



(12) BẢN MÔ TẢ SÁNG CHẾ THUỘC BẰNG ĐỘC QUYỀN SÁNG CHẾ
(19) Cộng hòa xã hội chủ nghĩa Việt Nam (VN) (11)
CỤC SỞ HỮU TRÍ TUỆ
(51)⁷ **G10L 15/04; G10L 25/87; G10L 15/26; (13) B**
G10L 21/04; G06K 9/00; G10L 15/08



1-0043663

(21) 1-2019-04432 (22) 23/11/2017
(86) PCT/KR2017/013397 23/11/2017 (87) WO 2018/135743 26/07/2018
(30) 10-2017-0007951 17/01/2017 KR
(45) 25/02/2025 443 (43) 25/05/2020 386A
(73) SAMSUNG ELECTRONICS CO., LTD. (KR)
129, Samsung-ro, Yeongtong-gu, Suwon-si, Gyeonggi-do, 16677, Republic of Korea
(72) KIM, Yong Ho (KR); PATERIYA, Sourabh (IN); KIM, Sunah (KR); JOO, Gahyun
(KR); HWANG, Sang-Woong (KR); JANG, Say (KR).
(74) Công ty TNHH Tầm nhìn và Liên danh (VISION & ASSOCIATES CO.LTD.)

(54) THIẾT BỊ ĐIỆN TỬ VÀ PHƯƠNG PHÁP VẬN HÀNH THIẾT BỊ ĐIỆN TỬ NÀY

(21) 1-2019-04432

(57) Sáng chế đề cập đến thiết bị và phương pháp vận hành thiết bị điện tử này, các phương án khác nhau của thiết bị này bao gồm micrô, bộ nhớ, và bộ xử lý được kết nối chức năng với micrô hoặc bộ nhớ, trong đó bộ xử lý được tạo cấu hình để: tính thời gian phát hiện điểm cuối (End Point Detection, EPD) dựa trên đầu vào thoại; khi hết thời gian EPD, thì xác định xem từ cuối cùng của đầu vào thoại có tương ứng với từ đã được tạo cấu hình trước đó được lưu trữ trong bộ nhớ hay không; và nếu từ cuối cùng này tương ứng với từ đã được tạo cấu hình trước đó, thì kéo dài thời gian EPD và chờ nhận đầu vào thoại. Ngoài ra, các phương án khác cũng là có thể.

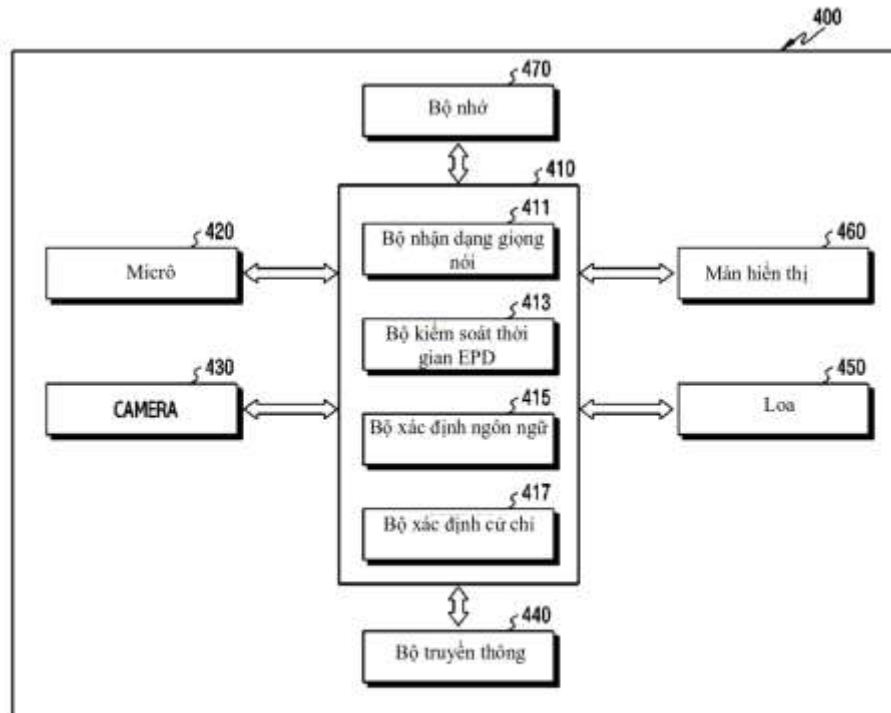


FIG.4

Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến phương pháp và thiết bị dùng để cảm biến sự kết thúc giọng nói.

Tình trạng kỹ thuật của sáng chế

Với sự cải thiện gần đây của công nghệ kỹ thuật số, nhiều loại thiết bị điện tử như thiết bị đầu cuối truyền thông di động, thiết bị hỗ trợ cá nhân kỹ thuật số (Personal Digital Assistant, PDA), thiết bị lập lịch biểu điện tử, điện thoại thông minh, máy tính cá nhân (personal computer, PC) dạng bảng, thiết bị đeo được, hoặc thiết bị tương tự đang được sử dụng rộng rãi. Để hỗ trợ và tăng cường chức năng của các thiết bị điện tử này, các bộ phận phần cứng và/hoặc các bộ phận phần mềm của các thiết bị điện tử này đang được phát triển liên tục.

Các thiết bị điện tử hiện đang cung cấp nhiều dịch vụ (hoặc chức năng) khác nhau bằng cách sử dụng công nghệ nhận dạng giọng nói. Công nghệ nhận dạng giọng nói bao gồm, ví dụ, công nghệ chuyển đổi tín hiệu âm thanh giọng nói thu được bởi thiết bị điện tử thông qua cảm biến âm thanh như micrô thành từ hoặc câu. Công nghệ nhận dạng giọng nói thường thực hiện bước loại bỏ nhiễu sau khi trích xuất tín hiệu âm thanh giọng nói, và sau đó thực hiện nhận dạng giọng nói bằng cách trích xuất các đặc điểm của tín hiệu âm thanh giọng nói và so sánh các đặc điểm này với cơ sở dữ liệu (Database, DB) về mẫu giọng nói. Các thiết bị điện tử hiện đang cung cấp nhiều loại dịch vụ thoại (hoặc dịch vụ nhận dạng giọng nói) khác nhau dựa trên công nghệ nhận dạng giọng nói. Theo một phương án, các thiết bị điện tử cung cấp nhiều loại dịch vụ tiện lợi cho cuộc sống như tìm kiếm trên thiết bị di động, quản lý lịch trình, gọi điện, ghi nhớ hoặc phát nhạc, dựa trên lệnh thoại của người sử dụng.

Bản chất kỹ thuật của sáng chế

Vấn đề kỹ thuật

Thiết bị điện tử có thể kích hoạt (vận hành) dịch vụ thoại theo nhiều phương pháp tương tác khác nhau. Ví dụ, phương pháp tương tác để kích hoạt dịch vụ thoại có

thể là ít nhất một trong số các phương pháp sau: phương pháp đánh thức bằng cảm ứng chạm, phương pháp đánh thức bằng giọng nói hoặc ấn nút để nói (push to talk), và phương pháp đánh thức bằng cách theo dõi cử chỉ hoặc khuôn mặt/mắt. Hầu hết các dịch vụ thoại mà không sử dụng phương pháp ấn nút để nói đều sử dụng phương pháp phát hiện điểm cuối (End Point Detection, EPD). Phương pháp EPD có thể chỉ việc phát hiện sự kết thúc đầu vào thoại của người sử dụng. Ví dụ, nếu xác định được rằng cường độ giọng nói của người sử dụng nhỏ hơn hoặc bằng giá trị ngưỡng trong thời gian lâu hơn thời gian định sẵn, thì thiết bị điện tử có thể nhận dạng rằng giọng nói của người sử dụng đã kết thúc và có thể truyền thông tin đã nhận dạng được đến máy chủ (ví dụ, máy chủ nhận dạng giọng nói) để xử lý thông tin. Theo cách khác, nếu không nhận được đầu vào người sử dụng trong khoảng thời gian định sẵn (ví dụ, 3 giây, 5 giây) sau khi bật micrô, thì thiết bị điện tử có thể xác định rằng người sử dụng có ý định kết thúc đầu vào thoại, và có thể không nhận đầu vào thoại. Tuy nhiên, khi EPD chỉ được sử dụng với thông tin về thời gian hoặc âm thanh giọng nói, thì thiết bị điện tử có thể không nhận đầu vào thoại hoặc có thể kết thúc dịch vụ thoại ngay cả khi người sử dụng không muốn kết thúc đầu vào thoại này.

Các phương án khác nhau đề xuất phương pháp và thiết bị phân tích ý định kết thúc (hoặc hoàn thành) giọng nói của người sử dụng theo các phương pháp khác nhau, như bằng cách sử dụng ngữ cảnh (hoặc trạng thái) của thiết bị điện tử, đặc trưng của người sử dụng, từ định sẵn (ví dụ, từ rỗng, liên từ, lệnh chờ, v.v.) hoặc cử chỉ định sẵn, và xác định xem đầu vào thoại đã kết thúc hay chưa.

Giải pháp để giải quyết vấn đề

Theo các phương án khác nhau, thiết bị điện tử có thể bao gồm: micrô; bộ nhớ; và bộ xử lý được kết nối chức năng với micrô hoặc bộ nhớ này, và bộ xử lý này có thể được tạo cấu hình để tính thời gian EPD dựa trên đầu vào thoại, để xác định xem từ cuối cùng của đầu vào thoại có tương ứng với từ định sẵn được lưu trữ trong bộ nhớ khi hết thời gian EPD hay không, và để kéo dài thời gian EPD khi từ cuối cùng tương ứng với từ định sẵn, và để chờ nhận đầu vào thoại.

Theo các phương án khác nhau, phương pháp vận hành thiết bị điện tử có thể bao gồm các bước: tính thời gian EPD dựa trên đầu vào thoại; khi hết thời gian EPD, xác định xem từ cuối cùng của đầu vào thoại có tương ứng với từ định sẵn được lưu

trữ trong bộ nhớ hay không; và khi từ cuối cùng tương ứng với từ định sẵn, thì kéo dài thời gian EPD và chờ nhận đầu vào thoại.

Hiệu quả đạt được bởi sáng chế

Theo các phương án khác nhau, ý định kết thúc (hoặc hoàn thành) giọng nói của người sử dụng có thể được phân tích theo nhiều phương pháp xác định khác nhau, như bằng cách sử dụng ngữ cảnh (hoặc trạng thái) của thiết bị điện tử, đặc trưng của người sử dụng, từ định sẵn (ví dụ, từ rỗng, liên từ, lệnh chờ, v.v.) hoặc cử chỉ định sẵn, và có thể xác định được việc liệu đầu vào thoại đã kết thúc hay chưa.

Theo các phương án khác nhau, khi người sử dụng do dự vì người sử dụng không nghĩ ra từ chính xác, ý định của người sử dụng không rõ ràng, hoặc người sử dụng không biết chức năng nào có thể thực hiện được bằng dịch vụ thoại, thì có thể tăng cường sự tiện lợi cho người sử dụng bằng cách nắm bắt ý định kết thúc giọng nói của người sử dụng và kéo dài thời gian EPD cho đến khi đầu vào thoại của người sử dụng hoàn thành.

Theo các phương án khác nhau, thiết bị điện tử không kết thúc dịch vụ thoại hoặc không ngừng nhận dạng giọng nói (ví dụ, nhận đầu vào thoại) trong khi người sử dụng đang nói, và chờ đến khi người sử dụng nhập giọng nói, để có thể cung cấp dịch vụ thoại với trải nghiệm đối thoại thoải mái hơn.

Theo các phương án khác nhau, các từ (hoặc ngôn ngữ) thông dụng tương ứng với tất cả người sử dụng hoặc các từ (hoặc ngôn ngữ) cá nhân đặc trưng cho người sử dụng thông qua học tập được thu thập, để có thể xác định chính xác hơn ý định kết thúc giọng nói của người sử dụng dựa trên các từ thông dụng và các từ cá nhân này.

Theo các phương án khác nhau, có thể xác định chính xác hơn ý định kết thúc giọng nói của người sử dụng bằng cách thay đổi thời gian EPD hoặc thời gian kéo dài EPD theo đặc điểm của ứng dụng mà người sử dụng muốn thực thi để nhận dịch vụ thoại hoặc đặc điểm cá nhân của người sử dụng.

Mô tả vắn tắt các hình vẽ

Fig.1 là hình vẽ minh họa thiết bị điện tử trong môi trường mạng theo các phương án khác nhau;

Fig.2 là sơ đồ khối minh họa cấu hình của thiết bị điện tử theo các phương án khác nhau;

Fig.3 là sơ đồ khối minh họa mô đun chương trình theo các phương án khác nhau;

Fig.4 là sơ đồ minh họa cấu hình của thiết bị điện tử theo các phương án khác nhau;

Fig.5A hoặc Fig.5B là lưu đồ minh họa phương pháp vận hành thiết bị điện tử theo các phương án khác nhau;

Fig.6A đến Fig.6C là các hình vẽ minh họa các ví dụ về giao diện người sử dụng theo các phương án khác nhau;

Fig.7 là lưu đồ minh họa phương pháp thu thập từ cá nhân trong thiết bị điện tử theo các phương án khác nhau;

Fig.8A đến Fig.8C là các hình vẽ minh họa các ví dụ về việc thu thập thông tin ngôn ngữ cá nhân trong thiết bị điện tử theo các phương án khác nhau;

Fig.9 là lưu đồ minh họa phương pháp cảm biến sự kết thúc giọng nói trong thiết bị điện tử theo các phương án khác nhau;

Fig.10 là lưu đồ khác minh họa phương pháp cảm biến sự kết thúc giọng nói trong thiết bị điện tử theo các phương án khác nhau;

Fig.11 là hình vẽ minh họa ví dụ về việc xác định xem câu có được hoàn thành hay không trong thiết bị điện tử theo các phương án khác nhau;

Fig.12 là lưu đồ minh họa phương pháp cảm biến sự kết thúc giọng nói bằng cách sử dụng ý định kết thúc của người sử dụng theo các phương án khác nhau;

Fig.13 là lưu đồ minh họa phương pháp kiểm soát thời gian EPD theo thông tin ngữ cảnh theo các phương án khác nhau; và

Fig.14 là hình vẽ minh họa các ví dụ về thiết bị điện tử theo các phương án khác nhau.

