



(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)



1-0025527

(51)⁷ G10L 25/51; G10L 15/06

(13) B

(21) 1-2017-00873

(22) 15/08/2014

(86) PCT/SG2014/000385 15/08/2014

(87) WO 2016/024914 18/02/2016

(45) 25/09/2020 390

(43) 25/07/2017 352A

(73) IQ-HUB PTE. LTD. (SG)

101 CECIL STREET, #11-10 TONG ENG BUILDING, SINGAPORE 069533

(72) TOPOLEWSKI, DAVID L. (CA); SCHOLZ, KARL W. (US).

(74) Công ty TNHH Trà và cộng sự (TRA & ASSOCIATES CO.,LTD)

(54) PHƯƠNG PHÁP VÀ HỆ THỐNG GIÚP CẢI THIỆN GIỌNG NÓI CỦA NGƯỜI DÙNG THEO NGÔN NGỮ CHỈ ĐỊNH

(57) Sáng chế đề cập đến hệ thống và phương pháp giúp cải thiện giọng nói của người dùng theo ngôn ngữ chỉ định, phương pháp bao gồm: nhận văn bản theo ngôn ngữ chỉ định, hoặc theo ngôn ngữ khác từ người dùng được người dùng nói theo ngôn ngữ chỉ định; xử lý văn bản để thu được một hoặc nhiều đặc tính giọng nói văn bản theo ngôn ngữ chỉ định; nhận âm thanh nói văn bản của người dùng theo ngôn ngữ chỉ định; xử lý âm thanh để thu được một hoặc nhiều đặc tính nói văn bản của người dùng; so sánh một hoặc nhiều đặc tính nói văn bản của người dùng với một trong số một hoặc nhiều đặc tính kỳ vọng tương ứng của giọng nói văn bản để xác định độ chính xác về việc nói văn bản của người dùng; và đưa ra phản hồi đến người dùng biểu thị độ chính xác.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến phương pháp và hệ thống giúp cải thiện giọng nói của người dùng theo một ngôn ngữ chỉ định; cụ thể hơn sáng chế đề cập đến việc nhận văn bản theo ngôn ngữ chỉ định hoặc theo ngôn ngữ khác từ người dùng được người dùng nói theo ngôn ngữ chỉ định.

Sáng chế có ứng dụng đặc biệt nhưng không phải là ứng dụng riêng trong việc so sánh các đặc tính nói văn bản của người dùng với các đặc tính kỳ vọng tương ứng của giọng nói văn bản để xác định độ chính xác về việc nói văn bản của người dùng và đưa ra phản hồi đến người dùng biểu thị độ chính xác của chúng.

Tình trạng kỹ thuật của sáng chế

Thông thường, nhiều người hy vọng học một ngôn ngữ mong muốn sẽ tham gia một khóa học phù hợp với giáo viên và các sinh viên khác trong khóa học. Trong khóa học, ví dụ, giáo viên đưa cho sinh viên một số văn bản để đọc và sau đó đưa ra phản hồi đến các sinh viên về cách họ nói văn bản đó so sánh với giọng nói kỳ vọng của văn bản đó. Sự phản hồi cũng có thể bao gồm các bình luận liên quan đến các đặc tính cụ thể của giọng nói văn bản, như sự lưu loát và việc phát âm. Ngoài ra, giáo viên có thể đưa ra các bài tập cho sinh viên để thực hành tại nhà. Tuy nhiên, trong trường hợp này, sinh viên không nhận được bất kỳ phản hồi nào ở nhà và có thể, trong một số trường hợp, phát triển các thói quen xấu làm giảm khả năng học và cải thiện giọng nói ngôn ngữ mong muốn của sinh viên.

Trong ví dụ hiện tại, các hệ thống phản hồi giọng nói điện tử được sử dụng để loại bỏ nhu cầu về một giáo viên hữu hình đối với các sinh viên học giọng nói một ngôn ngữ mong muốn. Trong ví dụ hiện tại, các từ có các đặc tính giọng nói đã biết, như phát âm, được trình bày đến một người dùng để người dùng đọc lại bằng miệng. Hệ thống phản hồi nhận âm thanh của giọng nói văn bản của người dùng về các từ xác định trước và ứng dụng các thuật toán nhận dạng giọng nói để xác định liệu người dùng đã nói các từ đó chính xác hay không. Tuy nhiên, trong ví dụ này, người dùng có thể tìm ra các từ lặp lại và đã xác định trước dài dòng và không liên quan, vì vậy, làm

mất tiềm tàng sự hứng thú trong việc cải thiện giọng nói của người dùng trong ngôn ngữ mong muốn.

Bản chất kỹ thuật của sáng chế

Theo khía cạnh thứ nhất của sáng chế, sáng chế đề xuất một phương pháp giúp cải thiện giọng nói của một người dùng theo một ngôn ngữ chỉ định, phương pháp bao gồm: nhận từ người dùng văn bản theo ngôn ngữ chỉ định, hoặc theo một ngôn ngữ khác được nói bởi người dùng theo ngôn ngữ chỉ định; xử lý văn bản để thu được một hoặc nhiều đặc tính kỳ vọng về giọng nói văn bản theo ngôn ngữ chỉ định; nhận âm thanh nói văn bản của người dùng theo ngôn ngữ chỉ định; xử lý âm thanh để thu được một hoặc nhiều đặc tính nói văn bản của người dùng; so sánh một hoặc nhiều đặc tính nói văn bản của người dùng đã nói với một hoặc nhiều đặc tính kỳ vọng của giọng nói văn bản tương ứng để xác định độ chính xác về việc nói văn bản của người dùng; và đưa ra phản hồi đến người dùng biểu thị độ chính xác.

Trong một phương án thực hiện, một hoặc nhiều đặc tính kỳ vọng của giọng nói văn bản và một hoặc nhiều đặc tính nói văn bản của người dùng bao gồm ít nhất một trong các đặc trưng về phát âm, sự lưu loát và các ngôn điệu. Các đặc trưng ngôn điệu của giọng nói bao gồm sự biến đổi trong độ dài âm tiết, âm lượng và cường độ. Những người có hiểu biết trung bình về lĩnh vực kỹ thuật tương ứng đánh giá cao rằng sự phát âm, sự lưu loát và các đặc trưng ngôn điệu kỳ vọng có thể được xác định bằng cách phân tích thực nghiệm giọng nói của người nói theo ngôn ngữ chỉ định. Thật vậy, các đặc trưng ngôn điệu có thể là biểu thị về trọng âm của ngôn ngữ chỉ định. Trong trường hợp này, ví dụ, các đặc trưng âm điệu của lời nói giọng người Mỹ ở Trung Tây được xác định bằng cách phân tích thực nghiệm và phương pháp có thể xác định và đưa ra phản hồi đến người dùng biểu thị độ chính xác của người dùng nói giọng người Mỹ ở Trung Tây.

Những người có hiểu biết trung bình về lĩnh vực kỹ thuật tương ứng sẽ đánh giá cao rằng việc xác định độ chính xác về việc nói văn bản của người dùng bao gồm việc xác định một thước đo chất lượng (ví dụ từ 1/100 đến 100/100) nói văn bản của người dùng dựa vào tổng các xác định độ chính xác trên mỗi đặc tính nói văn bản của người dùng với một trong các đặc tính kỳ vọng tương ứng của giọng nói văn bản. Theo đó, trong phương án thực hiện, phương pháp xác định độ chính xác về việc nói văn bản

của người dùng bằng cách xác định độ chính xác của mỗi đặc tính chỉ định, như tốc độ, xương âm, năng lực, phát âm, sự lưu loát v.v. về việc nói văn bản của người dùng.

Trong một phương án thực hiện, phương pháp bao gồm thêm việc phân tích cú pháp văn bản thành một hoặc nhiều câu hoặc đoạn các từ. Ví dụ, một tập hợp các từ được phân tích cú pháp thành một hoặc nhiều câu bằng cách tìm kiếm lần lượt văn bản có các từ đó đối với sự xuất hiện của các dấu chấm hết câu không trích dẫn, như dấu chấm câu, dấu chấm hỏi và dấu chấm than.

