó bước a) được thực hiện sau khi máy khách IM thứ hai nhận được đáp ứng truyền thông giọng nói biến đổi âm điệu từ máy khách IM thứ nhất, đáp ứng truyền thông giọng nói biến đổi âm điệu là đáp lại yêu cầu truyền thông giọng nói biến đổi âm điệu được gửi bởi máy khách IM thứ hai; hoặc

trong đó kênh truyền thông giọng nói biến đổi âm điệu được thiết lập giữa máy khách IM thứ hai và máy khách IM thứ nhất sau khi máy khách IM thứ hai nhận được đáp ứng truyền thông giọng nói được trả về bởi máy khách IM thứ nhất; trong đó đáp

ứng truyền thông giọng nói là đáp lại yêu cầu truyền thông giọng nói được gửi bởi máy khách IM thứ hai.

4. Phương pháp theo điểm 1, trong đó việc xử lý thông tin giọng nói gốc thông qua việc biến đổi âm điệu tại bước b) bao gồm bước:

thu thập thông tin giọng nói gốc đưa vào, biến đổi thông tin giọng nói gốc đưa vào thành thông tin giọng nói kỹ thuật số; và xử lý thông tin giọng nói kỹ thuật số thông qua việc biến đổi âm điệu.

5. Phương pháp theo điểm 1, trong đó phương pháp này còn bao gồm bước:

xác định, trước khi xác định cách biến đổi âm điệu, thông tin biến đổi âm điệu hiện hành và logic lựa chọn dịch vụ để xác định cách biến đổi âm điệu;

trong đó việc xác định cách biến đổi âm điệu bao gồm bước: xác định cách biến đổi âm điệu bằng logic lựa chọn dịch vụ dựa trên thông tin biến đổi âm điệu hiện hành.

6. Phương pháp theo điểm 5, trong đó phương pháp này còn bao gồm bước: lấy thông tin đặc tính của thông tin giọng nói gốc trước khi xác định cách biến đổi âm điệu;

trong đó việc xác định cách biến đổi âm điệu bao gồm bước: xác định cách biến đổi âm điệu bằng logic lựa chọn dịch vụ dựa trên thông tin đặc tính và/hoặc thông tin biến đổi âm điệu hiện hành.

7. Phương pháp theo điểm 6, trong đó đáp ứng truyền thông giọng nói biến đổi âm điệu bao gồm thông tin môi trường giọng nói của máy khách IM thứ nhất;

trong đó việc xác định cách biến đổi âm điệu bao gồm bước: xác định cách biến đổi âm điệu bằng logic lựa chọn dịch vụ dựa trên ít nhất một trong số thông tin môi trường giọng nói, thông tin biến đổi âm điệu hiện hành và thông tin đặc tính.

8. Phương pháp theo điểm 4, trong đó phương pháp này còn bao gồm bước:

thực hiện khử nhiễu đối với thông tin giọng nói kỹ thuật số trước khi xử lý thông tin giọng nói kỹ thuật số thông qua việc biến đổi âm điệu.

9. Phương pháp theo điểm 1, trong đó phương pháp này còn bao gồm bước:

trước khi gửi giọng nói biến đổi âm điệu tới máy khách IM thứ nhất thông qua kênh truyền thông giọng nói biến đổi âm điệu, thực hiện nén và mã hóa và/hoặc tăng cường lặp lại cho giọng nói biến đổi âm điệu;

và/hoặc

thực hiện việc nhóm lại và đóng gói giọng nói biến đổi âm điệu.

10. Phương pháp theo điểm 1, trong đó phương pháp này còn bao gồm các bước:

thiết lập kênh truyền thông giọng nói giữa ít nhất hai máy khách IM trước khi thiết lập kênh truyền thông giọng nói biến đổi âm điệu; và

hủy bỏ kênh truyền thông giọng nói sau khi thiết lập kênh truyền thông giọng nói biến đổi âm điệu.