Mô tả chi tiết sáng chế

Các phương án khác nhau của sáng chế được mô tả dưới đây dựa vào các hình vẽ kèm theo. Tuy nhiên, cần phải hiểu rằng sáng chế không bị giới hạn ở các dạng cụ thể được mô tả trong bản mô tả này; đúng hơn là, cần phải hiểu rằng sáng chế bao gồm các dạng cải biến, các dạng tương đương và/hoặc các dạng thay thế khác nhau của các phương án theo sáng chế. Khi mô tả các hình vẽ, các số chỉ dẫn giống nhau có thể được sử dụng để biểu thị các bộ phận cấu tạo giống nhau. Dạng thể hiện số ít cũng có thể bao gồm cả dạng thể hiện số nhiều, trừ khi ngữ cảnh có quy định khác rõ ràng. Trong bản mô tả này, cách diễn tả “A hoặc B” hoặc “ít nhất một trong số A và/hoặc B” có thể bao gồm tất cả các tổ hợp có thể có của các đối tượng được liệt kê. Cách diễn tả “thứ nhất”, “thứ hai”, “thứ nhất này” hoặc “thứ hai này” được sử dụng trong các phương án khác nhau của sáng chế có thể bỏ nghĩa cho các bộ phận khác nhau bất kể thứ tự và/hoặc mức độ quan trọng, nhưng không làm giới hạn các bộ phận tương ứng. Khi một bộ phận (ví dụ, bộ phận thứ nhất) được gọi là “được kết nối” hoặc “được ghép nối” (theo cách hoạt động được hoặc truyền thông được) với một bộ phận khác (ví dụ, bộ phận thứ hai), thì nó có thể được kết nối trực tiếp hoặc được ghép nối trực tiếp với bộ phận còn lại, hoặc có thể đặt bộ phận khác bất kỳ (ví dụ, bộ phận thứ ba) vào giữa chúng.

Cách diễn tả “được tạo cấu hình để” trong bản mô tả này có thể được thay thế bằng, ví dụ, “thích hợp để”, “có khả năng để”, “được thiết kế để”, “được làm cho phù hợp để”, “được tạo ra để” hoặc “có thể thực hiện để” tùy thuộc vào từng trường hợp. Thuật ngữ “được tạo cấu hình để” có thể không nhất thiết phải có nghĩa là “được thiết kế đặc biệt để” trong phần cứng. Theo các khác, trong một số trường hợp, cách diễn tả “thiết bị được tạo cấu hình để” có thể có nghĩa là thiết bị này, cùng với các thiết bị hoặc bộ phận khác, “có khả năng để”. Ví dụ, cách diễn tả “bộ xử lý được làm cho phù hợp để (hoặc được tạo cấu hình để) thực hiện A, B và C” có thể có nghĩa là bộ xử lý chuyên dụng (ví dụ, bộ xử lý nhúng) chỉ dùng để thực hiện các hoạt động tương ứng hoặc bộ xử lý có mục đích chung (ví dụ, bộ xử lý trung tâm (Central Processing Unit, CPU) hoặc bộ xử lý ứng dụng (Application Processor, AP)) mà có thể thực hiện các hoạt động tương ứng bằng cách thực hiện một hoặc nhiều chương trình phần mềm được lưu trữ trong thiết bị bộ nhớ.

Thiết bị điện tử theo các phương án khác nhau của sáng chế có thể bao gồm ít

nhất một trong số, ví dụ, điện thoại thông minh, máy tính cá nhân (personal computer, PC) dạng bảng, điện thoại di động, điện thoại video, thiết bị đọc sách điện tử, PC dạng để bàn, PC dạng laptop, máy tính dạng netbook, trạm làm việc, máy chủ, thiết bị hỗ trợ cá nhân kỹ thuật số (Personal Digital Assistant, PDA), thiết bị phát đa phương tiện di động (Portable Multimedia Player, PMP), thiết bị phát MPEG-1 lớp audio 3 (MPEG-1 audio layer-3, MP3), thiết bị y tế di động, camera và thiết bị đeo được. Theo các phương án khác nhau, thiết bị đeo được này có thể bao gồm ít nhất một loại phụ kiện (ví dụ, đồng hồ đeo tay, nhẫn, vòng tay, vòng chân, vòng cổ, kính, kính áp tròng hoặc thiết bị gắn được vào đầu (Head-Mounted Device, HMD)), loại tích hợp được vào vải hoặc quần áo (ví dụ, quần áo điện tử), loại gắn được vào cơ thể (ví dụ, miếng dán da hoặc hình xăm), và loại cấy sinh học được (ví dụ, mạch cấy được). Theo một số phương án, thiết bị điện tử này có thể là thiết bị gia dụng. Thiết bị gia dụng này có thể bao gồm ít nhất một trong số, ví dụ, vô tuyến, thiết bị phát đĩa video kỹ thuật số (Digital video disk, DVD), thiết bị phát âm thanh, tủ lạnh, máy điều hòa không khí, máy hút bụi, lò nướng, lò vi sóng, máy giặt, máy làm sạch không khí, hộp thu giải mã tín hiệu truyền hình, bảng điều khiển nhà tự động, bảng điều khiển bảo mật, hộp TV (ví dụ, Samsung HomeSync™, Apple TV™ hoặc Google TV™), bảng điều khiển trò chơi (ví dụ, Xbox™ và PlayStation™), từ điển điện tử, khóa điện tử, máy quay video và khung ảnh điện tử.

Theo phương án khác, thiết bị điện tử này có thể bao gồm ít nhất một trong số các loại thiết bị y tế (ví dụ, các thiết bị đo y tế di động khác nhau (thiết bị theo dõi đường máu, thiết bị theo dõi nhịp tim, thiết bị đo huyết áp, thiết bị đo thân nhiệt, v.v.), thiết bị chụp mạch cộng hưởng từ (Magnetic Resonance Angiography, MRA), thiết bị chụp ảnh cộng hưởng từ (Magnetic Resonance Imaging, MRI), thiết bị chụp cắt lớp vi tính (Computed Tomography, CT) và thiết bị siêu âm), thiết bị định vị, bộ thu hệ thống định vị toàn cầu (Global Positioning System, GPS), thiết bị ghi dữ liệu sự kiện (Event Data Recorder, EDR), thiết bị ghi dữ liệu chuyến bay (Flight Data Recorder, FDR), thiết bị thông tin giải trí trong xe, thiết bị điện tử dùng cho tàu biển (ví dụ, thiết bị định vị dùng cho tàu biển và la bàn con quay hồi chuyển), thiết bị điện tử hàng không, thiết bị bảo mật, hệ thống thiết bị đầu xe, robot gia dụng hoặc công nghiệp, máy giao dịch tự động (automatic teller's machine, ATM) trong ngân hàng, thiết bị dùng tại điểm bán lẻ (point of sales, POS) trong cửa hàng, hoặc thiết bị kết nối vạn vật (internet device of things, IoT) (ví dụ, bóng đèn, các bộ cảm biến khác nhau, đồng hồ đo điện hoặc khí

đốt, thiết bị phun nước, thiết bị báo cháy, thiết bị điều nhiệt, đèn đường, lò nướng bánh, đồ dùng thể thao, bình nước nóng, máy sưởi, nồi hơi, v.v.).

Theo một số phương án, thiết bị điện tử này có thể bao gồm ít nhất một trong số một phần của đồ nội thất hoặc tòa nhà/kiến trúc, bảng điện tử, thiết bị nhận chữ ký điện tử, máy chiếu, và các loại thiết bị đo khác (ví dụ, đồng hồ đo nước, đồng hồ đo điện, đồng hồ đo khí đốt hoặc đồng hồ đo sóng vô tuyến). Thiết bị điện tử theo các phương án khác nhau của sáng chế có thể là sự kết hợp của một hoặc nhiều thiết bị nêu trên. Thiết bị điện tử theo một số phương án của sáng chế có thể là thiết bị mềm dẻo. Ngoài ra, thiết bị điện tử theo một phương án của sáng chế không bị giới hạn ở các thiết bị nêu trên, và có thể bao gồm thiết bị điện tử mới theo sự phát triển của công nghệ. Thiết bị điện tử theo các phương án khác nhau sẽ được mô tả dưới đây dựa vào các hình vẽ kèm theo. Trong bản mô tả này, thuật ngữ “người sử dụng” có thể chỉ người sử dụng thiết bị điện tử hoặc thiết bị (ví dụ, thiết bị điện tử thông minh nhân tạo) mà sử dụng thiết bị điện tử.

Fig.1 minh họa môi trường mạng bao gồm thiết bị điện tử theo các phương án khác nhau của sáng chế.

Thiết bị điện tử 101 trong môi trường mạng 100 theo các phương án khác nhau sẽ được mô tả dựa vào Fig.1. Thiết bị điện tử 101 có thể gồm có bus 110, bộ xử lý 120, bộ nhớ 130, giao diện đầu vào/đầu ra 150, màn hiển thị 160 và giao diện truyền thông 170. Trong một số phương án, thiết bị điện tử 101 có thể lược bỏ ít nhất một trong số các bộ phận nêu trên hoặc có thể còn bao gồm các bộ phận khác.

Bus 110 có thể bao gồm, ví dụ, mạch để kết nối các bộ phận 110-170 và truyền thông tin truyền thông (ví dụ, tin nhắn và/hoặc dữ liệu điều khiển) giữa các bộ phận này.

Bộ xử lý 120 có thể bao gồm một hoặc nhiều trong số bộ xử lý trung tâm (Central Processing Unit, CPU), bộ xử lý ứng dụng (Application Processor, AP) và bộ xử lý truyền thông (Communication Processor, CP). Ví dụ, bộ xử lý 120 có thể thực hiện các hoạt động hoặc xử lý dữ liệu liên quan đến việc điều khiển và/hoặc truyền thông của ít nhất một bộ phận khác của thiết bị điện tử 101.

Bộ nhớ 130 có thể bao gồm bộ nhớ khả biến và/hoặc bộ nhớ không khả biến. Bộ

xử lý 130 có thể lưu trữ, ví dụ, các lệnh hoặc dữ liệu liên quan đến ít nhất một bộ phận khác của thiết bị điện tử 101. Theo một phương án, bộ nhớ 130 có thể lưu trữ phần mềm và/hoặc chương trình 140. Chương trình 140 có thể bao gồm, ví dụ, nhân 141, phần sụn 143, giao diện lập trình ứng dụng (Application Programming Interface, API) 145 và/hoặc chương trình ứng dụng (hoặc “ứng dụng”) 147. Ít nhất một vài trong số nhân 141, phần sụn 143 và API 145 có thể được gọi là hệ điều hành (Operating System, OS).

Nhân 141 có thể điều khiển hoặc quản lý các tài nguyên hệ thống (ví dụ, bus 110, bộ xử lý 120 hoặc bộ nhớ 130) được sử dụng để thực hiện hoạt động hoặc chức năng được thực hiện bởi các chương trình khác (ví dụ, phần sụn 143, API 145 hoặc chương trình ứng dụng 147). Ngoài ra, nhân 141 có thể tạo ra giao diện mà qua đó phần sụn 143, API 145 hoặc chương trình ứng dụng 147 có thể truy cập từng bộ phận riêng lẻ của thiết bị điện tử 101 để điều khiển hoặc quản lý tài nguyên hệ thống.

Phần sụn 143, ví dụ, có thể đóng vai trò làm trung gian để cho phép API 145 hoặc các chương trình ứng dụng 147 truyền thông với nhân 141 để trao đổi dữ liệu. Ngoài ra, phần sụn 143 có thể xử lý một hoặc nhiều yêu cầu hoạt động nhận được từ chương trình ứng dụng 147 theo mức độ ưu tiên. Ví dụ, phần sụn 143 có thể gán mức độ ưu tiên sử dụng các tài nguyên hệ thống của thiết bị điện tử 101 (ví dụ, bus 110, bộ xử lý 120, bộ nhớ 130, và bộ phận tương tự) cho ít nhất một trong số các chương trình ứng dụng 147. Ví dụ, phần sụn 143 có thể thực hiện việc lập lịch hoặc cân bằng tải cho một hoặc nhiều yêu cầu hoạt động bằng cách xử lý một hoặc nhiều yêu cầu hoạt động này theo mức độ ưu tiên được gán cho ít nhất một chương trình ứng dụng. API 145 là giao diện mà qua đó các ứng dụng 147 điều khiển các chức năng được cung cấp từ nhân 141 hoặc phần sụn 143, và có thể bao gồm, ví dụ, ít nhất một giao diện hoặc chức năng (ví dụ, lệnh) để điều khiển tệp tin, điều khiển cửa sổ, xử lý hình ảnh, hoặc điều khiển văn bản.

Ví dụ, giao diện đầu vào/đầu ra 150 có thể đóng vai trò làm giao diện có thể truyền lệnh hoặc dữ liệu đầu vào từ người sử dụng hoặc thiết bị ngoại vi khác đến (các) bộ phận khác của thiết bị điện tử 101. Ngoài ra, giao diện đầu vào/đầu ra 150 có thể kết xuất các lệnh hoặc dữ liệu nhận được từ (các) bộ phận khác của thiết bị điện tử 101 đến người sử dụng hoặc thiết bị ngoại vi khác.