Ngoài ra, tập hợp các từ được phân tích cú pháp thành tập hợp các đoạn đếm từ cố định thường có quy mô từ hai đến 8 từ. Sau đó các đoạn được đưa ra và vì vậy được dạy cho người dùng một cách liên tiếp, đầu tiên dạy một đoạn và sau đó bổ sung đoạn thứ hai v.v. cho đến khi dạy được toàn bộ tập hợp các từ được dạy. Trong một sự thay đổi, các đoạn có thể được xây dựng từ trái sang phải từ lúc bắt đầu tập hợp từ cho đến khi toàn bộ tập hợp được dạy. Ngoài ra, các đoạn có thể được xây dựng từ phải sang trái từ cuối của tập hợp từ phía đầu của tập hợp từ; ví dụ, nếu đếm từ của đoạn là ba, đầu tiên ba chữ cuối của tập hợp sẽ được dạy, sau đó là sáu chữ cuối, tiếp theo là chín chữ cuối v.v. cho đến khi kết thúc tập hợp.

Trong một phương án thực hiện khác, một ngữ pháp của một hoặc nhiều câu văn bản cũng được tạo ra. Những người có hiểu biết trung bình về lĩnh vực tương ứng sẽ đánh giá cao rằng ngữ pháp (ví dụ: các ngữ pháp) là một ngữ pháp nhận dạng giọng nói; đó là, cấu trúc ngữ pháp chính thức có thể được nhận dạng bởi một thiết bị nhận dạng giọng nói thực hiện các thuật toán nhận dạng giọng nói. Trong một ví dụ, lời nhắc cho người dùng để nói một câu đầu tiên trong số các câu được cung cấp đến người dùng, như qua văn bản của câu đầu tiên trong số các câu được hiển thị trên một màn hình của thiết bị người dùng. Phương pháp bao gồm thêm việc nhận âm thanh của người dùng nói câu đầu tiên trong số các câu văn bản sau đó đưa ra phản hồi đến người dùng biểu thị chính xác của người dùng nói câu đầu tiên trong số các câu. Sau khi đưa ra phản hồi, người dùng sau đó được nhắc để nói câu thứ hai trong số các câu.

Vì vậy, trong một ví dụ về việc sử dụng, đầu tiên người dùng nhập văn bản vào theo ngôn ngữ chỉ định mà họ muốn học để nói, hoặc theo ngôn ngữ khác và được dịch thành ngôn ngữ chỉ định để hiển thị cho người dùng. Sau đó văn bản được phân tích cú pháp thành các câu. Sau đó, người dùng, mong muốn cải thiện giọng nói của họ

theo ngôn ngữ chỉ định được nhắc để nói câu văn bản đầu tiên đã hiển thị. Âm thanh của người dùng nói câu được nhận và xử lý bằng thiết bị nhận dạng giọng nói sao cho người dùng có thể nhận phản hồi về việc phát âm, sự lưu loát của họ về câu v.v.. Sau khi nhận phản hồi, sau đó người dùng có thể đến câu tiếp theo và nhận được phản hồi tiếp v.v.

Ví dụ, phương pháp sử dụng tiêu chí ngữ pháp nhận dạng giọng nói (SRGS), để điều khiển tất cả các hoạt động nhận dạng cùng với một thiết bị nhận dạng giọng nói thực hiện các thuật toán nhận dạng giọng nói. SRGS là một tiêu chuẩn của Hiệp hội mạng lưới toàn cầu (W3C) cho các ngữ pháp nhận dạng giọng nói. Sẽ được đánh giá cao rằng một ngữ pháp nhận dạng giọng nói là một tập hợp các mẫu từ hướng dẫn thiết bị nhận dạng giọng nói về kỳ vọng những gì mà con người nói.

Trong một phương án thực hiện, phương pháp bao gồm thêm việc xử lý văn bản để diễn tả văn bản thành âm thanh theo ngôn ngữ chỉ định và sau đó đưa âm thanh đến người dùng qua một loa của thiết bị người dùng. Đó là, trong phương án thực hiện này, thiết bị tạo thành giọng văn bản nói được để giúp người dùng cải thiện giọng nói theo ngôn ngữ chỉ định. Ví dụ, khi sử dụng, người dùng nói một câu, nhận phản hồi về độ chính xác bằng cách người dùng nói câu và sau đó nghe giọng nói kỳ vọng của câu để so sánh. Thật vậy, trình tự sử dụng phương pháp có thể là, ví dụ (a): nhập văn bản, nhập giọng nói sau đó cung cấp đầu ra giọng nói kỳ vọng hoặc (b): nhập văn bản, cung cấp đến đầu ra giọng nói kỳ vọng, sau đó nhập giọng nói để so sánh.

Những người có hiểu biết trung bình về lĩnh vực kỹ thuật tương ứng sẽ đánh giá cao rằng bước xử lý âm thanh để thu được một hoặc nhiều đặc tính người dùng nói văn bản được thực hiện sử dụng các thuật toán nhận dạng giọng nói và bước xử lý văn bản để thu được một hoặc nhiều đặc tính giọng nói kỳ vọng của văn bản được thực hiện sử dụng các thuật toán tổng hợp giọng nói.

Tốt hơn, độ chính xác bao gồm nhiều hơn một giá trị tin cậy kết hợp với độ chính xác của một hoặc nhiều đặc tính nói văn bản của người dùng tương ứng với một hoặc nhiều đặc tính kỳ vọng giọng nói văn bản. Giá trị tin cậy có thể, ví dụ, là sự sắp xếp ĐẠT/KHÔNG ĐẠT dựa vào đặc tính đã nói của người dùng giống như thế nào với các đặc tính kỳ vọng. Ví dụ, đường cơ sở 70% phù hợp giữa sự lưu loát và phát âm đã nói và sự lưu loát và phát âm kỳ vọng của một ai đó nói văn bản được thiết lập và

giá trị ĐẠT được xác định nếu sự lưu loát đã nói và sự phát âm văn bản vượt quá 70% phù hợp. Trong một ví dụ, giá trị tin cậy kết hợp với độ chính xác bao gồm: CHÍNH XÁC CAO (ví dụ, phù hợp 80%+); màu cam cho CHÍNH XÁC KHÔNG ĐÁNG KÊ (ví dụ, phù hợp 50%-80%) và KÉM (ví dụ, phù hợp 20%-50%) và RẤT KÉM (ví dụ, phù hợp 0%-20%). Tham chiếu đến ví dụ này, sự phản hồi cũng bao gồm các màu sắc biểu thị các giá trị khác nhau của giá trị tin cậy về độ chính xác. Ví dụ, màu xanh lá cây được biểu thị trên màn hình của thiết bị người dùng cho CHÍNH XÁC CAO, màu cam cho CHÍNH XÁC KHÔNG ĐÁNG KÊ, màu nâu cho KÉM, và màu đỏ cho RẤT KÉM. Có thể dự kiến rằng sự phản hồi âm thanh cũng có thể được cung cấp đến người dùng, như đưa ra giọng nói “RẤT KÉM” qua loa của thiết bị người dùng. Các hình thức phản hồi khác bao gồm loại số (ví dụ 1-10), loại chữ (ví dụ A-F), dấu hiệu, hoặc các chỉ số hình ảnh phản hồi khác.

Trong phương án thực hiện khác, phương pháp bao gồm thêm việc xử lý văn bản để thu được các từ cần xác định trước trong văn bản sao cho các từ cần có thể được kiểm duyệt. Ví dụ, các từ chửi thề và các từ biểu thị giọng nói thù hận v.v. được xác định trước và lưu trong bộ nhớ sao cho có thể tiếp cận được trước khi thực hiện bước phân tích cú pháp văn bản thành các câu.