11. Thiết bị truyền thông giọng nói dựa trên hệ thống tin nhắn nhanh (IM) bao gồm:

bộ phận gửi yêu cầu, được làm tương thích để thiết lập kênh truyền thông giọng nói biến đổi âm điệu;

bộ phận thu thập giọng nói, được làm tương thích để thu thập thông tin giọng nói gốc được đưa vào;

bộ phận biến đổi âm điệu, được làm tương thích để xử lý thông tin giọng nói gốc thu thập được bởi bộ phận thu thập giọng nói thông qua việc biến đổi âm điệu để lấy được giọng nói biến đổi âm điệu;

bộ phận gửi giọng nói, được làm tương thích để gửi giọng nói biến đổi âm điệu lấy được bởi bộ phận biến đổi âm điệu tới thiết bị thu thông qua kênh truyền thông giọng nói biến đổi âm điệu đã được thiết lập bởi bộ phận truyền dẫn yêu cầu;

trong đó bộ phận biến đổi âm điệu được làm tương thích để xác định cách biến đổi âm điệu và thực hiện biến đổi âm điệu theo cách biến đổi âm điệu được xác định;

trong đó bộ phận biến đổi âm điệu được làm tương thích để xác định cách biến đổi âm điệu theo ít nhất một cách sau:

lấy thông tin môi trường giọng nói phía đối tác của thiết bị thu bằng cách phân tích các tín hiệu giọng nói thu thập được bởi mic (microphone) của thiết bị thu;

cung cấp ít nhất hai cách biến đổi âm điệu cho người dùng và xác định cách biến đổi âm điệu theo sự lựa chọn của người dùng từ ít nhất hai cách biến đổi âm điệu và thông tin môi trường giọng nói phía đối tác.

12. Thiết bị theo điểm 11, trong đó bộ phận thu thập giọng nói còn được làm tương thích để chuyển đổi thông tin giọng nói gốc thu thập được thành thông tin giọng nói kỹ thuật số;

bộ phận biến đổi âm điệu bao gồm:

môđun xác định thông tin biến đổi âm điệu, được làm tương thích để xác định và xuất ra thông tin biến đổi âm điệu hiện hành;

môđun logic dịch vụ, được làm tương thích để tạo và xuất ra logic lựa chọn dịch vụ được sử dụng bởi môđun biến đổi âm điệu để thực hiện việc biến đổi âm điệu;

môđun biến đổi âm điệu, được làm tương thích để xác định cách biến đổi âm điệu dựa trên thông tin biến đổi âm điệu được xuất ra bởi môđun xác định thông tin biến đổi âm điệu và dựa vào logic lựa chọn dịch vụ được đưa ra bởi môđun logic dịch vụ, thực hiện, theo cách biến đổi âm điệu, biến đổi âm điệu đối với thông tin giọng nói kỹ thuật số lấy được bởi bộ phận thu thập giọng nói, và xuất ra giọng nói biến đổi âm điệu tương ứng với thông tin giọng nói kỹ thuật số.

13. Thiết bị theo điểm 12, trong đó bộ phận biến đổi âm điệu còn bao gồm: môđun lấy đặc tính người dùng và/hoặc môđun lấy môi trường phía đối tác; trong đó

môđun lấy đặc tính người dùng được làm tương thích để lấy thông tin đặc tính từ thông tin giọng nói kỹ thuật số lấy được bởi bộ phận thu thập giọng nói, tạo và xuất ra thông tin đặc tính;

môđun lấy môi trường giọng nói phía đối tác được làm tương thích để lấy và xuất ra thông tin môi trường phía đối tác được mang trong đáp ứng truyền thông giọng nói biến đổi âm điệu nhận được bởi bộ phận gửi yêu cầu;

môđun biến đổi âm điệu được làm tương thích để xác định cách biến đổi âm điệu dựa trên thông tin biến đổi âm điệu hiện hành và thông tin đặc tính; hoặc dựa vào logic lựa chọn dịch vụ, thông tin biến đổi âm điệu hiện hành và thông tin môi trường giọng nói phía đối tác; hoặc dựa vào logic lựa chọn dịch vụ, thông tin biến đổi âm điệu hiện hành, thông tin đặc tính và thông tin môi trường giọng nói phía đối tác.