Ví dụ, màn hiển thị 160 có thể bao gồm màn hình tinh thể lỏng (Liquid Crystal Display, LCD), màn hình điốt phát quang (Light Emitting Diode, LED), màn hình điốt phát quang hữu cơ (Organic Light Emitting Diode, OLED), màn hình hệ thống vi cơ điện tử (Micro Electro Mechanical System, MEMS) hoặc màn hình giấy điện tử. Ví dụ, màn hiển thị 160 có thể hiển thị nhiều loại nội dung khác nhau (ví dụ, văn bản, hình ảnh, video, biểu tượng hoặc ký hiệu) cho người sử dụng. Màn hiển thị 160 có thể bao gồm màn hình chạm, và có thể nhận, ví dụ, đầu vào chạm, cử chỉ, tiệm cận hoặc lân cận bằng cách sử dụng bút điện tử hoặc bộ phận cơ thể của người sử dụng.

Ví dụ, giao diện truyền thông 170 có thể thiết lập truyền thông giữa thiết bị điện tử 101 và thiết bị ngoại vi (ví dụ, thiết bị điện tử ngoại vi thứ nhất 102, thiết bị điện tử ngoại vi thứ hai 104 hoặc máy chủ 106). Ví dụ, giao diện truyền thông 170 có thể được kết nối với mạng 162 thông qua truyền thông không dây hoặc truyền thông có dây để truyền thông với thiết bị ngoại vi (ví dụ, thiết bị điện tử ngoại vi thứ hai 104 hoặc máy chủ 106). Ví dụ, truyền thông không dây có thể bao gồm tiến hóa dài hạn (Long Term Evolution, LTE), LTE cải tiến (LTE-Advance, LTE-A), đa truy cập phân chia mã (Code Division Multiple Access, CDMA), CDMA băng thông rộng (Wideband CDMA, WCDMA), hệ thống viễn thông di động toàn cầu (Universal Mobile Telecommunications System, UMTS), WiBro (Wireless Broadband-băng thông rộng không dây) và hệ thống truyền thông di động toàn cầu (Global System for Mobile Communications, GSM), như giao thức truyền thông mạng tế bào. Theo một phương án, ví dụ, truyền thông không dây có thể bao gồm ít nhất một trong số Wi-Fi, Bluetooth, Bluetooth năng lượng thấp (Bluetooth low energy, BLE), Zigbee, truyền thông trường gần (Near Field Communication, NFC), truyền dữ liệu an toàn qua từ tính, tần số vô tuyến hoặc mạng vùng cơ thể (body area network, BAN).

GNSS có thể bao gồm ít nhất một trong số, ví dụ, hệ thống định vị toàn cầu (Global Positioning System, GPS), hệ thống vệ tinh định vị toàn cầu (Global Navigation Satellite System, Glonass), hệ thống vệ tinh định vị Beidou (dưới đây gọi là “Beidou”) và hệ thống định vị dựa trên vệ tinh toàn cầu châu Âu (Galileo), tùy theo vùng sử dụng, băng thông hoặc tương tự. Dưới đây, trong bản mô tả này, thuật ngữ “GPS” có thể được sử dụng theo cách có thể thay thế được với thuật ngữ “GNSS”. Truyền thông có dây có thể bao gồm ít nhất một trong số, ví dụ, bus nối tiếp vạn năng (Universal Serial Bus, USB), giao diện đa phương tiện độ phân giải cao (High

Definition Multimedia Interface, HDMI), chuẩn khuyến nghị 232 (Recommended Standard 232, RS-232) và hệ thống điện thoại thông thường (Plain Old Telephone Service, POTS). Mạng 162 có thể bao gồm ít nhất một trong số mạng truyền thông như mạng máy tính (ví dụ, LAN hoặc WAN), Internet và mạng điện thoại.

Mỗi trong số thiết bị điện tử ngoại vi thứ nhất 102 và thiết bị điện tử ngoại vi thứ hai 104 có thể là loại giống hoặc khác với loại thiết bị điện tử 101. Theo một phương án, máy chủ 106 có thể bao gồm nhóm gồm một hoặc nhiều máy chủ. Theo các phương án khác nhau, tất cả hoặc một số bước trong số các bước được thực hiện trong thiết bị điện tử 101 có thể được thực hiện trong thiết bị điện tử hoặc nhiều thiết bị điện tử khác (ví dụ, các thiết bị điện tử 102 và 104 hoặc máy chủ 106). Theo một phương án, khi thiết bị điện tử 101 cần phải thực hiện một số chức năng hoặc dịch vụ một cách tự động hoặc để phản hồi lại yêu cầu, thiết bị điện tử 101 có thể gửi yêu cầu thực hiện ít nhất một số chức năng liên quan đến thiết bị khác (ví dụ, thiết bị điện tử 102 hoặc 104 hoặc máy chủ 106) thay vì tự thực hiện các chức năng hoặc dịch vụ này hoặc thực hiện bổ sung thêm. Thiết bị điện tử khác có thể thực thi các chức năng được yêu cầu hoặc các chức năng bổ sung thêm, và có thể gửi kết quả thực hiện đến thiết bị điện tử 101. Thiết bị điện tử 101 có thể xử lý kết quả nhận được ở dạng nguyên trạng hoặc bổ sung thêm để cung cấp các chức năng hoặc dịch vụ được yêu cầu. Ví dụ, để đạt được mục đích này, công nghệ điện toán đám mây, công nghệ điện toán phân tán hoặc công nghệ điện toán máy khách - máy chủ có thể được sử dụng.

Máy chủ 106, ví dụ, bao gồm ít nhất một trong số máy chủ xử lý giọng nói, máy chủ nhận dạng giọng nói, máy chủ cung cấp dịch vụ thoại, máy chủ dữ liệu, máy chủ tìm kiếm, máy chủ thanh toán, máy chủ công ty thẻ, máy chủ ngân hàng, máy chủ xác thực, máy chủ ứng dụng, máy chủ quản lý, máy chủ tích hợp, máy chủ nhà cung cấp (hoặc máy chủ nhà điều hành truyền thông), máy chủ lưu trữ nội dung, máy chủ internet hoặc máy chủ đám mây.

Fig.2 là sơ đồ khối minh họa thiết bị điện tử theo các phương án khác nhau của sáng chế.

Fig.2 là sơ đồ khối của thiết bị điện tử 201 theo các phương án khác nhau. Ví dụ, thiết bị điện tử 201 có thể bao gồm toàn bộ hoặc một phần của thiết bị điện tử 101 được minh họa trong Fig.1. Thiết bị điện tử 201 có thể bao gồm ít nhất một bộ xử lý

(ví dụ, bộ xử lý ứng dụng (Application Processor, AP)) 210, môđun truyền thông 220, môđun nhận dạng thuê bao (Subscriber Identification Module, SIM) 224, bộ nhớ 230, môđun cảm biến 240, thiết bị đầu vào 250, màn hiển thị 260, giao diện 270, môđun audio 280, môđun camera 291, môđun quản lý nguồn 295, pin 296, bộ chỉ báo 297 và động cơ 298.

Bộ xử lý 210 có thể điều khiển nhiều bộ phận phần cứng hoặc phần mềm được kết nối với bộ xử lý 210 bằng cách điều khiển hệ điều hành hoặc chương trình ứng dụng và thực hiện việc xử lý các phần dữ liệu và các phép tính. Bộ xử lý 210 có thể được thực hiện bằng, ví dụ, hệ thống trên chip (SoC). Theo một phương án, bộ xử lý 210 có thể còn bao gồm bộ xử lý đồ họa (GPU) và/hoặc bộ xử lý tín hiệu hình ảnh. Bộ xử lý 210 có thể bao gồm ít nhất một số bộ phận (ví dụ, môđun mạng tế bào 221) trong số các bộ phận được minh họa trên Fig.2. Bộ xử lý 210 có thể tải, vào bộ nhớ khả biến, lệnh hoặc dữ liệu nhận được từ ít nhất một bộ phận (ví dụ, bộ nhớ không khả biến) trong số các bộ phận còn lại và có thể xử lý lệnh hoặc dữ liệu đã tải này, và có thể lưu trữ các dữ liệu trong bộ nhớ không khả biến.

Môđun truyền thông 220 có thể bao gồm, ví dụ, môđun mạng tế bào 221, môđun Wi-Fi 223, môđun Bluetooth (BT) 225, môđun GNSS 227 (ví dụ, môđun GPS, môđun Glonass, môđun Beidou hoặc môđun Galileo), môđun NFC 228 và môđun tần số vô tuyến (RF) 229. Môđun mạng tế bào 221 có thể cung cấp cuộc gọi thoại, cuộc gọi hình ảnh, dịch vụ tin nhắn văn bản hoặc dịch vụ Internet qua, ví dụ, mạng truyền thông. Theo một phương án, môđun mạng tế bào 221 có thể phân biệt và xác thực các thiết bị điện tử 201 trong mạng truyền thông bằng cách sử dụng môđun nhận dạng thuê bao (ví dụ, thẻ SIM 224). Theo một phương án của sáng chế, môđun mạng tế bào 221 có thể thực hiện ít nhất một vài trong số các chức năng mà bộ xử lý 210 có thể cung cấp. Theo một phương án, môđun mạng tế bào 221 có thể bao gồm bộ xử lý truyền thông (Communication Processor, CP).

Theo một số phương án của sáng chế, ít nhất một số (ví dụ, hai hoặc nhiều) trong số môđun mạng tế bào 221, môđun Wi-Fi 223, môđun BT 225, môđun GNSS 227 và môđun NFC 228 có thể được bao gồm trong một chip tích hợp (IC) hoặc gói IC. Môđun RF 229 có thể truyền/nhận, ví dụ, tín hiệu truyền thông (ví dụ, tín hiệu RF). Môđun RF 229 có thể bao gồm, ví dụ, bộ thu phát, môđun khuếch đại công suất (Power Amplifier Module, PAM), bộ lọc tần số, bộ khuếch đại nhiễu thấp (Low Noise

Amplifier, LNA) và ăng ten. Theo một phương án khác của sáng chế, ít nhất một trong số môđun mạng tế bào 221, môđun Wi-Fi 223, môđun BT 225, môđun GNSS 227 và môđun NFC 228 có thể truyền và nhận tín hiệu RF qua môđun RF riêng biệt. Môđun nhận dạng thuê bao 224 có thể bao gồm, ví dụ, thẻ bao gồm môđun nhận dạng thuê bao và/hoặc SIM nhúng, và có thể bao gồm thông tin nhận dạng đơn nhất (ví dụ, mã nhận dạng thẻ mạch tích hợp (Integrated Circuit Card Identifier, ICCID)) hoặc thông tin thuê bao (ví dụ, mã nhận dạng thuê bao di động quốc tế (International Mobile Subscriber Identity, IMSI)).

Bộ nhớ 230 (ví dụ, bộ nhớ 130) có thể bao gồm, ví dụ, bộ nhớ trong 232 hoặc bộ nhớ ngoài 234. Bộ nhớ nhúng 232 có thể bao gồm ít nhất một trong số bộ nhớ khả biến (ví dụ, bộ nhớ truy cập ngẫu nhiên động (Dynamic Random Access Memory, DRAM), RAM tĩnh (Static RAM, SRAM), RAM động đồng bộ (Synchronous Dynamic RAM, SDRAM), và tương tự) và bộ nhớ không khả biến (ví dụ, bộ nhớ chỉ đọc lập trình được một lần (One Time Programmable Read Only Memory, OTPROM), ROM lập trình được (Programmable ROM, PROM), ROM lập trình được và xóa được (Erasable and Programmable ROM, EPROM), ROM lập trình được và xóa được bằng điện (Electrically Erasable and Programmable ROM, EEPROM), ROM mặt nạ, ROM nhanh, bộ nhớ nhanh (ví dụ, bộ nhớ nhanh NAND hoặc bộ nhớ nhanh NOR), ổ đĩa cứng, ổ cứng thể rắn (Solid State Drive, SSD), và tương tự). Bộ nhớ ngoài 234 có thể còn bao gồm bộ nhớ nhanh (ổ đĩa flash), ví dụ, thẻ nhớ nhanh compac (Compact Flash, CF), thẻ nhớ kỹ thuật số bảo mật (Secure Digital, SD), thẻ nhớ kỹ thuật số bảo mật micro (Micro Secure Digital, Micro-SD), thẻ nhớ kỹ thuật số bảo mật mini (Mini Secure Digital, Mini-SD), thẻ nhớ kỹ thuật số cực biên (eXtreme Digital, xD), thẻ nhớ, hoặc tương tự. Bộ nhớ ngoài 234 có thể được kết nối chức năng và/hoặc vật lý với thiết bị điện tử 201 qua các giao diện khác nhau.

Môđun cảm biến 240 có thể đo đại lượng vật lý hoặc phát hiện trạng thái hoạt động của thiết bị điện tử 201, và có thể chuyển đổi thông tin đo được hoặc phát hiện được thành tín hiệu điện. Ví dụ, môđun cảm biến 240 có thể bao gồm ít nhất một trong số bộ cảm biến cử chỉ 240A, bộ cảm biến con quay hồi chuyển 240B, bộ cảm biến áp suất khí quyển 240C, bộ cảm biến từ 240D, bộ cảm biến gia tốc 240E, bộ cảm biến cảm máy (grip sensor) 240F, bộ cảm biến tiệm cận 240G, bộ cảm biến màu 240H (ví dụ, bộ cảm biến đỏ/lục/lam (Red/Green/Blue, RGB)), bộ cảm biến sinh trắc 240I, bộ

cảm biến nhiệt độ/độ ẩm 240J, bộ cảm biến ánh sáng 240K và bộ cảm biến tia cực tím (Ultra Violet, UV) 240M. Ngoài ra hoặc theo cách khác, môđun cảm biến 240 có thể bao gồm, ví dụ, bộ cảm biến khứu giác điện tử (E-nose), bộ cảm biến ghi điện cơ (electromyography, EMG), bộ cảm biến điện não đồ (electroencephalogram, EEG), bộ cảm biến điện tâm đồ (electrocardiogram, ECG), bộ cảm biến hồng ngoại (Infrared, IR), bộ cảm biến móng mắt và/hoặc bộ cảm biến dấu vân tay. Môđun cảm biến 240 có thể còn bao gồm mạch điều khiển để điều khiển một hoặc nhiều bộ cảm biến được bao gồm trong đó. Theo một số phương án của sáng chế, thiết bị điện tử 201 có thể còn bao gồm bộ xử lý được tạo cấu hình để điều khiển môđun cảm biến 240 dưới dạng một phần của hoặc tách biệt khỏi bộ xử lý 210, và có thể điều khiển môđun cảm biến 240 trong khi bộ xử lý 210 ở trạng thái ngủ.