Trong một phương án thực hiện, phương pháp bao gồm thêm việc ghi âm thanh nói văn bản của người dùng theo ngôn ngữ chỉ định trong bộ nhớ. Trong cách này, phương pháp có thể, chẳng hạn, xử lý các thí dụ âm thanh ghi trong cơ sở dữ liệu để thu được một hoặc nhiều đặc tính kỳ vọng của giọng nói văn bản. Đó là, ví dụ, sự phát âm kỳ vọng văn bản có thể được xác định bằng cách phân tích các bản ghi âm của nhiều người dùng nói văn bản. Ngoài ra, giọng nói đã ghi âm của một người dùng cụ thể có thể được sử dụng cho phân tích sau, như chỉ ra sự tiến bộ của người dùng đó nói ngôn ngữ chỉ định. Trong ví dụ tiếp theo, tất cả văn bản được người dùng nhập và nhận được giữ lại để phân tích ngoại tuyến. Đó là, việc thu thập chuỗi văn bản được tìm kiếm để lặp lại (ví dụ: nhập từ các người dùng khác nhau) và các từ hoặc chủ đề cụ thể phổ biến với nhiều người dùng.

Theo một khía cạnh khác của sáng chế, sáng chế đề xuất một hệ thống giúp cải thiện giọng nói của người dùng theo ngôn ngữ chỉ định, hệ thống bao gồm bộ xử lý có: mô đun nhập được cấu hình để: nhận văn bản theo ngôn ngữ chỉ định, hoặc theo ngôn ngữ khác từ người dùng được người dùng nói theo ngôn ngữ chỉ định; và nhận âm

thanh nói văn bản của người dùng theo ngôn ngữ chỉ định; mô đun xử lý được cấu hình để: xử lý văn bản để thu được một hoặc nhiều đặc tính kỳ vọng của giọng nói văn bản theo ngôn ngữ chỉ định, xử lý âm thanh để thu được một hoặc nhiều đặc tính nói văn bản của người dùng; và so sánh một hoặc nhiều đặc tính nói văn bản của người dùng với một trong số một hoặc nhiều các đặc tính kỳ vọng giọng nói văn bản đã nói tương ứng để xác định độ chính xác về việc nói văn bản của người dùng; và mô đun xuất cấu hình để đưa ra phản hồi đến người dùng biểu thị độ chính xác.

Theo một khía cạnh khác của sáng chế, sáng chế đề xuất hệ thống giúp cải thiện giọng nói của người dùng theo ngôn ngữ chỉ định, hệ thống bao gồm: một màn hình được cấu hình để hiển thị văn bản theo ngôn ngữ chỉ định, hoặc theo ngôn ngữ khác; phương tiện nhập văn bản được cấu hình để người dùng nhập văn bản đã nói theo ngôn ngữ chỉ định, hoặc theo ngôn ngữ khác đã nói bởi người dùng được nói theo ngôn ngữ chỉ định; một micrô được cấu hình để nhập âm thanh nói văn bản của người dùng theo ngôn ngữ chỉ định; và một bộ xử lý có: mô đun nhập được cấu hình để: nhận văn bản được nhập theo ngôn ngữ chỉ định đã nói, hoặc theo ngôn ngữ khác đã nói; và nhận âm thanh nói văn bản của người dùng theo ngôn ngữ chỉ định; một mô đun xử lý được cấu hình để: xử lý văn bản để thu được một hoặc nhiều đặc tính lời nói đoán trước của văn bản theo ngôn ngữ chỉ định; xử lý âm thanh để thu được một hoặc nhiều đặc tính nói văn bản của người dùng; và so sánh một hoặc nhiều đặc tính kỳ vọng giọng nói đã nói tương ứng của văn bản để xác định độ chính xác về việc nói văn bản của người dùng, và mô đun xuất được cấu hình để đưa ra phản hồi đến người dùng biểu thị độ chính xác.

Trong một phương án thực hiện, hệ thống bao gồm thêm một máy chủ có bộ xử lý đề cập trên đây trong truyền thông dữ liệu qua một mạng với thiết bị người dùng bao gồm màn hình, phương tiện nhập dữ liệu vào và micrô. Đó là, trong phương án thực hiện này, người dùng có một thiết bị người dùng có chức năng nhập và xuất (ví dụ: máy tính bảng, máy tính cá nhân, hoặc điện thoại thông minh) đó là truyền thông dữ liệu qua một mạng (ví dụ: mạng Internet) với máy chủ lưu trữ bộ xử lý. Theo đó, người dùng nhập văn bản được nói theo ngôn ngữ chỉ định, bằng cách có thể là, văn bản hoặc dán điện tử văn bản từ một tài liệu vào hộp văn bản đã phân bổ. Văn bản được truyền thông qua mạng đến máy chủ để phân tích các cú pháp văn bản thành các

câu và đưa ra câu nói bởi người dùng qua mạng được hiển thị trên màn hình của thiết bị người dùng. Sau đó người dùng có thể nói câu đầu tiên và âm thanh của người dùng đang nói được thu nạp bởi micro và được truyền đến máy chủ để xử lý sao cho phản hồi có thể được xác định và được đưa ra qua, lời nói, màn hình của thiết bị người dùng đến người dùng như mô tả trên đây. Đó là, màn hình của thiết bị người dùng được cấu hình thêm để hiển thị phản hồi đến người dùng.

Trong một phương án ưu tiên khác, thiết bị người dùng bao gồm màn hình, phương tiện nhập văn bản và micrô và bộ xử lý. Vì vậy, trong phương án thực hiện này; việc xử lý được thực hiện tại chỗ trên thiết bị người dùng.

Tốt hơn, mô đun xử lý được cấu hình thêm để phân tích cú pháp văn bản thành một hoặc nhiều câu và mô đun xuất được cấu hình thêm để đưa ra lời nhắc cho người dùng để nói câu đầu tiên trong số các câu. Ngoài ra, mô đun nhập được cấu hình thêm để nhận âm thanh của người dùng nói câu đầu tiên trong số các câu của văn bản và mô đun xuất được cấu hình thêm để đưa ra phản hồi đến người dùng biểu thị độ chính xác của một hoặc nhiều đặc tính của người dùng nói câu đầu tiên trong số các câu của văn bản tương ứng với một hoặc nhiều đặc tính kỳ vọng giọng nói của câu đầu tiên trong số các câu của văn bản.

Ngoài ra, mô đun xử lý được cấu hình thêm để tạo ra một ngữ pháp của một hoặc nhiều câu của văn bản và lời nhắc bao gồm văn bản của câu đầu tiên trong số các câu.

Trong một phương án thực hiện, mô đun xử lý được cấu hình thêm để xử lý văn bản thu được một giọng nói kỳ vọng của văn bản theo ngôn ngữ chỉ định. Trong phương án thực hiện, mô đun xuất được cấu hình để đưa ra giọng nói kỳ vọng của văn bản được đưa đến người dùng qua loa của thiết bị người dùng. Ví dụ, mô đun xử lý bao gồm một mô đun tổng hợp giọng nói để xử lý văn bản để thu được một giọng nói kỳ vọng được đưa ra đến người dùng, ví dụ sau khi người dùng nói câu của họ, sao cho người dùng có thể so sánh và cải thiện thêm giọng nói của họ.

Trong một phương án thực hiện, mô đun xử lý được cấu hình thêm để xử lý văn bản để thu được các từ cấm xác định trước trong văn bản sao cho các từ cấm có thể được kiểm duyệt như đã mô tả. Các từ cấm được lưu trong bộ nhớ có thể được đặt từ xa với bộ xử lý và có thể truy cập qua một mạng hoặc có thể được đặt tại chỗ.

Theo một khía cạnh khác của sáng chế, sáng chế đề xuất mã chương trình máy tính mà khi chạy sẽ thực hiện các phương pháp trên đây.

Theo một khía cạnh khác của sáng chế, sáng chế đề xuất phương tiện đọc được bằng máy tính hữu hình bao gồm mã chương trình trên đây.

Theo một khía cạnh khác của sáng chế, sáng chế đề xuất tập tin dữ liệu bao gồm mã chương trình trên đây.