14. Thiết bị theo điểm 12, trong đó thiết bị này còn bao gồm:

bộ phận khử nhiễu, được làm tương thích để nhận thông tin giọng nói lấy được bởi bộ phận thu thập giọng nói, thực hiện khử nhiễu đối với thông tin giọng nói kỹ thuật số, và lấy thông tin giọng nói kỹ thuật số được khử nhiễu; và/hoặc

bộ phận mã hóa và/hoặc bộ phận tối ưu hóa;

trong đó bộ phận mã hóa được làm tương thích để nén và mã hóa giọng nói biến đổi âm điệu lấy được bởi bộ phận biến đổi âm điệu, và lấy các luồng bit (bit streams) giọng nói biến đổi âm điệu;

bộ phận tối ưu hóa được làm tương thích để thực hiện tăng cường lặp lại và/hoặc nhóm lại và đóng gói giọng nói biến đổi âm điệu lấy được bởi bộ phận biến đổi âm điệu hoặc các luồng bit (bit streams) giọng nói biến đổi âm điệu lấy được bởi bộ phận mã hóa, và xuất ra dữ liệu giọng nói biến đổi âm điệu lấy được bởi bộ phận tối ưu hóa thông qua việc xử lý tới bộ phận gửi giọng nói.

15. Phương pháp truyền thông giọng nói dựa trên hệ thống tin nhắn nhanh (IM), bao gồm các bước:

thiết lập kênh truyền thông giọng nói giữa ít nhất hai máy khách IM;

xử lý thông tin giọng nói gốc thông qua việc biến đổi âm điệu để lấy được giọng nói biến đổi âm điệu sau khi xác định để thực hiện việc truyền thông giọng nói biến đổi âm điệu; và truyền dẫn giọng nói biến đổi âm điệu tới máy khách IM thứ nhất trong ít nhất hai máy khách IM thông qua kênh truyền thông giọng nói;

trong đó phương pháp này còn bao gồm các bước:

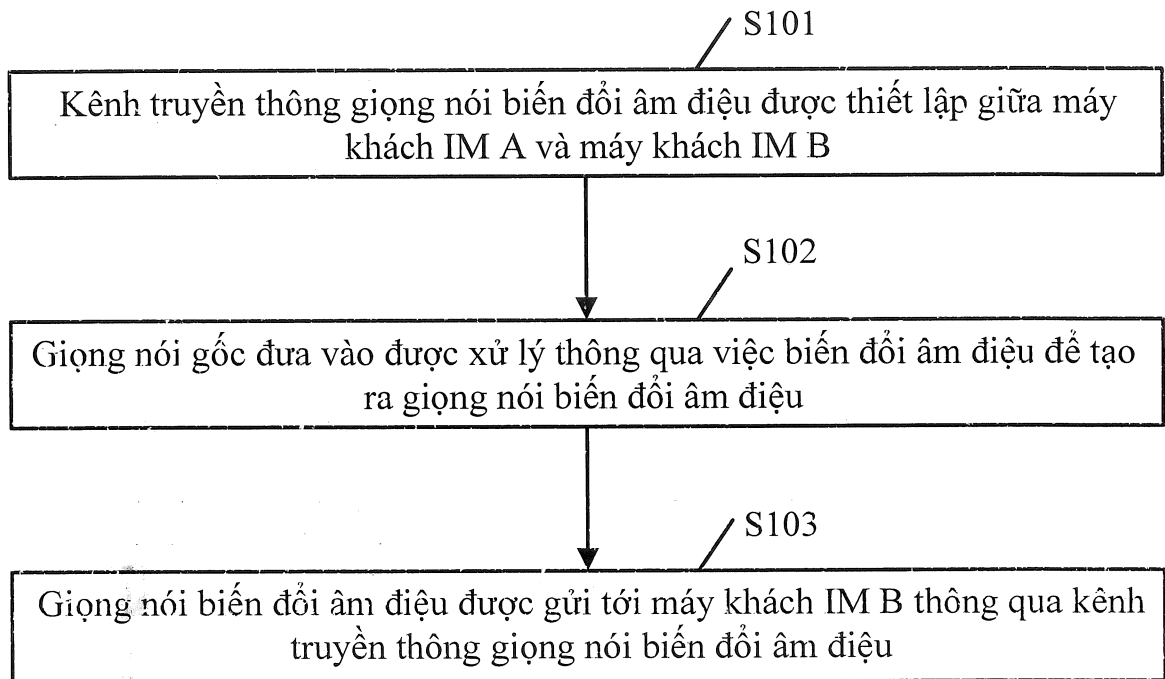
xác định cách biến đổi âm điệu; và

thực hiện biến đổi âm điệu theo cách biến đổi âm điệu được chọn;