Thiết bị đầu vào 250 có thể bao gồm, ví dụ, bảng điều khiển chạm 252, bộ cảm biến bút (kỹ thuật số) 254, phím 256 hoặc thiết bị đầu vào siêu âm 258. Bảng điều khiển chạm 252 có thể sử dụng ít nhất một trong số, ví dụ, loại điện dung, loại điện trở, loại hồng ngoại và loại siêu âm. Ngoài ra, bảng điều khiển chạm 252 có thể còn bao gồm mạch điều khiển. Bảng điều khiển chạm 252 có thể còn bao gồm lớp xúc giác và tạo ra phản ứng xúc giác cho người sử dụng. Bộ cảm biến bút (kỹ thuật số) 254 có thể bao gồm, ví dụ, cảm nhận dạng dưới dạng một phần của bảng điều khiển chạm hoặc tách biệt khỏi bảng điều khiển chạm. Phím 256 có thể bao gồm, ví dụ, nút vật lý, phím quang hoặc bàn phím. Thiết bị đầu vào siêu âm 258 có thể phát hiện sóng siêu âm được tạo ra bởi công cụ đầu vào qua micrô (ví dụ, micrô 288) và nhận dạng dữ liệu tương ứng với sóng siêu âm đã phát hiện được.

Màn hiển thị 260 (ví dụ, màn hiển thị 160) có thể bao gồm panel 262, thiết bị tạo ảnh toàn ký 264 hoặc dụng cụ chiếu 266. Panel 262 có thể bao gồm cấu hình giống hoặc tương tự với màn hiển thị 160 được minh họa trên Fig.1. Panel 262 có thể được tạo ra dưới dạng, ví dụ, mềm dẻo, trong suốt hoặc dẻo được. Panel 262 và bảng điều khiển chạm 252 có thể được tạo ra dưới dạng một môđun.

Thiết bị tạo ảnh toàn ký 264 có thể chiếu hình ảnh ba chiều trong không khí bằng cách sử dụng hiện tượng giao thoa ánh sáng. Dụng cụ chiếu 266 có thể hiển thị hình ảnh bằng cách chiếu ánh sáng lên màn hình. Màn hình này có thể được đặt, ví dụ, bên trong hoặc bên ngoài thiết bị điện tử 201. Theo một phương án, màn hiển thị 260 có thể còn bao gồm mạch điều khiển để điều khiển panel 262, thiết bị tạo ảnh toàn ký

264 hoặc dụng cụ chiếu 266. Giao diện 270 có thể bao gồm, ví dụ, giao diện đa phương tiện độ phân giải cao (High-Definition Multimedia Interface, HDMI) 272, bus nối tiếp vạn năng (USB) 274, giao diện quang học 276 hoặc đầu nối D tinh xảo (D-sub) 278. Giao diện 270 có thể được bao gồm trong, ví dụ, giao diện truyền thông 170 được minh họa trên Fig.1. Ngoài ra hoặc theo cách khác, giao diện 270 có thể bao gồm, ví dụ, giao diện liên kết di động độ phân giải cao (Mobile High-definition Link, MHL), giao diện thẻ kỹ thuật số bảo mật (Secure Digital, SD)/thẻ đa phương tiện (Multi-Media Card, MMC), hoặc giao diện chuẩn hiệp hội dữ liệu hồng ngoại (Infrared Data Association, IrDA).

Môđun audio 280 có thể có thể chuyển đổi hai chiều, ví dụ, âm thanh và tín hiệu điện. Ít nhất một số bộ phận của môđun audio 280 có thể được bao gồm trong, ví dụ, giao diện đầu vào/đầu ra 145 được minh họa trên Fig.1. Môđun audio 280 có thể xử lý thông tin âm thanh được nhập vào hoặc kết xuất qua, ví dụ, loa 282, thiết bị thu 284, tai nghe 286, micro 288, hoặc bộ phận tương tự. Môđun camera 291 là thiết bị có thể chụp ảnh tĩnh và ảnh động. Theo một phương án, môđun camera 291 có thể bao gồm một hoặc nhiều bộ cảm biến hình ảnh (ví dụ, bộ cảm biến trước hoặc bộ cảm biến sau), thấu kính, bộ xử lý tín hiệu hình ảnh (Image Signal Processor, ISP) hoặc đèn chớp (ví dụ, LED hoặc đèn xenon).

Môđun quản lý nguồn 295 có thể quản lý, ví dụ, nguồn của thiết bị điện tử 201. Theo một phương án, môđun quản lý nguồn 295 có thể bao gồm mạch tích hợp quản lý nguồn (Power Management Integrated Circuit, PMIC), mạch tích hợp (Integrated Circuit, IC) sạc, hoặc bộ đo pin hoặc nhiên liệu. PMIC có thể sử dụng phương pháp sạc có dây và/hoặc không dây. Ví dụ về phương pháp sạc không dây có thể bao gồm, ví dụ, phương pháp cộng hưởng từ, phương pháp cảm ứng từ, phương pháp điện từ và phương pháp tương tự. Các mạch bổ sung (ví dụ, cuộn dây vòng, mạch cộng hưởng, bộ chỉnh lưu, v.v.) để sạc không dây có thể được thêm vào. Bộ đo pin có thể đo, ví dụ, dung lượng còn lại của pin 296, và điện áp, dòng điện hoặc nhiệt độ trong khi sạc. Pin 296 có thể bao gồm, ví dụ, pin có thể sạc được hoặc pin mặt trời.

Bộ chỉ báo 297 có thể hiển thị trạng thái cụ thể (ví dụ, trạng thái khởi động, trạng thái tin nhắn, trạng thái sạc, hoặc trạng thái tương tự) của thiết bị điện tử 201 hoặc một bộ phận (ví dụ, bộ xử lý 210). Động cơ 298 có thể chuyển đổi tín hiệu điện thành tín hiệu rung cơ học, và có thể tạo ra hiệu ứng rung, hiệu ứng xúc giác, hoặc

hiệu ứng tương tự. Mặc dù không được minh họa, nhưng thiết bị điện tử 201 có thể bao gồm bộ xử lý (ví dụ, GPU) để hỗ trợ vô tuyến (television, TV) di động. Ví dụ, bộ xử lý để hỗ trợ TV di động có thể xử lý dữ liệu đa phương tiện theo tiêu chuẩn nhất định như phát rộng đa phương tiện kỹ thuật số (Digital Multimedia Broadcasting, DMB), phát rộng video kỹ thuật số (Digital Video Broadcasting, DVB) hoặc mediaFLO™. Mỗi bộ phận thành phần được mô tả ở trên của phần cứng theo sáng chế có thể được tạo cấu hình với một hoặc nhiều thành phần, và tên gọi của các bộ phận thành phần tương ứng có thể thay đổi dựa theo loại thiết bị điện tử. Thiết bị điện tử theo các phương án khác nhau của sáng chế có thể bao gồm ít nhất một trong số các bộ phận nêu trên. Một số bộ phận có thể được lược bỏ hoặc các bộ phận bổ sung khác cũng có thể được thêm vào thiết bị điện tử. Ngoài ra, một số thành phần của phần cứng theo các phương án khác nhau có thể được kết hợp vào một thực thể, mà có thể thực hiện các chức năng giống với các chức năng của các thành phần tương ứng trước khi kết hợp.

Fig.3 là sơ đồ khối của môđun chương trình theo các phương án khác nhau của sáng chế.

Theo một phương án, môđun chương trình 310 (ví dụ, chương trình 140) có thể bao gồm hệ điều hành (OS) để điều khiển tài nguyên liên quan đến thiết bị điện tử (ví dụ, thiết bị điện tử 101) và/hoặc các ứng dụng khác (ví dụ, các chương trình ứng dụng 147) được thực thi trong hệ điều hành. Hệ điều hành này có thể là, ví dụ, Android, iOS, Windows, Symbian, Tizen, Bada hoặc hệ điều hành tương tự. Môđun chương trình 310 có thể bao gồm nhân 320, phần sụn 330, API 360 và/hoặc ứng dụng 370. Ít nhất một vài trong số các môđun chương trình 310 có thể được tải sẵn trên thiết bị điện tử, hoặc có thể được tải xuống từ thiết bị điện tử ngoại vi (ví dụ, thiết bị điện tử 102 hoặc 104, hoặc máy chủ 106).

Nhân 320 (ví dụ, nhân 141) có thể bao gồm, ví dụ, trình quản lý tài nguyên hệ thống 321 và/hoặc trình điều khiển thiết bị 323. Trình quản lý tài nguyên hệ thống 321 có thể thực hiện việc điều khiển, phân bổ, truy tìm hoặc tương tự đối với tài nguyên hệ thống. Theo một phương án của sáng chế, trình quản lý tài nguyên hệ thống 321 có thể bao gồm trình quản lý xử lý, trình quản lý bộ nhớ, trình quản lý hệ thống tệp tin, hoặc tương tự. Trình điều khiển thiết bị 323 có thể bao gồm, ví dụ, trình điều khiển hiển thị, trình điều khiển camera, trình điều khiển Bluetooth, trình điều khiển bộ nhớ chia sẻ,

trình điều khiển USB, trình điều khiển bàn phím, trình điều khiển Wi-Fi, trình điều khiển audio, hoặc trình điều khiển truyền thông liên quy trình (Inter-Process Communication, IPC).

Phần sụn 330 có thể cung cấp chức năng được yêu cầu bởi các ứng dụng 370 nói chung hoặc cung cấp các chức năng khác nhau cho các ứng dụng 370 thông qua API 360 sao cho các ứng dụng 370 có thể sử dụng một cách hiệu quả tài nguyên hệ thống giới hạn trong thiết bị điện tử. Theo một phương án, phần sụn 330 (ví dụ, phần sụn 143) có thể bao gồm, ví dụ, ít nhất một trong số thư viện thời gian chạy 335, trình quản lý ứng dụng 341, trình quản lý cửa sổ 342, trình quản lý đa phương tiện 343, trình quản lý tài nguyên 344, trình quản lý nguồn 345, trình quản lý cơ sở dữ liệu 346, trình quản lý gói 347, trình quản lý kết nối 348, trình quản lý thông báo 349, trình quản lý vị trí 350, trình quản lý đồ họa 351 và trình quản lý bảo mật 352.

Thư viện thời gian chạy 335 có thể bao gồm môđun thư viện được sử dụng bởi trình biên dịch để bổ sung chức năng mới thông qua ngôn ngữ lập trình trong khi ứng dụng 370 đang được chạy. Thư viện thời gian chạy 335 có thể thực hiện việc quản lý đầu vào/đầu ra, quản lý bộ nhớ, chức năng cho hàm số học, hoặc chức năng tương tự. Trình quản lý ứng dụng 341 có thể quản lý, ví dụ, thời hạn sử dụng của ít nhất một trong số các ứng dụng 370. Trình quản lý cửa sổ 342 có thể quản lý tài nguyên giao diện đồ họa người sử dụng (Graphical User Interface, GUI) được sử dụng cho màn hình. Trình quản lý đa phương tiện 343 có thể xác định định dạng cần thiết để phát các tệp tin đa phương tiện khác nhau, và có thể mã hóa hoặc giải mã tệp tin đa phương tiện bằng cách sử dụng bộ mã hóa/bộ giải mã (bộ lập-giải mã) phù hợp với định dạng tương ứng.

Trình quản lý tài nguyên 344 có thể quản lý các tài nguyên, như mã nguồn, bộ nhớ, không gian lưu trữ và tài nguyên tương tự của ít nhất một trong số các ứng dụng 370. Trình quản lý nguồn 345 có thể hoạt động cùng với hệ thống đầu vào/đầu ra cơ bản (Basic Input/Output System, BIOS) để quản lý pin hoặc nguồn và có thể cung cấp thông tin nguồn cần cho hoạt động của thiết bị điện tử. Trình quản lý cơ sở dữ liệu 346 có thể tạo, tìm kiếm và/hoặc thay đổi cơ sở dữ liệu được sử dụng bởi ít nhất một trong số các ứng dụng 370. Trình quản lý gói 347 có thể quản lý việc cài đặt hoặc cập nhật ứng dụng được phân phối ở dạng tệp tin gói.

Trình quản lý kết nối 348 có thể quản lý sự kết nối không dây như, ví dụ, Wi-Fi hoặc Bluetooth. Trình quản lý thông báo 349 có thể hiển thị hoặc thông báo sự kiện, như tin nhắn đến, cuộc hẹn, thông báo sự tiếm cận, hoặc tương tự, theo cách không làm phiền đến người sử dụng. Trình quản lý vị trí 350 có thể quản lý thông tin vị trí của thiết bị điện tử. Trình quản lý đồ họa 351 có thể quản lý hiệu ứng đồ họa được cung cấp đến người sử dụng, hoặc giao diện người sử dụng liên quan đến hiệu ứng đồ họa này. Trình quản lý bảo mật 352 có thể cung cấp các chức năng bảo mật cần thiết để bảo mật hệ thống, xác thực người sử dụng, hoặc chức năng tương tự.