Mô tả vắn tắt các hình vẽ

Để có thể hiểu rõ hơn về sáng chế, ví dụ của các phương án thực hiện sẽ được mô tả cùng với việc tham chiếu đến các hình vẽ kèm theo, trong đó:

Fig.1 là lưu đồ của phương pháp giúp cải thiện giọng nói của người dùng theo ngôn ngữ chỉ định theo một phương án thực hiện của sáng chế;

Fig.2 là hình vẽ dưới dạng giản đồ của hệ thống giúp cải thiện giọng nói của người dùng theo ngôn ngữ chỉ định theo phương án thực hiện của sáng chế; và

Fig.3 là hình vẽ dưới dạng giản đồ của hệ thống trên Fig.2 thể hiện hệ thống truyền thông với thiết bị của người dùng qua một mạng.

Mô tả chi tiết sáng chế

Theo một phương án thực hiện của sáng chế, sáng chế đề xuất phương pháp 10 giúp cải thiện giọng nói của người dùng theo ngôn ngữ chỉ định, như thể hiện trên Fig.1. Phương pháp 10 bao gồm các bước nhận văn bản theo ngôn ngữ chỉ định 12, hoặc bằng ngôn ngữ khác, từ người dùng được người dùng nói theo ngôn ngữ chỉ định, xử lý văn bản để thu được một hoặc nhiều đặc tính kỳ vọng giọng nói văn bản theo ngôn ngữ chỉ định 14, nhận âm thanh nói văn bản của người dùng theo ngôn ngữ chỉ định 16, xử lý âm thanh để thu được một hoặc nhiều đặc tính nói văn bản của người dùng với một hoặc nhiều đặc tính kỳ vọng, giọng nói văn bản 18, so sánh một hoặc nhiều đặc tính nói văn bản của người dùng với một hoặc nhiều đặc tính kỳ vọng giọng nói văn bản để xác định sự chính xác về việc nói văn bản của người dùng, và đưa ra phản hồi đến người dùng biểu thị độ chính xác 20.

Như mô tả trên đây, một hoặc nhiều đặc tính kỳ vọng giọng nói văn bản và các đặc tính nói văn bản của người dùng bao gồm ít nhất một trong số các đặc trưng về phát âm, sự lưu loát, và ngôn điệu. Các đặc trưng ngôn điệu của giọng nói bao gồm sự

biến đổi về độ dài âm tiết của từ, âm lượng và cường độ. Vì vậy, trong ví dụ về việc sử dụng, sự lưu loát và, có thể là, phát âm của giọng nói kỳ vọng của văn bản đã nhập được thu từ văn bản và so sánh với sự lưu loát và phát âm thu được từ âm thanh nói văn bản của người dùng đã nhận để xác định độ chính xác về việc nói văn bản của người dùng. Độ chính xác của người dùng, về phát âm và sự lưu loát, sau đó được phản hồi lại người dùng để giúp người dùng cải thiện giọng nói người dùng theo ngôn ngữ chỉ định (ví dụ: tiếng Anh).

Trong một phương án thực hiện khác của sáng chế, sáng chế đề xuất một hệ thống 24, như thể hiện trên Fig.2, giúp cải thiện giọng nói của người dùng theo ngôn ngữ chỉ định để thực hiện phương pháp 10. Hệ thống 24 bao gồm bộ xử lý 26 có một số mô đun để thực hiện phương pháp 10. Cụ thể, bộ xử lý 26 bao gồm mô đun nhập 28, mô đun xử lý 30 và mô đun xuất 32. Bộ xử lý 26 được bố trí để nhận và truyền thông tin qua, có thể là, một mạng và/hoặc giữa các bộ phận khác của hệ thống 24, như bộ nhớ 52 (thể hiện trên Fig.3) qua kênh truyền thông 34. Trong phương án thực hiện, bộ xử lý 26 được thực thi bằng máy tính truyền thông qua kênh truyền thông 34 với các thiết bị nhập và các thiết bị xuất có trong thiết bị của người dùng 38 (thể hiện trên Fig.3). Như mô tả trên đây, tuy nhiên, có thể dự tính được rằng bộ nhớ 52 cũng có thể lưu trữ trên, có thể là, máy chủ xa bộ xử lý 26 có thể truy cập được qua mạng. Trong mọi trường hợp, những người có hiểu biết trung bình về lĩnh vực tương ứng sẽ đánh giá cao rằng mô đun nhập 28 và mô đun xuất 32 có các giao diện phù hợp tiếp xúc với mạng, các mô đun trong hệ thống 24, và thiết lập kênh truyền thông 34. Hơn nữa, người ta cũng đánh giá rằng các thiết bị nhập và các thiết bị xuất không cần có trong cùng thiết bị của người dùng 38. Ví dụ, văn bản được nói bởi người dùng có thể được hiển thị trên tivi liên lạc với hệ thống 24 qua mạng trong khi âm thanh được ghi bởi người dùng sử dụng điện thoại thông minh liên lạc với hệ thống 24 qua mạng.

Trong mọi trường hợp, như mô tả trên đây, mô đun nhập 28 được cấu hình để nhận văn bản theo ngôn ngữ chỉ định, hoặc theo ngôn ngữ khác, từ người dùng và nhận âm thanh nói văn bản của người dùng theo ngôn ngữ chỉ định. Mô đun xử lý 30 được cấu hình để xử lý văn bản để thu được các đặc tính kỳ vọng (ví dụ: sự lưu loát) giọng nói văn bản đã nhận từ mô đun nhập 28, sử dụng, ví dụ, các thuật toán nhận dạng giọng nói và xử lý âm thanh đã nhận từ mô đun nhập 28 để thu được các đặc tính người dùng nói văn bản. Mô đun xử lý 30 cũng được sử dụng để phân tích cú pháp văn

bản nhờ một hoặc nhiều câu. Đó là, văn bản đã nhận, bao gồm tập hợp các từ, được phân tích thành các câu bằng mô đun xử lý 30 tìm kiếm lần lượt văn bản về sự xuất hiện của các dấu chấm câu không trích dẫn, như dấu chấm câu, dấu chấm hỏi và dấu chấm than. Vì vậy, khi sử dụng, người dùng được nhắc để nói một trong số các câu đã phân tích cú pháp và âm thanh của người dùng nói câu đầu tiên trong số các câu của văn bản được nhận bởi mô đun nhập 28 v.v.

Trong một ví dụ, câu: Anh ấy đã nói, *“Tôi không làm được!”* sau đó anh ấy tiếp tục nói được phân tích cú pháp thành một câu đơn: Anh ấy đã nói, *“Tôi không làm được”* sau đó anh ấy tiếp tục nói.

Tuy nhiên, câu sau: *Anh ấy đã nói, Tôi không làm được! Sau đó anh ấy tiếp tục nói.* được phân tích cú pháp thành hai câu: 1. *Anh ấy nói, Tôi không làm được!;* 2. *Sau đó anh ấy tiếp tục nói.* để nhắc người dùng nói những câu này. Hơn nữa, một khi tập hợp các từ từ văn bản đã nhập được phân tích cú pháp thành các câu riêng biệt, mỗi câu được chuyển thành ngữ pháp nhận dạng giọng nói phù hợp với các đặc tính kỳ vọng. Câu được trình bày dưới dạng chữ viết hoặc âm thanh đến người dùng, sau đó người dùng được nhắc để nói những gì mà anh ấy hoặc cô ấy đã đọc hoặc nghe thấy, và phản hồi bằng lời nói của người dùng được chuyển từ mô đun nhập 28 đến thiết bị nhận dạng lời nói thực hiện bởi mô đun xử lý 30, với ngữ pháp câu, để phân tích.

Theo đó, mô đun xử lý 30 so sánh các đặc tính nói văn bản của người dùng các đặc tính kỳ vọng của giọng nói văn bản, ví dụ được áp đặt bởi nhà cung cấp nhận dạng giọng nói, để xác định độ chính xác về việc nói văn bản của người dùng liên quan đến các đặc tính. Mô đun xuất 32 nhận sự xác định độ chính xác từ mô đun xử lý 30 và đưa ra phản hồi đến người dùng biểu thị độ chính xác về việc nói văn bản của người dùng.