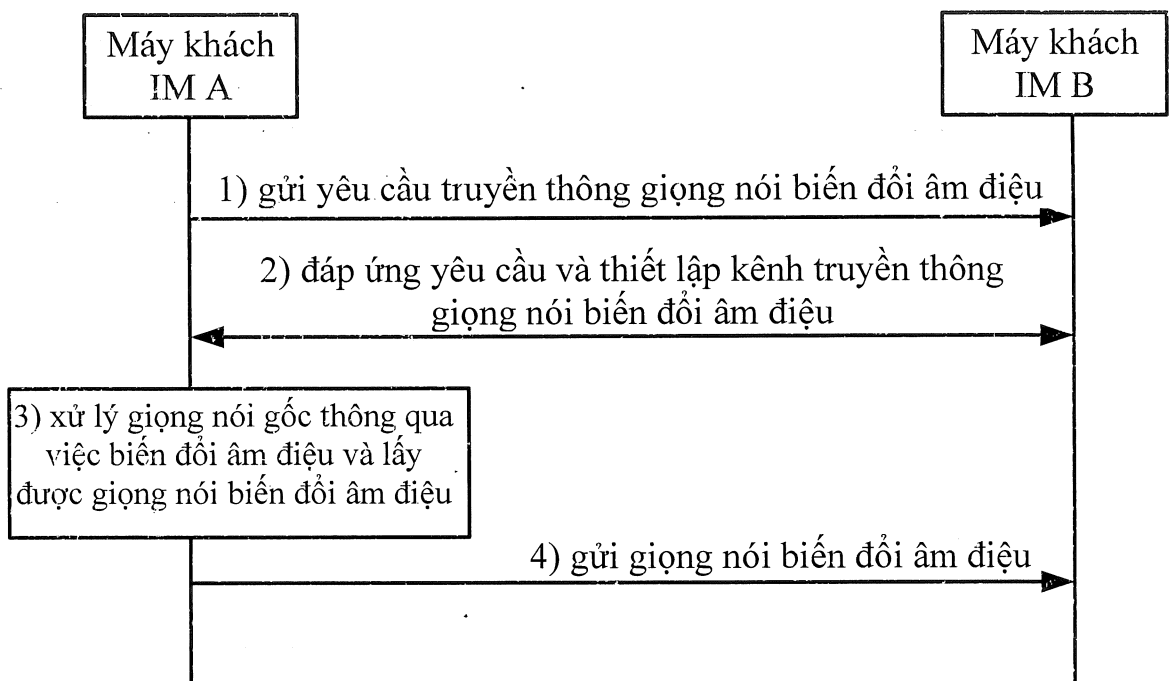
trong đó việc xác định cách biến đổi âm điệu bao gồm các bước:

lấy thông tin môi trường giọng nói của máy khách IM thứ nhất bằng cách phân tích các tín hiệu giọng nói thu thập được bằng mic (microphone) của máy khách IM thứ nhất;