Theo một phương án của sáng chế, khi thiết bị điện tử (ví dụ, thiết bị điện tử 101) có chức năng gọi điện thoại, phần sụn 330 có thể còn bao gồm trình quản lý điện thoại để quản lý chức năng gọi thoại hoặc chức năng gọi video của thiết bị điện tử này. Phần sụn 330 có thể bao gồm môđun phần sụn tạo ra tổ hợp các chức năng của các bộ phận nêu trên. Phần sụn 330 có thể cung cấp môđun chuyên dụng cho từng loại OS để cung cấp chức năng chuyên biệt. Ngoài ra, phần sụn 330 có thể tự động xóa một vài trong số các bộ phận hiện có, hoặc có thể bổ sung thêm các bộ phận mới. API 360 (ví dụ, API 145) là, ví dụ, tập hợp các chức năng lập trình API, và có thể được cung cấp cấu hình khác nhau tùy theo OS. Ví dụ, trong trường hợp của Android hoặc iOS, một bộ API có thể được cung cấp cho mỗi nền tảng. Trong trường hợp của Tizen, hai hoặc nhiều bộ API có thể được cung cấp cho mỗi nền tảng.

Các ứng dụng 370 (ví dụ, chương trình ứng dụng 147) có thể bao gồm, ví dụ, một hoặc nhiều ứng dụng mà có thể cung cấp các chức năng như màn hình chính (home) 371, ứng dụng quay số 372, ứng dụng SMS/MMS 373, ứng dụng tin nhắn tức thời (IM) 374, ứng dụng trình duyệt 375, ứng dụng camera 376, ứng dụng báo thức 377, ứng dụng danh bạ 378, ứng dụng quay số bằng giọng nói 379, ứng dụng thư điện tử 380, ứng dụng lịch 381, ứng dụng phát đa phương tiện 382, ứng dụng sưu tập ảnh (album) 383, ứng dụng đồng hồ 384, ứng dụng chăm sóc sức khỏe (ví dụ, đo mức độ tập luyện hoặc đường máu), hoặc ứng dụng cung cấp thông tin về môi trường (ví dụ, thông tin về áp suất không khí, độ ẩm hoặc nhiệt độ). Theo một phương án của sáng chế, các ứng dụng 370 có thể bao gồm ứng dụng (dưới đây được gọi là “ứng dụng trao đổi thông tin” để thuận tiện cho việc mô tả) hỗ trợ cho việc trao đổi thông tin giữa thiết bị điện tử (ví dụ, thiết bị điện tử 101) và thiết bị điện tử ngoại vi (ví dụ, thiết bị điện tử 102 hoặc 104). Ứng dụng liên quan đến việc trao đổi thông tin có thể bao gồm, ví dụ,

ứng dụng chuyển tiếp thông báo để chuyển thông tin cụ thể đến thiết bị điện tử ngoại vi, hoặc ứng dụng quản lý thiết bị để quản lý thiết bị điện tử ngoại vi.

Ví dụ, ứng dụng chuyển tiếp thông báo có thể bao gồm chức năng phân phối, đến thiết bị điện tử ngoại vi (ví dụ, thiết bị điện tử 102 hoặc 104), thông tin thông báo được tạo ra bởi các ứng dụng khác (ví dụ, ứng dụng SMS/MMS, ứng dụng thư điện tử, ứng dụng chăm sóc sức khỏe, ứng dụng cung cấp thông tin môi trường, v.v.) của thiết bị điện tử 101. Ngoài ra, ứng dụng chuyển tiếp thông báo có thể nhận thông tin thông báo từ, ví dụ, thiết bị điện tử ngoại vi và cung cấp thông tin thông báo nhận được đến người sử dụng. Ứng dụng quản lý thiết bị có thể quản lý (ví dụ, cài đặt, xóa hoặc cập nhật), ví dụ, chức năng cho ít nhất một bộ phận của thiết bị điện tử ngoại vi (ví dụ, thiết bị điện tử 102 hoặc 104) mà truyền thông với thiết bị điện tử (ví dụ, bật/tắt chính thiết bị điện tử ngoại vi này (hoặc một số bộ phận của nó) hoặc điều chỉnh độ sáng (hoặc độ phân giải) của màn hiển thị), các ứng dụng được thực hiện trong thiết bị điện tử ngoại vi, hoặc các dịch vụ được cung cấp từ thiết bị điện tử ngoại vi (ví dụ, dịch vụ gọi điện thoại hoặc dịch vụ nhắn tin). Theo một phương án, các ứng dụng 370 có thể bao gồm các ứng dụng (ví dụ, ứng dụng chăm sóc sức khỏe của thiết bị y tế di động hoặc tương tự) được chỉ định theo thuộc tính của thiết bị điện tử ngoại vi 102 hoặc 104. Theo một phương án của sáng chế, ứng dụng 370 có thể bao gồm ứng dụng nhận được từ thiết bị điện tử ngoại vi (ví dụ, máy chủ 106, hoặc thiết bị điện tử 102 hoặc 104). Theo một phương án của sáng chế, ứng dụng 370 có thể bao gồm ứng dụng đã tải sẵn hoặc ứng dụng của bên thứ ba mà có thể được tải xuống từ máy chủ. Tên gọi của các bộ phận của môđun chương trình 310 theo các phương án nêu trên của sáng chế có thể thay đổi tùy theo loại OS. Theo các phương án khác nhau của sáng chế, ít nhất một vài trong số các môđun chương trình 310 có thể được thực hiện trong phần mềm, phần sụn, phần cứng, hoặc tổ hợp của hai hoặc nhiều loại trong số này. Ít nhất một vài trong số các môđun chương trình 310 có thể được thực hiện (ví dụ, được thực thi) bằng, ví dụ, bộ xử lý (ví dụ, bộ xử lý 210). Ít nhất một vài trong số các môđun chương trình 310 có thể bao gồm, ví dụ, môđun, chương trình, đoạn chương trình, tập hợp lệnh và/hoặc quy trình thực hiện một hoặc nhiều chức năng.

Thuật ngữ “môđun” trong bản mô tả này có thể, ví dụ, có nghĩa là đơn vị bao gồm một trong số phần cứng, phần mềm và phần sụn, và “môđun” có thể được sử dụng theo cách có thể thay thế được với, ví dụ, thuật ngữ “đơn vị”, “logic”, “khối

logic”, “bộ phận” hoặc “mạch”. “Môđun” có thể là đơn vị tối thiểu của bộ phận thành phần tích hợp hoặc một phần của nó. “Môđun” có thể là đơn vị tối thiểu để thực hiện một hoặc nhiều chức năng hoặc một phần của nó. “Môđun” có thể được thực hiện theo cách cơ học hoặc điện tử. Ví dụ, “môđun” theo sáng chế có thể bao gồm ít nhất một trong số chip mạch tích hợp đặc hiệu ứng dụng (Application-Specific Integrated Circuit, ASIC), mảng cổng lập trình được dạng trường (Field-Programmable Gate Arrays, FPGA) và thiết bị logic lập trình được để thực hiện các hoạt động đã biết hoặc sẽ được phát triển sau đây. Theo các phương án tiêu biểu khác nhau, ít nhất một phần của thiết bị (ví dụ, môđun hoặc chức năng của nó) hoặc phương pháp được thực hiện dưới dạng lệnh được lưu trữ trong vật ghi đọc được bằng máy tính (ví dụ, bộ nhớ (130)) ở dạng môđun chương trình. Nếu lệnh được thực thi bởi bộ xử lý (ví dụ, bộ xử lý (120)), thì bộ xử lý này thực hiện chức năng tương ứng với lệnh này.

Fig.4 là sơ đồ minh họa cấu hình của thiết bị điện tử theo các phương án khác nhau.

Dựa vào Fig.4, thiết bị điện tử 400 có thể bao gồm bộ xử lý 410, micrô 420, camera 430, bộ truyền thông 440, loa 450, màn hiển thị 460 và bộ nhớ 470. Vì các bộ phận được minh họa trên Fig.4 không phải là các bộ phận thiết yếu, nên thiết bị điện tử 400 theo các phương án khác nhau có thể được triển khai để bao gồm nhiều hoặc ít bộ phận hơn so với trên Fig.4.

Bộ xử lý 410 có thể thực hiện vai trò phát hiện giọng nói, nhận dạng giọng nói đã phát hiện được, và thực hiện chức năng tương ứng với giọng nói đã nhận dạng được. Khi nhận được đầu vào thoại từ người sử dụng, bộ xử lý 410 có thể tính thời gian EPD, và khi hết thời gian EPD, bộ xử lý 410 có thể phân tích ý định kết thúc giọng nói của người sử dụng và có thể xác định rằng đầu vào thoại đã được hoàn thành, và có thể cung cấp dịch vụ thoại tương ứng với đầu vào thoại này. Để đạt được điều này, bộ xử lý 410 có thể bao gồm bộ nhận dạng giọng nói 411, bộ kiểm soát thời gian EPD 413, bộ xác định ngôn ngữ 415 và bộ xác định cử chỉ 417. Bộ nhận dạng giọng nói 411, bộ kiểm soát thời gian EPD 413, bộ xác định ngôn ngữ 415 và bộ xác định cử chỉ 417 có thể được bao gồm dưới dạng môđun phần cứng, hoặc có thể được bao gồm dưới dạng môđun phần mềm.

Bộ nhận dạng giọng nói 411 có thể thực hiện vai trò xử lý âm thanh được nhập

vào thông qua micrô 420 thành tín hiệu thoại (ví dụ, xử lý trước), và nhận dạng giọng nói (ví dụ, nhận dạng giọng nói tự động (automatic speech recognition, ASR)). Ngoài ra, bộ nhận dạng giọng nói 411 có thể thực hiện xử lý ngôn ngữ (ví dụ, nắm bắt ngôn ngữ tự nhiên (natural language understanding, NLU), quản lý đối thoại (dialogue management, DM), hoặc chuyển văn bản thành giọng nói (text to speech, TTS)) đối với giọng nói đã nhận dạng được. Bộ nhận dạng giọng nói 411 có thể vận hành kết hợp với máy chủ xử lý giọng nói để xử lý và nhận dạng giọng nói. Tức là, bộ nhận dạng giọng nói 411 có thể truyền tín hiệu thoại đến máy chủ xử lý giọng nói, và có thể nhận thông tin thoại (hoặc thông tin văn bản tương ứng với thông tin thoại) từ máy chủ xử lý giọng nói. Bộ nhận dạng giọng nói 411 có thể truyền dữ liệu về việc liệu giọng nói có được nhận dạng (hoặc được phát hiện) hay không, thông tin thoại hoặc thông tin văn bản tương ứng với thông tin thoại đã nhận dạng được đến bộ kiểm soát thời gian EPD 413 hoặc bộ xác định ngôn ngữ 415. Bộ nhận dạng giọng nói 411 có thể lưu trữ tạm thời thông tin thoại được nhập vào thông qua micrô 420 trong khoảng thời gian EPD trong bộ nhớ 470 (ví dụ, lưu trữ trong bộ đệm).

Khi nhận được đầu vào thoại từ người sử dụng, bộ kiểm soát thời gian EPD 413 có thể tính thời gian EPD. Ví dụ, thời gian EPD có thể là thời gian chờ để xác định sự kết thúc đầu vào thoại. Thời gian EPD có thể được hiểu là thời gian chờ đầu vào hoặc thời gian chờ ghi âm micrô. Ví dụ, bộ kiểm soát thời gian EPD 413 có thể bắt đầu tính từ thời điểm phát hiện đầu vào thoại (ví dụ, đầu vào thoại thứ nhất) (hoặc thời điểm kết thúc đầu vào thoại thứ nhất), và khi đầu vào thoại (ví dụ, đầu vào thoại thứ hai) được phát hiện (hoặc đầu vào thoại thứ hai được kết thúc) lần nữa trước khi hết thời gian EPD, thì bộ kiểm soát thời gian EPD 413 có thể tính lại thời gian EPD.

Ví dụ, khi thời gian EPD là 5 giây, thì bộ kiểm soát thời gian EPD 413 có thể tính thời gian EPD từ thời điểm phát hiện đầu vào thoại thứ nhất, và khi đầu vào thoại thứ hai được phát hiện tại thời điểm thời gian EPD được tính tới 3 giây, thì bộ kiểm soát thời gian EPD 413 có thể xóa bỏ (bỏ qua) thời gian EPD đã tính đến thời điểm hiện tại và có thể tính lại thời gian EDP từ thời điểm phát hiện đầu vào thoại thứ hai. Theo cách khác, khi không phát hiện được đầu vào thoại bổ sung trong thời gian định sẵn, thì bộ kiểm soát thời gian EPD 413 có thể tính thời gian EPD. Thời gian định sẵn có thể là giống với thời gian EPD hoặc có thể ngắn hơn hoặc dài hơn thời gian EDP. Theo cách khác, khi không phát hiện được đầu vào thoại bổ sung trong thời gian định

sẵn, thì bộ kiểm soát thời gian EPD 413 có thể xác định rằng thời gian EPD đã hết.

Khi hết thời gian EPD, bộ kiểm soát thời gian EPD 413 có thể thông báo cho bộ xác định ngôn ngữ 415 hoặc bộ xác định cử chỉ 417 rằng thời gian EPD đã hết. Ngoài ra, bộ kiểm soát thời gian EPD 413 có thể kéo dài thời gian EPD theo kết quả nhận được từ bộ xác định ngôn ngữ 415 hoặc bộ xác định cử chỉ 417.