Fig.3 thể hiện một hệ thống 36 để giúp cải thiện giọng nói của người dùng theo ngôn ngữ chỉ định bao gồm một thiết bị của người dùng 38 để nhập và xuất thông tin đến bộ xử lý 26 qua một mạng 40, như mạng Internet. Như đã mô tả trên đây, người ta dự tính rằng, trong phương án thực hiện khác không được thể hiện trên các hình vẽ, hệ thống 24 tự chứa và gồm có các bộ phận phù hợp để nhập và xuất thông tin đến bộ xử lý 26 để thực hiện phương pháp 10. Tuy nhiên, bộ xử lý 26 được thể hiện như được thực hiện bằng một máy chủ 54 có thể truy cập trên mạng Internet 40. Cũng có thể dự tính rằng bộ xử lý 26 và bộ nhớ 52 có thể được thực hiện như, có thể là, một dịch vụ

đám mây với các máy chủ ảo hoặc qua nhiều máy chủ hữu hình. Trong một ví dụ, máy chủ 54 là có thể truy cập qua Địa chỉ nguồn tài nguyên (URL) có thể được nhúng vào các trang web cho người dùng truy cập qua các thiết bị sử dụng 38 của họ. Trong ví dụ khác, mạng 40 là mạng cục bộ (LAN) và thiết bị của người dùng 38 truyền thông với máy chủ qua, có thể là, WiFi.

Thiết bị của người dùng 38 thể hiện trên Fig.3 bao gồm màn hình 44 được cấu hình để hiển thị văn bản theo ngôn ngữ chỉ định, hoặc theo ngôn ngữ khác đến người dùng. Có thể thấy trong ví dụ được minh họa trên Fig.3, văn bản hiển thị là “Này Joe, bạn đang làm gì” và thiết bị của người dùng là một máy tính bảng có màn hình cảm ứng hoặc có các khả năng đọc động tác. Như đã mô tả trên đây, các thiết bị của người dùng khác được dự tính để sử dụng với hệ thống 36, như điện thoại thông minh và các máy tính cá nhân.

Với việc tham chiếu đến một ví dụ sử dụng phương án thực hiện này, người dùng đã nhập văn bản “Này Joe, bạn đang làm gì” vào sử dụng một phương tiện nhập 46 ở dạng bàn phím của màn hình cảm ứng làm một phần văn bản được nói bởi người dùng để giúp người dùng học tiếng Anh. Một khi người dùng kết thúc đánh máy văn bản, hoặc dán văn bản từ tài liệu khác, sau đó người dùng có thể nhấn vào nút “trình” (không được thể hiện) và chuyển văn bản qua mạng Internet 40 đến máy chủ 54. Văn bản nhập được chuyển qua các kênh truyền thông 42 và 34, qua mạng Internet 40, đến mô đun nhập 28 sao cho mô đun xử lý 30 có thể xử lý văn bản để thu được các đặc tính kỳ vọng giọng nói văn bản. Như đã mô tả trên đây, mô đun xử lý 30 phân tích cú pháp văn bản thành các câu được hiển thị đến người dùng được nói một câu tại một thời điểm để nhận phản hồi biểu thị độ chính xác về việc để văn bản được nói như thế nào vào lúc kết thúc mỗi câu. Ngoài ra, trong một phương án thực hiện, mô đun xử lý 30 chuyển từng câu thành ngữ pháp phi ngữ cảnh, đây là cú pháp yêu cầu bởi thiết bị nhận dạng giọng nói mục tiêu sử dụng bởi mô đun xử lý 30 để phân tích giọng nói của người dùng nói các câu. Đó là, trong phương án thực hiện này, ngữ pháp đại diện cho cấu trúc của văn bản được đóng gói chuyển đến thiết bị nhận dạng giọng nói.

Tham chiếu đến cùng ví dụ, văn bản “Này Joe, bạn đang làm gì” được xuất trở lại thiết bị người dùng 38, qua mô đun xuất 32, và hiển thị trên màn hình 44 như là lời nhắc cho câu đầu tiên được nói. Đó là, lời nhắc bao gồm văn bản của câu đầu tiên được nói. Micro 48 được sử dụng để ghi người dùng nói câu này và âm thanh của

người dùng nói câu được chuyển qua mạng Internet 40, đến mô đun nhập 28 sao cho mô đun xử lý 30 có thể xử lý âm thanh để thu được các đặc tính nói văn bản của người dùng. Sau đó, mô đun xử lý 30 so sánh các đặc tính âm thanh đã thu được với các đặc tính kỳ vọng để xác định độ chính xác về việc nói văn bản của người dùng sử dụng các thuật toán nhận dạng giọng nói mục tiêu.

Sau đó mô đun xuất 32 đưa ra phản hồi đến người dùng biểu thị độ chính xác đã xác định bởi mô đun xử lý 30, qua mạng Internet 40, như các màu sắc được hiển thị trên màn hình 44 biểu thị các mức độ khác nhau của các giá trị tin cậy của độ chính xác đã xác định. Như đã mô tả trên đây, màu xanh lá cây được sử dụng để biểu thị giá trị lòng tin liên quan đến sự chính xác là “CHÍNH XÁC CAO” (ví dụ, phù hợp 80%+). Trong ví dụ khác, mô đun xuất 32 đưa ra phản hồi đến người dùng biểu thị độ chính xác được xác định bởi mô đun xử lý 30 như âm thanh đưa ra từ loa 50 của thiết bị người dùng 38. Ví dụ, phản hồi âm thanh khẳng định “Độ chính xác cao” qua loa 50 khi mô đun xử lý 30 xác định giá trị lòng tin lớn hơn 80%. Như mô tả trên đây màn hình 44 cũng có thể được cấu hình để hiển thị các giao thức biểu thị các mức độ khác nhau của giá trị lòng tin về độ chính xác đã xác định, như các loại số và chữ. Hơn nữa, trong một phương án thực hiện, mô đun xử lý 30 bao gồm các thuật toán tổng hợp giọng nói, như đã mô tả trên đây, và ở đây mô đun xuất 32 có thể được cấu hình để đưa ra giọng nói kỳ vọng đã tổng hợp của văn bản được đưa ra từ loa 50 đến người dùng để giúp người dùng cải thiện giọng nói của họ.

Sau khi đưa ra phản hồi, mô đun xuất 32 được cấu hình thêm để đưa ra câu tiếp theo (không thể hiện) đã phân tích cú pháp bởi mô đun xử lý 30 từ văn bản đã nhập bởi người dùng như là lời nhắc thêm cho người dùng nói câu thứ hai. Lời nhắc gồm văn bản của câu thứ hai cũng được hiển thị trên màn hình 44 và micrô 48 được sử dụng để ghi người dùng nói câu này. Như với câu thứ nhất, âm thanh của người dùng nói câu này được chuyển vào mô đun nhập 28 sao cho mô đun xử lý 30 có thể xử lý âm thanh để thu được các đặc tính nói văn bản của người dùng. Sau đó mô đun xử lý 30 so sánh các đặc tính âm thanh đã thu được với các đặc tính kỳ vọng để xác định độ chính xác về việc nói văn bản của người dùng và mô đun xuất 32 đưa ra phản hồi đến người dùng biểu thị độ chính xác được xác định bởi mô đun xử lý 30 như các màu sắc hiển thị trên màn hình 44. Sau đó mô đun xuất đưa ra câu tiếp theo và lặp lại quy trình

cho đến khi tất cả các câu đã phân tích cú pháp từ văn bản đã nhập bởi người dùng được nói hoặc người dùng chấm dứt quy trình sớm.