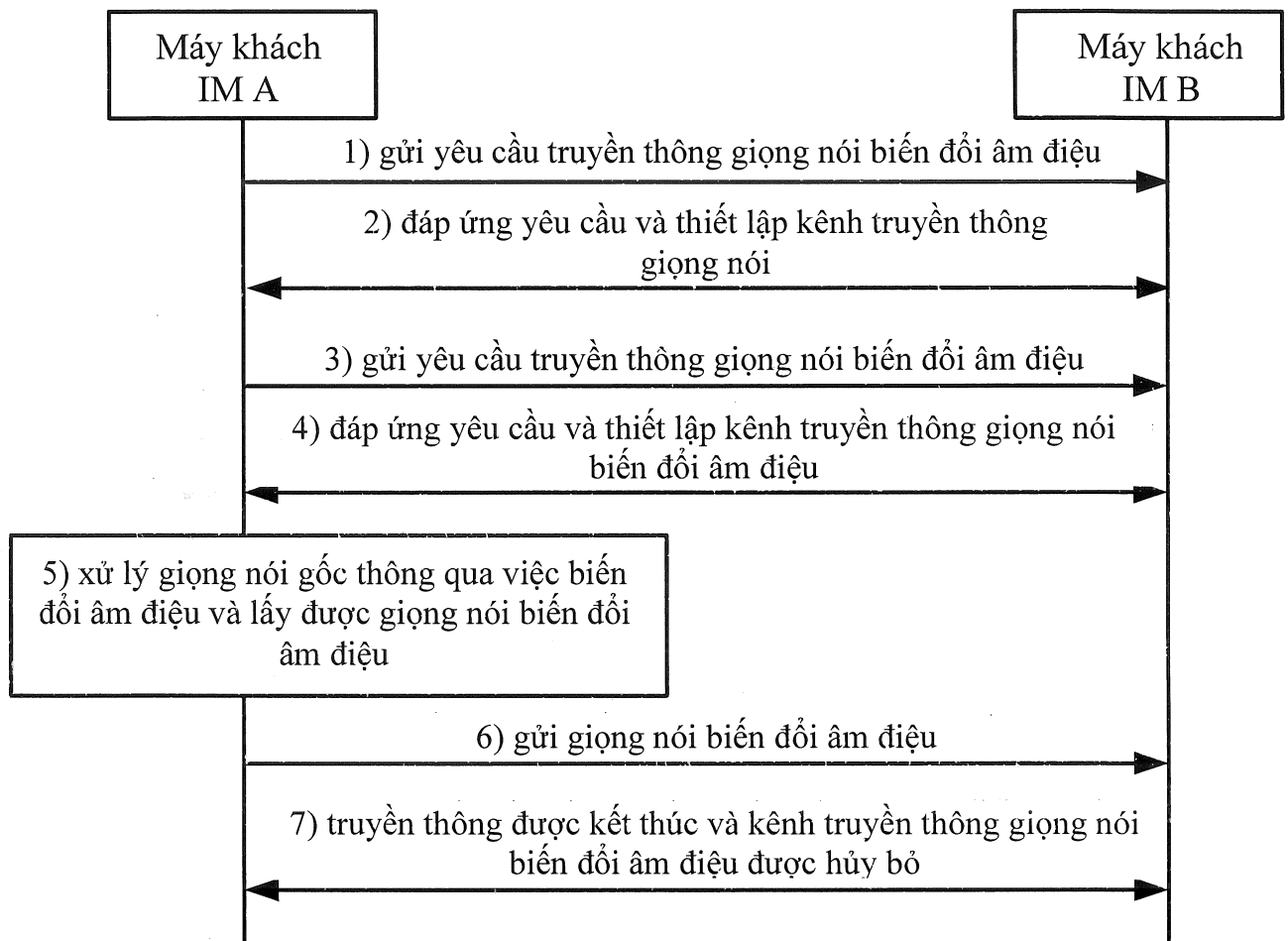
cung cấp tối thiểu hai cách biến đổi âm điệu cho người dùng và xác định cách biến đổi âm điệu theo sự lựa chọn của người dùng từ ít nhất hai cách biến đổi âm điệu và thông tin môi trường giọng nói.