Bộ kiểm soát thời gian EPD 413 theo các phương án khác nhau có thể thiết lập thời gian EPD khác nhau theo ngữ cảnh (hoặc trạng thái) của thiết bị điện tử 400 hoặc đặc trưng của người sử dụng. Ngoài ra, bộ kiểm soát thời gian EPD 413 có thể điều chỉnh thời gian kéo dài của thời gian EPD khác nhau theo ngữ cảnh (hoặc trạng thái) của thiết bị điện tử 400 hoặc đặc trưng của người sử dụng. Tức là, bộ kiểm soát thời gian EPD 413 có thể thay đổi thời gian EPD hoặc thời gian kéo dài EPD theo đặc điểm của ứng dụng mà người sử dụng muốn thực thi để nhận dịch vụ thoại hoặc đặc điểm cá nhân của người sử dụng. Bộ kiểm soát thời gian EPD 413 theo các phương án khác nhau có thể thiết lập thời gian EPD hoặc thời gian kéo dài EPD khác nhau dựa trên ít nhất một trong số khoảng thời gian mà người sử dụng không nói, sự thay đổi về thông tin người sử dụng hoặc sự thay đổi về tốc độ nói của người sử dụng.

Bộ xác định ngôn ngữ 415 có thể xác định (hoặc nắm bắt) ý định kết thúc giọng nói của người sử dụng dựa trên từ định sẵn được lưu trữ trong bộ nhớ 470. Ví dụ, bộ xác định ngôn ngữ 415 có thể xác định xem từ cuối cùng được nhập vào bằng giọng nói có phải là từ định sẵn hay không. Ví dụ, người sử dụng có thể do dự vì người sử dụng không nghĩ ra từ chính xác, ý định của người sử dụng không rõ ràng, hoặc người sử dụng không biết chức năng nào có thể thực hiện được bằng dịch vụ thoại. Trong trường hợp này, người sử dụng có thể nói từ định sẵn bao gồm ít nhất một trong số từ vô nghĩa (ví dụ, từ rỗng), liên từ (ví dụ, “so”, “therefore”, “and”, v.v.) hoặc lệnh chờ (ví dụ, “just a moment”, “wait”, v.v.) giữa quá trình nhập giọng nói. Tuy nhiên, từ định sẵn có thể là ngôn ngữ không bắt buộc để nhận dạng giọng nói, mà có thể thể hiện ý định tiếp tục nhập giọng nói của người sử dụng. Khi hết thời gian EPD, nhưng ngôn ngữ cuối cùng được nói bởi người sử dụng là từ định sẵn, thì bộ xác định ngôn ngữ 415 có thể kiểm soát để kéo dài thời gian EPD. Tức là, khi từ cuối cùng được nhập vào là từ định sẵn, thì bộ xác định ngôn ngữ 415 có thể ra lệnh cho bộ kiểm soát thời gian EPD 413 kéo dài thời gian EPD. Khi thời gian EPD được kéo dài bởi bộ kiểm soát thời gian EPD 413, thì bộ nhận dạng giọng nói 411 có thể chờ nhận đầu vào

thoại mà không ngừng nhận đầu vào thoại.

Bộ xác định cử chỉ 417 có thể xác định (hoặc nắm bắt) ý định kết thúc giọng nói của người sử dụng dựa trên cử chỉ định sẵn được lưu trữ trong bộ nhớ 470. Bộ xác định cử chỉ 417 có thể xác định xem cử chỉ của người sử dụng đã phát hiện được từ camera 430 hoặc bộ cảm biến (ví dụ, bộ cảm biến chuyển động) có phải là cử chỉ định sẵn hay không. Người sử dụng có thể tạo ra cử chỉ cụ thể trong khi nghĩ về từ mà người sử dụng định nói vì người sử dụng không nghĩ ra từ chính xác giữa quá trình nhập giọng nói. Ví dụ, người sử dụng có thể nhìn lên hoặc nghiêng đầu sang trái hoặc phải trong khi nghĩ. Khi hết thời gian EPD, nhưng cử chỉ của người sử dụng đã phát hiện được là cử chỉ định sẵn, thì bộ xác định cử chỉ 417 có thể kiểm soát để kéo dài thời gian EPD. Tức là, khi cử chỉ của người sử dụng đã phát hiện được là cử chỉ định sẵn, thì bộ xác định cử chỉ 417 có thể ra lệnh cho bộ kiểm soát thời gian EPD 413 kéo dài thời gian EPD. Khi thời gian EPD được kéo dài bởi bộ kiểm soát thời gian EPD 413, thì bộ nhận dạng giọng nói 411 có thể chờ nhận đầu vào thoại mà không ngừng nhận đầu vào thoại.

Micrô 420 có thể nhận đầu vào là tín hiệu âm thanh bên ngoài và có thể xử lý tín hiệu này thành dữ liệu giọng nói điện. Trong micrô 420, các thuật toán khử nhiễu khác nhau có thể được thực hiện để loại bỏ nhiễu xuất hiện trong quá trình nhận đầu vào là tín hiệu âm thanh bên ngoài. Micrô 420 có thể thực hiện chức năng nhập luồng âm thanh như lệnh thoại (lệnh thoại để kích hoạt/hủy kích hoạt dịch vụ thoại, v.v.). Micrô 420 có thể được kích hoạt trong thời gian EPD.

Camera 430 (ví dụ, môđun camera 291 trên Fig.2) có thể thể hiện cấu hình để hỗ trợ chức năng quay chụp của thiết bị điện tử 400. Camera 430 có thể quay chụp đối tượng nhất định dưới sự điều khiển của bộ xử lý 410 và có thể truyền dữ liệu quay chụp được (ví dụ, hình ảnh) đến màn hiển thị 460 và bộ xử lý 410.

Bộ truyền thông 440 có thể có cấu hình giống như hoặc tương tự với môđun truyền thông 220 trên Fig.2. Bộ truyền thông 440 có thể bao gồm một hoặc nhiều môđun cho phép truyền thông không dây giữa thiết bị điện tử 400 và thiết bị ngoại vi (ví dụ, thiết bị điện tử khác 102, 104, máy chủ xử lý giọng nói, v.v.). Ví dụ, bộ truyền thông 440 có thể được tạo cấu hình để bao gồm môđun truyền thông di động, môđun mạng vùng cục bộ (LAN) không dây, môđun truyền thông tầm gần và môđun tính toán

vị trí. Theo các phương án khác nhau, bộ truyền thông 440 có thể bao gồm môđun để truyền thông với các thiết bị ngoại vi lân cận (ví dụ, môđun truyền thông tầm gần, môđun truyền thông tầm xa, v.v.).

Loa 450 có thể kết xuất dữ liệu âm thanh nhận được từ bộ truyền thông 440 hoặc được lưu trữ trong bộ nhớ 470. Loa 450 có thể kết xuất tín hiệu âm thanh giọng nói liên quan đến các hoạt động (các chức năng) khác nhau được thực hiện trong thiết bị điện tử 400.

Màn hiển thị 460 có thể hiển thị (kết xuất) các thông tin đã được xử lý trong thiết bị điện tử 400. Ví dụ, màn hiển thị 460 có thể hiển thị các giao diện người sử dụng (UI) hoặc UI đồ họa (GUI) liên quan đến việc sử dụng thiết bị điện tử 400. Theo các phương án khác nhau, màn hiển thị 460 có thể cung cấp các cấu hình màn hình khác nhau để phản hồi trạng thái kích hoạt (hoặc hết thời gian) đến người sử dụng trong dịch vụ thoại. Màn hiển thị 460 có thể sử dụng các màn hiển thị khác nhau (ví dụ, màn hiển thị 160 trên Fig.1). Theo các phương án khác nhau, màn hiển thị 460 có thể sử dụng màn hình cong.

Bộ nhớ 470 (ví dụ, bộ nhớ 130, 230 trên Fig.1 hoặc Fig.2) có thể lưu trữ một hoặc nhiều chương trình được thực thi bởi bộ xử lý 410, và có thể thực hiện chức năng để lưu trữ tạm thời dữ liệu được nhập vào/kết xuất (ví dụ, đệm). Dữ liệu được nhập vào/kết xuất có thể bao gồm, ví dụ, video, hình ảnh, ảnh hoặc âm thanh. Bộ nhớ 470 có thể lưu trữ ít nhất một trong số từ định sẵn (ví dụ, từ thông dụng, từ cá nhân), cử chỉ định sẵn, giá trị trọng số hoặc điểm trong từng điều kiện để phân tích ý định kết thúc giọng nói, thời gian EPD theo ngữ cảnh của thiết bị điện tử hoặc đặc điểm cá nhân, và thời gian kéo dài EPD theo ngữ cảnh của thiết bị điện tử hoặc đặc điểm cá nhân.

Từ định sẵn theo các phương án khác nhau có thể là từ thông dụng thường được sử dụng, hoặc được thiết lập khác nhau theo đặc trưng của người sử dụng không giống như từ thông dụng. Từ thông dụng có thể thay đổi theo quốc gia (ví dụ, theo đặc điểm ngôn ngữ của mỗi quốc gia). Bộ nhớ 470 có thể bao gồm các từ thông dụng của mỗi quốc gia theo phân loại. Theo cách khác, các từ thông dụng được lưu trữ trong bộ nhớ 470 có thể thay đổi theo quốc gia mà thiết bị điện tử 400 được bán ở đó.

Bảng 1 thể hiện các từ thông dụng của từng ngôn ngữ được lưu trữ trong bộ nhớ 470.

Bảng 1

Phân loại	Tiếng Hàn	Tiếng Anh	Tiếng Trung	Tiếng Đức
Từ rỗng	응, 음, 어, 그, 에	Um, Ahh, well, Hmm	这个 zhe ge, 那个 na ge...	ah/s:/, so/zo :/, tja, halt, eigentlich
Liên từ	그러니까, 그래서, 그리고, 그런데	For, And, Nor, But, Or Yet, So	从而 cóng'ér (do đó), 但是 dànshì (nhưng), 而且 érqǐě (và), 反而 fǎn'ér (thay vào đó)	Und (và), aber (nhưng), oder (hoặc), denn (bởi vì)
Lệnh chờ	잠깐만, 기다려봐, 잠만	wait, wait a second, hold on	等 一 下 děngyīxià	Entschuldigung

Bộ xử lý 410 có thể tiếp tục cập nhật các từ cá nhân được nhập vào khác nhau theo đặc trưng của người sử dụng thông qua việc học bởi thiết bị điện tử 400. Ví dụ, bộ nhớ 470 có thể lưu trữ các từ thông dụng và các từ cá nhân như được thể hiện trong bảng 2.

Bảng 2 thể hiện các ngôn ngữ thông dụng và các ngôn ngữ cá nhân được lưu trữ trong bộ nhớ 470.

Bảng 2

DB thông dụng	DB cá nhân		
	Người sử dụng A	Người sử dụng B	Người sử dụng C
Yeah, Um, Well, Hum, ...	so..., what it is..., so to speak...	Whatchamacallit, so, whatsit...	What was it..., and then...

Ví dụ, một số từ cá nhân của người sử dụng A sử dụng thiết bị điện tử 101, người sử dụng B sử dụng thiết bị điện tử 102 và người sử dụng C sử dụng thiết bị điện tử 104 có thể là tương tự với nhau hoặc khác nhau. Theo các phương án khác nhau, bộ xử lý 410 có thể chia sẻ các từ cá nhân của chính nó với người sử dụng khác theo định kỳ hoặc theo thiết lập, và có thể cập nhật các từ thông dụng bằng cách phân tích các từ cá nhân của người sử dụng khác này. Ví dụ, bộ xử lý 410 có thể cập nhật các từ giống nhau giữa các từ cá nhân của người sử dụng A và các từ cá nhân của người sử dụng B làm từ thông dụng.

Theo các phương án khác nhau, cử chỉ được tạo ra khi người sử dụng đang

ngữ có thể bao gồm cử chỉ thông dụng chung, hoặc thay đổi theo đặc trưng của người sử dụng không giống như cử chỉ thông dụng. Cử chỉ cá nhân thay đổi theo đặc trưng của người sử dụng có thể tiếp tục được cập nhật thông qua việc học bởi thiết bị điện tử 400. Cử chỉ định sẵn có thể thay đổi theo quốc gia (theo đặc điểm cử chỉ của mỗi quốc gia). Bộ nhớ 470 có thể lưu trữ các cử chỉ thông dụng của mỗi quốc gia. Theo cách khác, các cử chỉ thông dụng được lưu trữ trong bộ nhớ 470 có thể thay đổi theo quốc gia mà thiết bị điện tử 400 được bán ở đó.

Thiết bị điện tử 400 theo các phương án khác nhau có thể bao gồm micrô 420, bộ nhớ 470, và bộ xử lý 410 được kết nối chức năng với micrô hoặc bộ nhớ, và bộ xử lý này có thể được tạo cấu hình để tính thời gian EPD dựa trên đầu vào thoại, để xác định xem từ cuối cùng của đầu vào thoại có tương ứng với từ định sẵn được lưu trữ trong bộ nhớ khi hết thời gian EPD hay không, và để kéo dài thời gian EPD khi từ cuối cùng tương ứng với từ định sẵn, và để chờ nhận đầu vào thoại.

Bộ xử lý có thể được tạo cấu hình để kéo dài thời gian EPD khi từ cuối cùng tương ứng với từ định sẵn bao gồm ít nhất một trong số từ rỗng, liên từ hoặc lệnh chờ.

Bộ xử lý có thể được tạo cấu hình để kéo dài thời gian EPD khi phát hiện được đầu vào thoại bổ sung trước khi hết thời gian EPD.

Từ định sẵn có thể bao gồm từ thông dụng và từ cá nhân, và bộ xử lý có thể được tạo cấu hình để xác định mức độ tương tự giữa lệnh thoại được nhận dạng sau khi lỗi lệnh thoại và lệnh thoại trước đó, và để thu thập từ cá nhân dựa trên mức độ tương tự này.

Bộ xử lý có thể được tạo cấu hình để phân tích thông tin văn bản thay đổi giữa lệnh thoại và lệnh thoại trước đó, và để cập nhật thông tin văn bản bằng từ cá nhân khi thông tin văn bản thay đổi được phát hiện với số lần định sẵn hoặc nhiều hơn.