Như đã mô tả trên đây, bộ nhớ 52 có thể được dùng để ghi âm thanh nói văn bản của người dùng theo ngôn ngữ chỉ định. Theo đó, mô đun xử lý 30 có thể xử lý thêm các mẫu âm thanh lưu trữ trong bộ nhớ 52 để thu và lọc một hoặc nhiều đặc tính kỳ vọng giọng nói văn bản như việc phát âm văn bản kỳ vọng có thể được xác định từ sự phân tích các bản ghi âm người dùng nói văn bản đó. Ngoài ra, giọng nói đã ghi được lưu trong bộ nhớ 52 kết hợp với dữ liệu biểu thị người dùng cụ thể sao cho các bản ghi có thể được truy tìm cho phân tích sau, như để biểu thị sự tiến bộ của một người dùng cụ thể trong giọng nói của họ, có thể là, tiếng Anh. Những người có hiểu biết trung bình về lĩnh vực kỹ thuật tương ứng cũng sẽ đánh giá rằng phương pháp và hệ thống có thể được sử dụng cho người dùng mong muốn cải thiện giọng nói của họ theo các ngôn ngữ khác, như tiếng Pháp, tiếng Trung, tiếng Nhật v.v.

Khía cạnh khác của phương pháp sẽ là rõ ràng từ phần mô tả trên đây của hệ thống. Người có hiểu biết trung bình về lĩnh vực kỹ thuật tương ứng cũng sẽ đánh giá cao rằng phương pháp có thể được bao gồm trong một mã chương trình. Mã chương trình có thể được cung cấp theo nhiều cách, ví dụ trên một phương tiện đọc được bằng máy tính hữu hình, như ổ đĩa hoặc bộ nhớ hoặc tín hiệu dữ liệu hoặc tập tin dữ liệu.

Những người có hiểu biết trung bình về lĩnh vực kỹ thuật tương ứng có thể hiểu được rằng nhiều thay đổi, bổ sung và/hoặc sửa đổi có thể được thực hiện đối với sáng chế mà không tách rời khỏi tinh thần và phạm vi của sáng chế.

Phải hiểu rằng, nếu bất kỳ lĩnh vực kỹ thuật nào đã biết được đề cập trong tài liệu này, thì sự tham khảo đó không tạo ra sự thừa nhận rằng lĩnh vực kỹ thuật đã biết tạo thành một phần của kiến thức chung phổ biến trong lĩnh vực kỹ thuật tương ứng ở bất kỳ quốc gia nào.

Sáng chế có thể được sử dụng làm cơ sở ưu tiên đối với một hoặc nhiều đơn trong tương lai, và các yêu cầu bảo hộ của các đơn trong tương lai có thể được hướng đến bất kỳ một đặc trưng hoặc sự kết hợp các đặc trưng nào được mô tả trong sáng chế này. Như vậy, các đơn trong tương lai bao gồm một hoặc nhiều các yêu cầu bảo hộ dưới đây, được đưa ra làm ví dụ và không giới hạn các đối tượng có thể được bảo hộ trong bất kỳ đơn nào trong tương lai.

Yêu cầu bảo hộ

1. Phương pháp giúp cải thiện giọng nói của người dùng theo ngôn ngữ chỉ định, phương pháp bao gồm:

nhận văn bản theo ngôn ngữ chỉ định, hoặc ngôn ngữ khác từ người dùng được người dùng nói theo ngôn ngữ chỉ định;

xử lý văn bản để thu được một hoặc nhiều đặc tính kỳ vọng giọng nói văn bản theo ngôn ngữ chỉ định.

nhận âm thanh nói văn bản của người dùng theo ngôn ngữ chỉ định;

xử lý âm thanh để thu được một hoặc nhiều đặc tính nói văn bản của người dùng;

so sánh một hoặc nhiều đặc tính nói văn bản của người dùng đã nói với một trong số một hoặc nhiều các đặc tính kỳ vọng giọng nói tương ứng của văn bản đã nói để xác định độ chính xác về việc nói văn bản của người dùng; và

đưa ra phản hồi đến người dùng biểu thị độ chính xác về việc nói văn bản của người dùng, trong đó phương pháp bao gồm thêm:

phân tích cú pháp văn bản thành một hoặc nhiều câu hoặc đoạn các từ được nói bởi người dùng;

nhận âm thanh của người dùng nói các câu hoặc đoạn văn bản; và

đưa ra phản hồi đã nói đến người dùng biểu thị độ chính xác của người dùng nói các câu hoặc đoạn văn bản.

2. Phương pháp theo điểm 1, phương pháp bao gồm thêm việc đưa ra một lời nhắc cho người dùng để nói câu đầu tiên trong số các câu hoặc đoạn và nhận âm thanh của người dùng nói câu đầu tiên trong số các câu hoặc đoạn văn bản.

3. Phương pháp theo điểm 2, phương pháp bao gồm thêm việc đưa các phản hồi đã nói đến người dùng biểu thị độ chính xác của người dùng nói câu đầu tiên trong số các câu hoặc đoạn văn bản.

4. Phương pháp theo điểm 3, phương pháp bao gồm thêm việc đưa ra thêm lời nhắc cho người dùng để nói câu thứ hai trong số các câu hoặc đoạn thứ hai sau khi đưa ra

phản hồi đến người dùng biểu thị độ chính xác của người dùng nói câu đầu tiên trong số các câu hoặc đoạn văn bản.

5. Phương pháp theo điểm 2, trong đó lời nhắc bao gồm văn bản của câu đầu tiên trong số các câu hoặc đoạn được hiển thị trên một màn hình của thiết bị người dùng đến người dùng.

6. Phương pháp theo điểm 1, phương pháp bao gồm thêm việc tạo ra ngữ pháp nhận dạng giọng nói văn bản.

7. Phương pháp theo bất kỳ điểm nào từ 1 đến 6, phương pháp bao gồm thêm việc xử lý văn bản để thu được giọng nói văn bản kỳ vọng đã phân tích theo ngôn ngữ chỉ định.

8. Phương pháp theo điểm 7, phương pháp bao gồm thêm việc đưa ra giọng nói văn bản kỳ vọng đã phân tích đến người dùng qua loa của thiết bị người dùng.

9. Phương pháp theo bất kỳ điểm nào từ 1 đến 8, trong đó độ chính xác bao gồm các giá trị lòng tin liên quan đến độ chính xác của một hoặc nhiều đặc tính nói văn bản của người dùng tương ứng với một hoặc nhiều các đặc tính kỳ vọng giọng nói văn bản.

10. Phương pháp theo điểm 9, trong đó sự phản hồi bao gồm các màu sắc biểu thị các giá trị lòng tin khác nhau về độ chính xác.

11. Phương pháp theo bất kỳ điểm nào từ 1 đến 10, phương pháp bao gồm thêm việc xử lý văn bản để thu được các từ cấm xác định trước trong văn bản sao cho các từ cấm có thể được kiểm duyệt.

12. Phương pháp theo bất kỳ điểm nào từ điểm 1 đến 11, trong đó một hoặc nhiều đặc tính kỳ vọng giọng nói trong số các đặc trưng về văn bản và một hoặc nhiều đặc tính nói văn bản của người dùng bao gồm ít nhất một sự phát âm, sự lưu loát và ngôn điệu.

13. Phương pháp theo bất kỳ điểm nào từ 1 đến 12, phương pháp bao gồm thêm việc xử lý âm thanh để thu được một hoặc nhiều đặc tính nói văn bản của người dùng đã nói sử dụng thuật toán nhận dạng giọng nói.

14. Phương pháp theo bất kỳ điểm nào từ 1 đến 13, phương pháp bao gồm thêm việc xử lý văn bản để thu được một hoặc nhiều các đặc tính kỳ vọng đã nói của văn bản sử dụng thuật toán tổng hợp giọng nói.