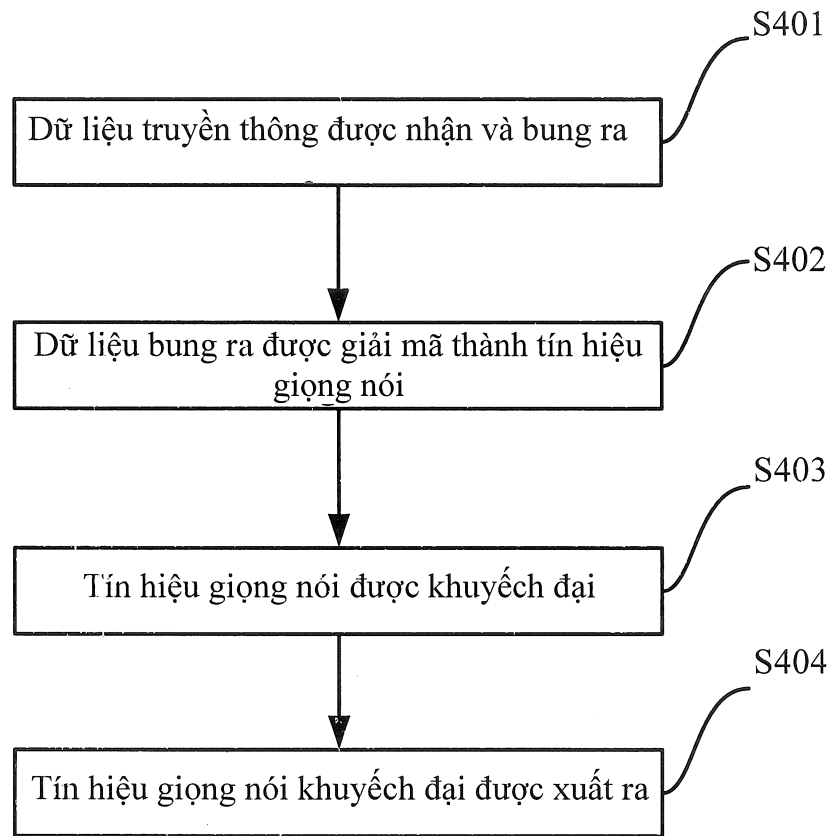
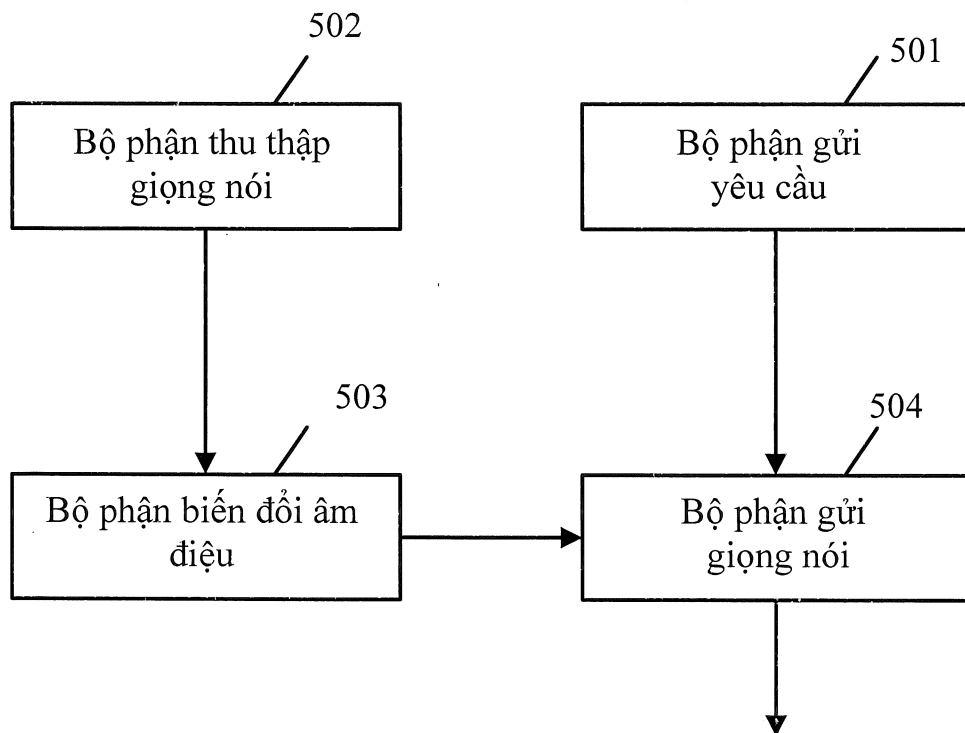