Bộ xử lý có thể được tạo cấu hình để xác định xem có phát hiện được cử chỉ định sẵn khi hết thời gian EPD hay không, và để kéo dài thời gian EPD khi phát hiện được cử chỉ định sẵn.

Bộ xử lý có thể được tạo cấu hình để xác định xem câu theo đầu vào thoại có được hoàn thành khi hết thời gian EPD hay không, và để kéo dài thời gian EPD khi

xác định được rằng câu này không được hoàn thành.

Bộ xử lý có thể được tạo cấu hình để xác định xem có thực hiện bước xác định xem câu có được hoàn thành hay không, dựa trên loại lệnh thoại theo đầu vào thoại.

Bộ xử lý có thể được tạo cấu hình để kéo dài thời gian EPD theo giá trị cố định, hoặc để thay đổi thời gian EPD đến giá trị tương ứng với sự nhận dạng ngữ cảnh, và để kéo dài thời gian EPD theo giá trị đã thay đổi.

Bộ xử lý có thể được tạo cấu hình để xác định thời gian EPD hoặc thời gian kéo dài EPD, dựa trên thông tin ngữ cảnh của thiết bị điện tử và thông tin đặc trưng của người sử dụng.

Bộ xử lý có thể được tạo cấu hình để phân tích ý định kết thúc giọng nói của người sử dụng dựa trên ít nhất một trong số thông tin ngữ cảnh của thiết bị điện tử, thông tin đặc trưng của người sử dụng, việc có phát hiện được đầu vào thoại bổ sung hay không, việc có phát hiện được từ định sẵn hay không, việc có phát hiện được cử chỉ định sẵn hay không, hoặc việc có hoàn thành câu hay không.

Bộ xử lý có thể được tạo cấu hình để gán giá trị trọng số hoặc điểm cho ít nhất một trong số thời gian phát hiện sự im lặng, việc có phát hiện được từ định sẵn hay không, việc có phát hiện được cử chỉ định sẵn hay không, hoặc việc có hoàn thành câu hay không, và để kéo dài thời gian EPD khi chỉ số ý định kết thúc lớn hơn hoặc bằng chỉ số định sẵn.

Bộ xử lý có thể được tạo cấu hình để thay đổi thời gian kéo dài EPD dựa trên chỉ số ý định kết thúc.

Fig.5A và Fig.5B là các lưu đồ minh họa phương pháp vận hành thiết bị điện tử theo các phương án khác nhau.

Fig.5A minh họa phương pháp vận hành để kiểm soát thời gian EPD khi không phát hiện được đầu vào thoại bổ sung trong thời gian định sẵn.

Dựa vào Fig.5B, ở bước 501, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể nhận đầu vào thoại. Bộ xử lý 410 (ví dụ, bộ nhận dạng giọng nói 411) có thể phát hiện xem liệu âm thanh có được nhập vào từ micrô 420 hay không. Theo các phương án

khác nhau, bộ xử lý 410 có thể bắt đầu dịch vụ thoại (ví dụ, có thể kích hoạt chức năng thoại), dựa trên phương pháp đánh thức bằng giọng nói. Để đạt được điều này, bộ xử lý 410 có thể còn bao gồm bộ phận xử lý đánh thức (hoặc công cụ đánh thức). Theo các phương án khác nhau, thiết bị điện tử 400 có thể bao gồm bộ xử lý thứ nhất ở chế độ nguồn thấp bao gồm bộ xử lý đánh thức, và bộ xử lý thứ hai bao gồm bộ nhận dạng giọng nói 411, bộ kiểm soát thời gian EPD 413, bộ xác định ngôn ngữ 415 và bộ xác định cử chỉ 417.

Theo một phương án, bộ xử lý 410 có thể sử dụng lệnh thoại cụ thể (ví dụ, từ làm lệnh gọi thoại (ví dụ, từ đánh thức), và có thể không thực hiện việc nhận dạng giọng nói đối với giọng nói của người sử dụng trước khi lệnh gọi thoại được nhập vào. Lệnh gọi thoại (ví dụ, từ đánh thức) có thể sử dụng, ví dụ, tên gọi của thiết bị điện tử (hoặc công cụ hỗ trợ của thiết bị điện tử (hoặc trí tuệ nhân tạo (artificial intelligence, AI)) làm từ đánh thức. Người sử dụng nên bao gồm lệnh gọi thoại để sử dụng công cụ nhận dạng giọng nói (ví dụ, công cụ hỗ trợ giọng nói) để sử dụng dịch vụ thoại trong thiết bị điện tử 400, và lệnh gọi thoại có thể được gọi là lệnh đánh thức hoặc từ đánh thức. Khi lệnh gọi thoại được nhận dạng, thì thiết bị điện tử có thể điều khiển công cụ nhận dạng giọng nói và sau đó có thể nhận dạng giọng nói của người sử dụng và cung cấp dịch vụ thoại.

Theo các phương án khác nhau, bộ xử lý 410 có thể bắt đầu dịch vụ thoại dựa trên việc chọn nút cụ thể (ví dụ, nút màn hình chính (home)) hoặc việc nhập vào của người sử dụng như việc nhập lệnh cụ thể vào. Ngoài ra, khi phát hiện được cử chỉ, khuôn mặt hoặc mắt định sẵn để gọi dịch vụ thoại, bộ xử lý 410 có thể bắt đầu dịch vụ thoại. Bộ xử lý 410 (ví dụ, bộ nhận dạng giọng nói 411) có thể bắt đầu dịch vụ thoại, có thể phát hiện âm thanh được nhập vào từ micrô 420, và có thể thực hiện việc xử lý tín hiệu thoại. Việc xử lý tín hiệu thoại chỉ việc xử lý trước để nhận dạng giọng nói, và ví dụ, có thể chỉ việc chuyển đổi âm thanh thành tín hiệu thoại. Ngoài ra, bộ xử lý 410 (ví dụ, bộ nhận dạng giọng nói 411) có thể nhận dạng giọng nói liên quan đến đầu vào thoại đã phát hiện được. Ví dụ, bộ xử lý 410 (ví dụ, bộ nhận dạng giọng nói 411) có thể xử lý ngôn ngữ liên quan đến tín hiệu thoại đã được chuyển đổi và có thể chuyển đổi tín hiệu thoại thành văn bản (ví dụ, giọng nói thành văn bản). Vì công nghệ nhận dạng giọng nói đã được biết rõ, nên mô tả chi tiết về công nghệ này được lược bỏ. Bộ xử lý 410 có thể vận hành micrô 420 và có thể ghi âm giọng nói nhận được từ micrô

420.

Ở bước 503, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể xác định xem có phát hiện được đầu vào thoại trong thời gian định sẵn hay không. Ví dụ, thời gian định sẵn có thể là giống với thời gian EPD hoặc có thể là ngắn hơn hoặc dài hơn thời gian EPD. Khi không phát hiện được đầu vào thoại trong thời gian định sẵn, thì bộ xử lý 410 (ví dụ, bộ kiểm soát thời gian EPD 413) có thể tính thời gian EPD. Theo cách khác, khi không phát hiện được đầu vào thoại trong thời gian định sẵn (ví dụ, thời gian im lặng), thì bộ xử lý 410 có thể xác định rằng thời gian EPD đã hết. Ví dụ, khi thời gian định sẵn không giống với thời gian EPD và không phát hiện được đầu vào thoại trong thời gian định sẵn, thì bộ xử lý 410 có thể tính thời gian EPD. Theo cách khác, khi thời gian định sẵn là giống với thời gian EPD và không phát hiện được đầu vào thoại trong thời gian định sẵn, thì bộ xử lý 410 có thể xác định rằng thời gian EPD đã hết. Tức là, khi thời gian im lặng trong đó không phát hiện được đầu vào thoại là giống với thời gian EPD, bộ xử lý 410 có thể xác định rằng thời gian EPD đã hết.

Khi phát hiện được đầu vào thoại trong thời gian định sẵn, thì bộ xử lý 410 có thể thực hiện bước 511, và khi không phát hiện được đầu vào thoại trong thời gian định sẵn, thì bộ xử lý 410 có thể thực hiện bước 505.

Khi phát hiện được đầu vào thoại trong thời gian định sẵn, thì thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể kéo dài thời gian EPD ở bước 511. Khi phát hiện được đầu vào thoại mới trong thời gian định sẵn, thì bộ xử lý 410 (ví dụ, bộ kiểm soát thời gian EPD 413) có thể tính lại thời gian EPD. Bộ xử lý 410 có thể kéo dài thời gian EPD và có thể quay trở lại bước 503.

Khi không phát hiện được đầu vào thoại trong thời gian định sẵn, thiết bị điện tử (ví dụ, bộ xử lý 410) có thể xác định xem từ cuối cùng có phải là từ được chọn hay không ở bước 505. Từ cuối cùng có thể là toàn bộ hoặc một phần của thông tin thoại được nhập vào cuối cùng bởi người sử dụng. Tức là, bộ xử lý 410 có thể hiểu, dựa trên thông tin văn bản tương ứng với thông tin thoại cuối cùng, là từ cuối cùng. Ví dụ, bộ xử lý 410 có thể xác định xem từ cuối cùng có tương ứng với từ định sẵn bao gồm ít nhất một trong số từ rỗng (ví dụ, “yeah”, “um”, “ahh”, “well”, “hmm”...), liên từ (ví dụ, “so”, “therefore”, “but”, v.v.), hoặc lệnh chờ (ví dụ, “wait”, “wait a second”, “hold on”, v.v.) hay không.

Khi từ cuối cùng là từ đã chọn, bộ xử lý 410 có thể thực hiện bước 511, và khi từ cuối cùng không phải là từ đã chọn, bộ xử lý 410 có thể thực hiện bước 507.

Ở bước 507, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể xác định rằng đầu vào thoại đã được hoàn thành. Ví dụ, khi hết thời gian EPD và từ cuối cùng không phải là từ đã chọn, bộ xử lý 410 (ví dụ, bộ nhận dạng giọng nói 411) có thể xác định rằng người sử dụng có ý định kết thúc đầu vào thoại. Khi đầu vào thoại đã được hoàn thành, bộ xử lý 410 có thể hủy kích hoạt (ví dụ, không vận hành) micrô 420. Tức là, bộ xử lý 410 có thể ngừng nhận đầu vào thoại. Bộ xử lý 410 có thể ngừng ghi âm giọng nói nhận được từ micrô 420.

Ở bước 509, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể xử lý nhận dạng giọng nói. Bộ xử lý 410 (ví dụ, bộ nhận dạng giọng nói 411) có thể cung cấp dịch vụ theo đầu vào thoại từ người sử dụng. Ví dụ, bộ xử lý 410 có thể truyền tệp tin thoại (ví dụ, tín hiệu thoại) ghi âm giọng nói nhận được từ micrô 420 đến máy chủ xử lý giọng nói, và có thể nhận lệnh tương ứng với tệp tin thoại từ máy chủ xử lý giọng nói và có thể thực hiện chức năng (ví dụ, dịch vụ thoại) tương ứng với lệnh đã nhận được. Theo các phương án khác nhau, bộ xử lý 410 có thể cung cấp các chức năng khác nhau tương ứng với giọng nói đã nhận dạng được bằng cách sử dụng công nghệ nhận dạng giọng nói. Ví dụ, bộ xử lý 410 có thể cung cấp các chức năng khác nhau, như tìm kiếm trên thiết bị di động, quản lý lịch trình, gọi điện, ghi nhớ hoặc phát nhạc, dựa trên văn bản được chuyển đổi từ giọng nói.

Fig.5B minh họa phương pháp vận hành để kiểm soát thời gian EPD theo dữ liệu về việc liệu đầu vào thoại có được phát hiện thêm trong khoảng thời gian EPD hay không.

Dựa vào Fig.5B, ở bước 551, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể nhận đầu vào thoại. Bộ xử lý 410 (ví dụ, bộ nhận dạng giọng nói 411) có thể phát hiện xem liệu âm thanh có được nhập vào từ micrô 420 hay không. Theo các phương án khác nhau, bộ xử lý 410 có thể bắt đầu dịch vụ thoại (ví dụ, có thể kích hoạt chức năng thoại) dựa trên phương pháp đánh thức bằng giọng nói.

Theo các phương án khác nhau, bộ xử lý 410 có thể bắt đầu dịch vụ thoại dựa trên việc chọn nút cụ thể (ví dụ, nút màn hình chính (home)), hoặc việc nhập vào của người sử dụng như việc nhập lệnh cụ thể vào. Ngoài ra, khi phát hiện được cử chỉ,

khuôn mặt hoặc mắt định sẵn để gọi dịch vụ thoại, bộ xử lý 410 có thể bắt đầu dịch vụ thoại. Bộ xử lý 410 (ví dụ, bộ nhận dạng giọng nói 411) có thể bắt đầu dịch vụ thoại, có thể phát hiện âm thanh được nhập vào từ micrô 420, và có thể thực hiện việc xử lý tín hiệu thoại. Bộ xử lý 410 có thể vận hành micrô 420 và ghi âm giọng nói nhận được từ micrô 420.

Ở bước 553, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể tính thời gian EPD. Ví dụ, thời gian EPD có thể là thời gian chờ để xác định rằng đầu vào thoại đã kết thúc. Bộ xử lý 410 (ví dụ, bộ kiểm soát thời gian EPD 413) có thể bắt đầu tính thời gian EPD từ thời điểm phát hiện được đầu vào thoại (ví dụ, đầu vào thoại ban đầu) hoặc thời điểm kết thúc đầu vào thoại (ví dụ, đầu vào thoại ban đầu).