15. Hệ thống giúp cải thiện giọng nói của người dùng theo ngôn ngữ chỉ định, hệ thống bao gồm bộ xử lý có:

mô đun nhập được cấu hình để:

nhận văn bản theo ngôn ngữ chỉ định, hoặc theo ngôn ngữ khác, từ người dùng được người dùng nói theo ngôn ngữ chỉ định; và

nhận âm thanh nói văn bản của người dùng theo ngôn ngữ chỉ định;

mô đun xử lý được cấu hình để:

xử lý văn bản để thu được một hoặc nhiều đặc tính kỳ vọng giọng nói văn bản theo ngôn ngữ chỉ định;

xử lý âm thanh để thu được một hoặc nhiều đặc tính nói văn bản của người dùng; và

so sánh một hoặc nhiều đặc tính nói văn bản của người dùng đã nói với một trong số một hoặc nhiều đặc tính kỳ vọng đã nói tương ứng của giọng nói văn bản đã nói để xác định độ chính xác về việc nói văn bản của người dùng; và mô đun xuất được cấu hình để đưa ra phản hồi đến người dùng biểu thị độ chính xác về việc nói văn bản của người dùng, trong đó:

mô đun xử lý được cấu hình thêm để phân tích cú pháp văn bản thành một hoặc nhiều câu hoặc đoạn từ;

mô đun nhập được cấu hình thêm để nhận âm thanh của người dùng nói một hoặc nhiều câu hoặc đoạn văn bản; và

mô đun xuất được cấu hình thêm để đưa ra phản hồi đã nói đến người dùng biểu thị chính xác của người dùng nói các câu hoặc đoạn văn bản.

16. Hệ thống theo điểm 15, trong đó mô đun xuất được cấu hình thêm để đưa ra lời nhắc cho người dùng để nói một câu đầu tiên trong số các câu hoặc đoạn.

17. Hệ thống theo điểm 16, trong đó mô đun xuất được cấu hình thêm để nhận âm thanh của người dùng nói câu đầu tiên trong số các câu hoặc đoạn văn bản.

18. Hệ thống theo điểm 17, trong đó mô đun xuất được cấu hình thêm để đưa ra phản hồi đã nói đến người dùng biểu thị độ chính xác của người dùng nói câu đầu tiên trong số các câu hoặc đoạn văn bản.

19. Hệ thống theo điểm 18, trong đó mô đun xuất được cấu hình thêm để đưa ra thêm một lời nhắc cho người dùng để nói câu thứ hai trong số các câu hoặc đoạn sau khi đưa ra phản hồi đến người dùng biểu thị độ chính xác của người dùng nói câu đầu tiên trong số các câu hoặc đoạn văn bản.
20. Hệ thống theo điểm 16, trong đó lời nhắc bao gồm văn bản của câu đầu tiên trong số các câu hoặc đoạn được hiển thị trên màn hình của thiết bị người dùng đến người dùng.
21. Hệ thống theo điểm 15, trong đó mô đun xử lý được cấu hình thêm để tạo ra một ngữ pháp nhận dạng giọng nói văn bản.
22. Hệ thống theo bất kỳ điểm nào từ 15 đến 21, trong đó mô đun xử lý được cấu tạo thêm để xử lý văn bản để thu được giọng nói kỳ vọng đã tổng hợp dưới dạng văn bản theo ngôn ngữ chỉ định.
23. Hệ thống theo điểm 22, trong đó mô đun xuất được cấu hình thêm để đưa ra giọng nói văn bản kỳ vọng đã tổng hợp đến người dùng qua loa của một thiết bị người dùng.
24. Hệ thống theo bất kỳ điểm nào từ 15 đến 23, trong đó độ chính xác bao gồm các giá trị lòng tin kết hợp với độ chính xác của một hoặc nhiều đặc tính nói văn bản của người dùng tương ứng với một hoặc nhiều đặc tính kỳ vọng của giọng nói văn bản.
25. Hệ thống theo điểm 24, trong đó việc phản hồi bao gồm các màu sắc biểu thị các giá trị lòng tin khác nhau về độ chính xác.
26. Hệ thống theo bất kỳ điểm nào từ 15 đến 25, trong đó mô đun xử lý được cấu hình thêm để xử lý văn bản để thu được các từ cấm xác định trước trong văn bản sao cho các từ cấm có thể được kiểm duyệt.
27. Hệ thống theo bất kỳ điểm nào từ 15 đến 26, trong đó hệ thống bao gồm thêm bộ nhớ để ghi âm thanh nói văn bản của người dùng.
28. Hệ thống theo điểm 27, trong đó mô đun xử lý được cấu hình thêm để xử lý các mẫu âm thanh ghi trong bộ nhớ để thu được một hoặc nhiều các đặc tính kỳ vọng của giọng nói văn bản.
29. Hệ thống theo bất kỳ điểm nào từ 15 đến 28, trong đó một hoặc nhiều đặc tính kỳ vọng nói văn bản của người dùng bao gồm ít nhất một trong các đặc trưng về phát âm, sự lưu loát và ngôn điệu.

30. Hệ thống giúp cải thiện giọng nói của người dùng theo ngôn ngữ chỉ định, hệ thống bao gồm:

một màn hình được cấu hình để hiển thị văn bản theo ngôn ngữ chỉ định, hoặc theo ngôn ngữ khác

phương tiện nhập văn bản được cấu hình để nhập văn bản đã nói theo ngôn ngữ chỉ định, hoặc theo ngôn ngữ khác đã nói bởi người dùng được người dùng nói theo ngôn ngữ chỉ định;

micro được cấu hình để nhập âm thanh của người nói văn bản vào theo ngôn ngữ chỉ định và;

bộ xử lý có:

mô đun nhập được cấu hình để:

nhận văn bản đã nhập theo ngôn ngữ chỉ định đã nói, hoặc theo ngôn ngữ khác đã nói; và

nhận âm thanh nói văn bản của người dùng theo ngôn ngữ chỉ định;

mô đun xử lý được cấu hình để:

xử lý văn bản để thu được một hoặc nhiều các đặc tính kỳ vọng của giọng nói văn bản theo ngôn ngữ chỉ định;

xử lý âm thanh để thu được một hoặc nhiều đặc tính nói văn bản của người dùng; và

so sánh một hoặc nhiều đặc tính nói văn bản của người dùng đã nói với một trong số một hoặc nhiều các đặc tính kỳ vọng giọng nói văn bản đã nói để xác định độ chính xác về việc nói văn bản của người dùng; và

mô đun xuất được cấu hình để đưa ra phản hồi đến người dùng biểu thị độ chính xác về việc văn bản của người dùng, trong đó:

mô đun xử lý được cấu hình thêm để phân tích cú pháp văn bản thành một hoặc nhiều câu hoặc đoạn từ;

mô đun nhập được cấu hình thêm để nhận âm thanh nói một hoặc nhiều câu hoặc đoạn văn bản của người dùng; và

mô đun xuất được cấu hình thêm để đưa ra phản hồi đã nói đến người dùng biểu thị độ chính xác về việc nói các câu hoặc đoạn văn bản của người dùng.

31. Hệ thống theo điểm 30, trong đó màn hình được cấu hình thêm để hiển thị phản hồi đến người dùng.

32. Hệ thống theo điểm 31, hệ thống bao gồm thêm một máy chủ có bộ xử lý liên tục dữ liệu qua một mạng với thiết bị người dùng bao gồm màn hình, phương tiện nhập liệu và micrô.

33. Vật ghi đọc được bằng máy tính trong đó ghi mã chương trình máy tính mà khi chạy sẽ thực hiện các phương pháp theo bất kỳ điểm nào từ 1 đến 14.

34. Vật ghi đọc được bằng máy tính trong đó ghi tập tin dữ liệu của mã chương trình máy tính được lưu trong vật ghi theo điểm 33.