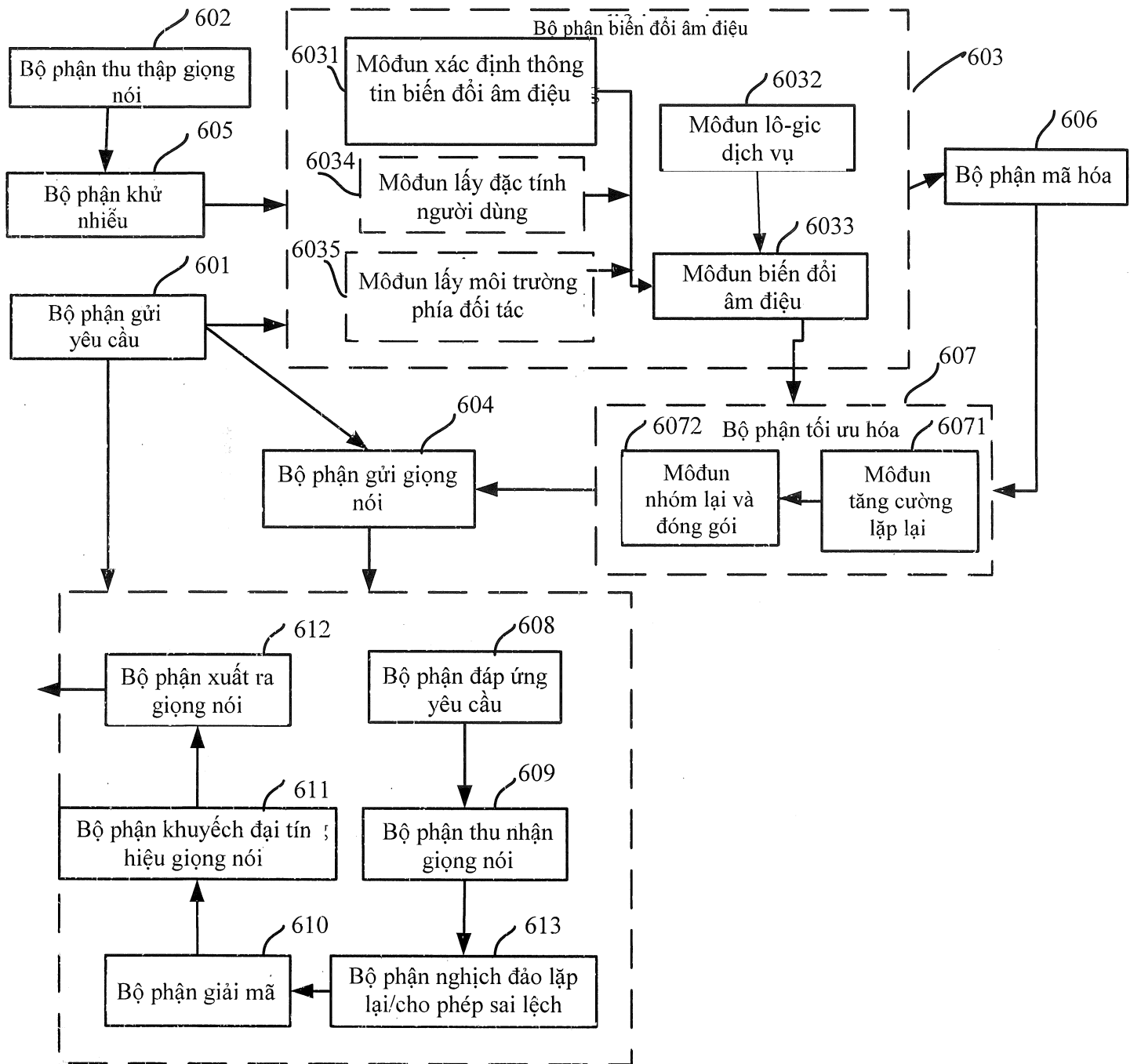
Hình 1



Hình 2

**Hình 3**

**Hình 4****Hình 5**



Hình 6