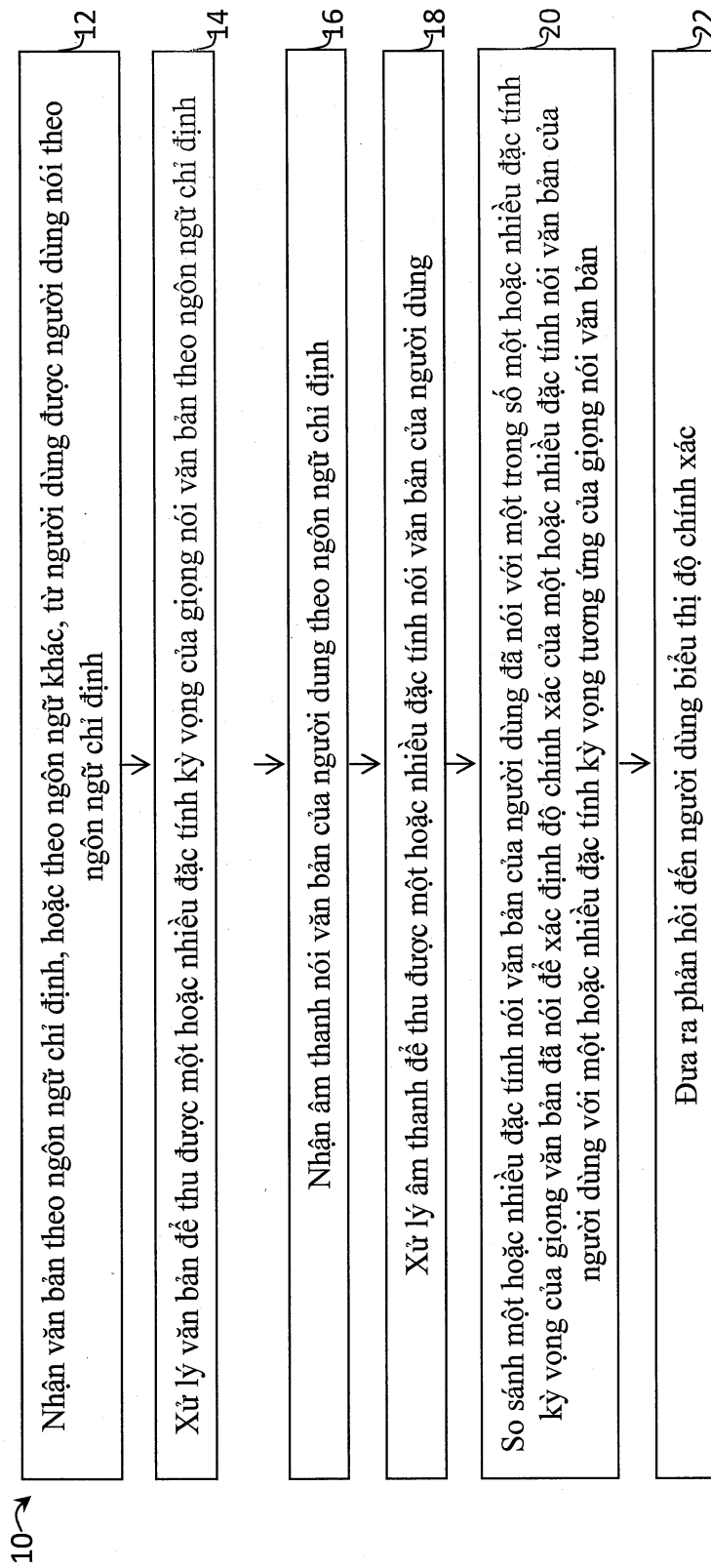


Fig.1

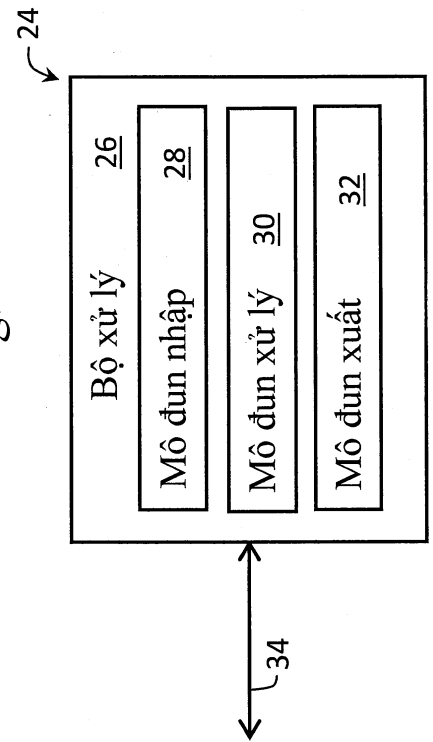


Fig.2

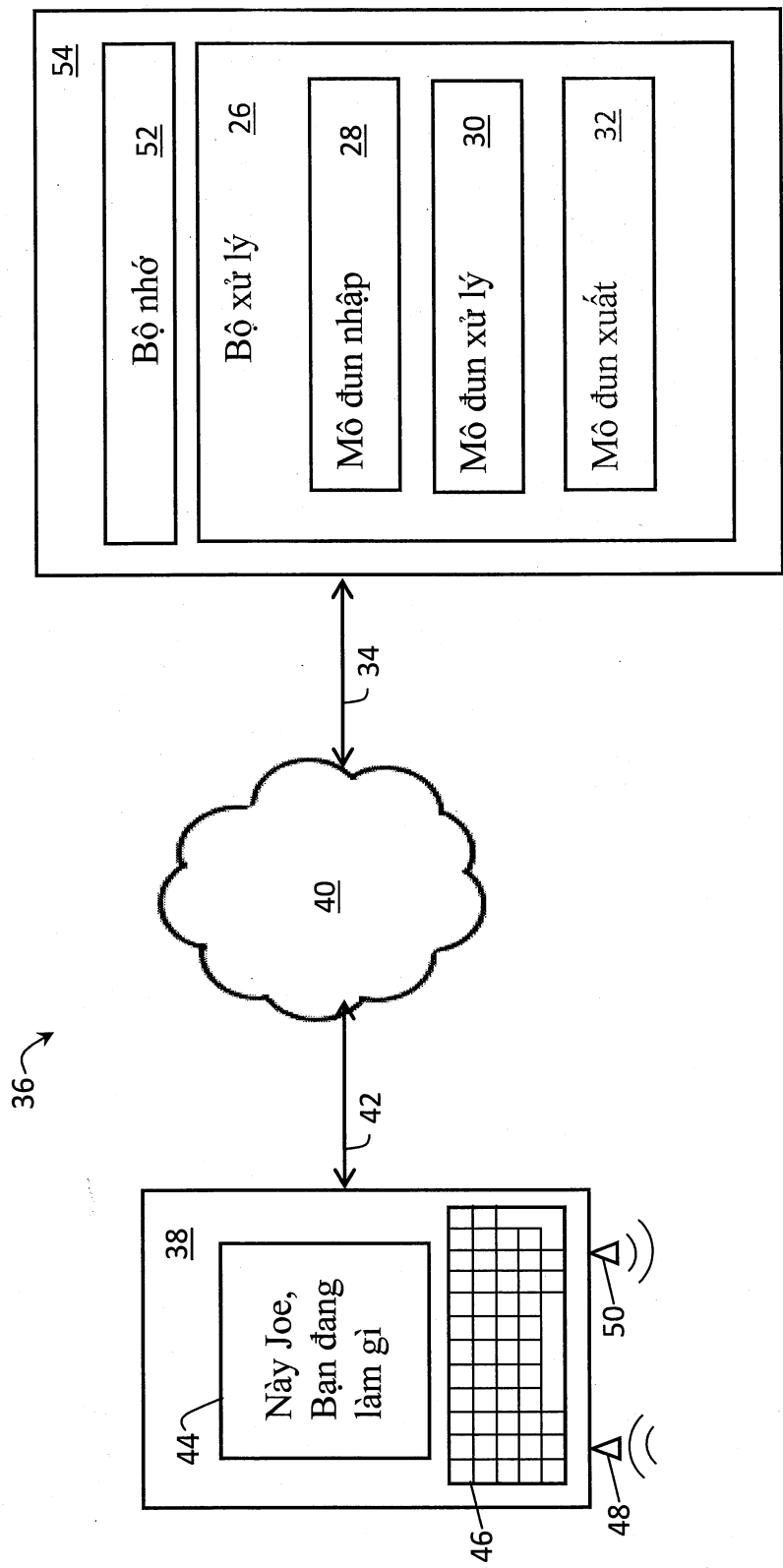


Fig.3