Ở bước 555, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể xác định xem thời gian EPD đã hết chưa. Bộ xử lý 410 (ví dụ, bộ kiểm soát thời gian EPD 413) có thể tính thời gian EPD và sau đó có thể tiếp tục xác định xem thời gian EPD đã hết chưa. Khi phát hiện được đầu vào thoại mới, thời gian EPD có thể được tính lại. Bộ xử lý 410 có thể tiếp tục xác định xem có phát hiện được đầu vào thoại mới trước khi hết thời gian EPD hay không.

Khi chưa hết thời gian EPD, bộ xử lý 410 có thể thực hiện bước 563, và khi hết thời gian EPD, bộ xử lý 410 có thể thực hiện bước 557.

Khi chưa hết thời gian EPD, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể xác định xem có phát hiện được đầu vào thoại ở bước 563 hay không. Bộ xử lý 410 (ví dụ, bộ nhận dạng giọng nói 411) có thể xác định xem có phát hiện được đầu vào thoại mới trước khi hết thời gian EPD hay không.

Khi phát hiện được đầu vào thoại, bộ xử lý 410 có thể thực hiện bước 565, và khi không phát hiện được đầu vào thoại, bộ xử lý 410 có thể quay trở lại bước 555. Tức là, bộ xử lý 410 có thể thực hiện bước 557 hoặc 565 bằng cách lặp lại việc xác định ở bước 555 và bước 563 trong thời gian EPD.

Khi phát hiện được đầu vào thoại, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể kéo dài thời gian EPD ở bước 565. Khi phát hiện được đầu vào thoại trong thời gian EPD, bộ xử lý 410 (ví dụ, bộ kiểm soát thời gian EPD 413) có thể tính lại thời gian EPD. Ví dụ, khi thời gian EPD là 3 giây và phát hiện được đầu vào thoại mới

trong vòng 3 giây, bộ xử lý 410 có thể kéo dài thời gian EPD thêm 3 giây từ thời điểm phát hiện được đầu vào thoại mới. Bộ xử lý 410 có thể kéo dài thời gian EPD và quay trở lại bước 503. Theo đó, bộ xử lý 410 có thể lặp lại bước 503, 505, 513 hoặc 515 sau khi nhận đầu vào thoại ban đầu.

Theo các phương án khác nhau, bộ xử lý 410 có thể thiết lập thời gian EPD khác nhau theo ngữ cảnh (hoặc trạng thái) của thiết bị điện tử 400 hoặc đặc trưng của người sử dụng (ví dụ, tốc độ nói, thói quen nói). Ngoài ra, bộ kiểm soát thời gian EPD 413 có thể điều chỉnh thời gian được kéo dài cho thời gian EPD (ví dụ, thời gian kéo dài EPD) khác nhau theo ngữ cảnh (hoặc trạng thái) của thiết bị điện tử 400 hoặc đặc trưng của người sử dụng.

Khi hết thời gian EPD, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể xác định xem từ cuối cùng có phải là từ được chọn hay không ở bước 557. Từ cuối cùng có thể là toàn bộ hoặc một phần của thông tin thoại được nhập vào cuối cùng bởi người sử dụng. Tức là, bộ xử lý 410 có thể hiểu thông tin văn bản tương ứng với thông tin thoại cuối cùng là từ cuối cùng. Ví dụ, khi không phát hiện được đầu vào thoại sau khi từ cuối cùng và thời gian EPD đã hết, bộ xử lý 410 (ví dụ, bộ xác định ngôn ngữ 415) có thể xác định xem từ cuối cùng có phải là từ định sẵn hay không. Ví dụ, bộ xử lý 410 có thể xác định xem từ cuối cùng có tương ứng với từ định sẵn bao gồm ít nhất một trong số từ rỗng (ví dụ, “yeah”, “um”, “ahh”, “well”, “hmm”...), liên từ (ví dụ, “so”, “therefore”, “but”, v.v.), hoặc lệnh chờ (ví dụ, “wait”, “wait a second”, “hold on”, v.v.) hay không.

Khi từ cuối cùng là từ đã chọn, bộ xử lý 410 có thể thực hiện bước 565, và khi từ cuối cùng không phải là từ đã chọn, bộ xử lý 410 có thể thực hiện bước 559.

Ở bước 559, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể xác định rằng đầu vào thoại đã được hoàn thành. Ví dụ, khi hết thời gian EPD và từ cuối cùng không phải là từ đã chọn, bộ xử lý 410 (ví dụ, bộ nhận dạng giọng nói 411) có thể xác định rằng người sử dụng có ý định kết thúc đầu vào thoại. Khi đầu vào thoại đã được hoàn thành, bộ xử lý 410 có thể hủy kích hoạt (ví dụ, có thể không vận hành) micrô 420. Tức là, bộ xử lý 410 có thể ngừng nhận đầu vào thoại. Bộ xử lý 410 có thể ngừng ghi âm giọng nói nhận được từ micrô 420.

Ở bước 561, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể xử lý nhận dạng

giọng nói. Bộ xử lý 410 (ví dụ, bộ nhận dạng giọng nói 411) có thể cung cấp dịch vụ theo đầu vào thoại từ người sử dụng. Ví dụ, bộ xử lý 410 có thể truyền tệp tin thoại (ví dụ, tín hiệu thoại) ghi âm giọng nói nhận được từ micrô 420 đến máy chủ xử lý giọng nói, và có thể nhận lệnh tương ứng với tệp tin thoại từ máy chủ xử lý giọng nói và có thể thực hiện chức năng (ví dụ, dịch vụ thoại) tương ứng với lệnh đã nhận được. Theo các phương án khác nhau, bộ xử lý 410 có thể cung cấp các chức năng khác nhau tương ứng với giọng nói đã nhận dạng được bằng cách sử dụng công nghệ nhận dạng giọng nói. Ví dụ, bộ xử lý 410 có thể cung cấp các chức năng khác nhau, như tìm kiếm trên thiết bị di động, quản lý lịch trình, gọi điện, ghi nhớ hoặc phát nhạc, dựa trên văn bản được chuyển đổi từ giọng nói.

Fig.6A đến Fig.6C là các hình vẽ minh họa các ví dụ về giao diện người sử dụng theo các phương án khác nhau.

Fig.6A minh họa một ví dụ về việc kéo dài thời gian EPD theo việc phát hiện từ rỗng.

Dựa vào Fig.6A, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể nhận lệnh thoại về việc gửi tin nhắn từ người sử dụng. Ví dụ, bộ xử lý 410 có thể nhận đầu vào thoại thứ nhất 613 như “Send message to wife honey I’ll be...uh...well”. Bộ xử lý 410 có thể hiển thị giao diện người sử dụng thứ nhất 610 bao gồm thông tin văn bản 611 tương ứng với đầu vào thoại thứ nhất 613 trên màn hiển thị 460. Theo các phương án khác nhau, bộ xử lý 410 có thể cung cấp một phần hoặc toàn bộ đầu vào thoại thứ nhất 613 dưới dạng thông tin văn bản 611.

Ví dụ, bộ xử lý 410 có thể nhận dạng “Send message” trong đầu vào thoại thứ nhất 613 là lệnh thực hiện ứng dụng tin nhắn, có thể nhận dạng “to wife” trong đầu vào thoại thứ nhất 613 là người nhận tin nhắn, và có thể nhận dạng “honey I’ll be” trong đầu vào thoại thứ nhất 613 là nội dung tin nhắn. Bộ xử lý 410 có thể xử lý thông tin văn bản tương ứng với đầu vào thoại thứ nhất 613 khi phân tích đầu vào thoại thứ nhất 613 và thực hiện chức năng tương ứng với đầu vào thoại này. Ví dụ, bộ xử lý 410 có thể nhận dạng “honey I’ll be...uh... well” trong đầu vào thoại thứ nhất 613 là nội dung tin nhắn, và có thể hiểu “uh....well” là từ vô nghĩa (ví dụ, từ rỗng) và có thể xóa bỏ từ vô nghĩa khi gửi tin nhắn. Bộ xử lý 410 có thể cung cấp màn hình thực hiện của ứng dụng tin nhắn thể hiện rằng người nhận là “wife” và thông tin văn bản được bao

gồm trong nội dung tin nhắn dưới dạng giao diện người sử dụng thứ nhất 610 đến giao diện người sử dụng thứ ba 620.

Theo các phương án khác nhau, bộ xử lý 410 có thể đã thực hiện ứng dụng tin nhắn trước khi nhận được lệnh thoại. Theo cách khác, bộ xử lý 410 có thể nhận lệnh thoại bằng lệnh gọi thoại khi màn hiển thị 460 đã được tắt.

Khi phát hiện được đầu vào thoại thứ nhất 613 hoặc kết thúc đầu vào thoại thứ nhất 613, bộ xử lý 410 có thể tính thời gian EPD và có thể xác định xem có phát hiện được đầu vào thoại mới trước khi hết thời gian EPD hay không. Khi không phát hiện được đầu vào thoại mới trước khi hết thời gian EPD và thời gian EDP đã hết, bộ xử lý 410 có thể xác định xem từ cuối cùng có phải là từ định sẵn hay không. Ví dụ, khi từ cuối cùng (ví dụ, “uh”, “well”) tương ứng với từ định sẵn (ví dụ, từ rỗng) dựa trên thông tin được lưu trữ trong bộ nhớ 470, bộ xử lý 410 có thể kéo dài thời gian EPD và có thể chờ đầu vào thoại từ người sử dụng. Bộ xử lý 410 có thể nhận đầu vào thoại thứ hai 619 từ người sử dụng trong thời gian chờ đầu vào thoại (ví dụ, thời gian EPD). Ví dụ, bộ xử lý 410 có thể nhận đầu vào thoại thứ hai 619 trong khi đang hiển thị giao diện người sử dụng thứ hai 615 bao gồm thông tin văn bản 617 giống hoặc tương tự với thông tin văn bản của giao diện người sử dụng thứ nhất 610 trên màn hiển thị 460. Theo cách khác, bộ xử lý 410 có thể hiển thị thông tin văn bản 621 tương ứng với đầu vào thoại thứ nhất 613 và đầu vào thoại thứ hai 619 trên giao diện người sử dụng thứ hai 615 cùng lúc với việc nhận đầu vào thoại thứ hai 619.

Bộ xử lý 410 có thể hiển thị giao diện người sử dụng thứ ba 620 bao gồm thông tin văn bản 621 tương ứng với đầu vào thoại thứ nhất 613 và đầu vào thoại thứ hai 619 trên màn hiển thị 460. Khi không phát hiện được đầu vào thoại mới sau khi đầu vào thoại thứ hai 619, thời gian EPD đã hết, và từ cuối cùng (ví dụ, “today”) không tương ứng với từ định sẵn, bộ xử lý 410 có thể xử lý nhận dạng giọng nói tương ứng với thông tin văn bản 621. Ví dụ, bộ xử lý 410 có thể truyền thông tin văn bản 621 đến máy chủ xử lý giọng nói, và có thể nhận lệnh để gửi tin nhắn bao gồm thông tin văn bản 621 đến người nhận (ví dụ, “wife”) từ máy chủ xử lý giọng nói, và có thể gửi tin nhắn theo lệnh này. Khi việc gửi tin nhắn được hoàn thành theo quy trình xử lý nhận dạng giọng nói, bộ xử lý 410 có thể hiển thị giao diện người sử dụng thứ tư 625 bao gồm tin nhắn hoàn thành việc gửi 627 trên màn hiển thị 460. Khi quy trình xử lý nhận dạng giọng nói được hoàn thành, bộ xử lý 410 có thể kết xuất thông tin thoại

tương ứng với tin nhắn hoàn thành việc gửi 627 thông qua loa 450.

Fig.6B là hình vẽ minh họa một ví dụ về việc kéo dài thời gian EDP theo việc phát hiện liên từ.

Dựa vào Fig.6B, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể nhận lệnh thoại liên quan đến việc điều khiển thiết bị điện tử ngoại vi từ người sử dụng. Ví dụ, bộ xử lý 410 có thể nhận đầu vào thoại thứ nhất 633 như “Turn on living room light... and”. Bộ xử lý 410 có thể hiển thị giao diện người sử dụng thứ nhất 630 bao gồm thông tin văn bản 631 tương ứng với đầu vào thoại thứ nhất 633 trên màn hiển thị 460. Theo các phương án khác nhau, bộ xử lý 410 có thể cung cấp một phần hoặc toàn bộ đầu vào thoại thứ nhất 633 dưới dạng thông tin văn bản 631. Ví dụ, bộ xử lý 410 có thể nhận dạng “Turn on” trong đầu vào thoại thứ nhất 633 là lệnh điều khiển thiết bị, và có thể nhận dạng “living room light” trong đầu vào thoại thứ nhất 633 là tín hiệu nhận dạng của thiết bị. Theo cách khác, khi đầu vào thoại thứ nhất 633 là lệnh điều khiển thiết bị, bộ xử lý 410 có thể hiểu “and” là từ vô nghĩa (ví dụ, liên từ), và có thể xóa bỏ “and” khi nhận dạng giọng nói.

Theo các phương án khác nhau, bộ xử lý 410 có thể được ghép nối với thiết bị điện tử ngoại vi trước khi nhận lệnh thoại. Theo cách khác, bộ xử lý 410 có thể cố gắng để được ghép nối với thiết bị điện tử ngoại vi theo lệnh gọi thoại khi màn hiển thị 460 đã được tắt. Bộ xử lý 410 có thể cung cấp, dưới dạng giao diện người sử dụng thứ nhất 630 đến giao diện người sử dụng thứ ba 640, màn hình thực hiện của ứng dụng điều khiển thiết bị mà hiển thị kết cấu bố trí của thiết bị điện tử ngoại vi hoặc vị trí của thiết bị điện tử ngoại vi tương ứng với lệnh thoại mà cần phân biệt với các thiết bị điện tử ngoại vi khác.

Khi phát hiện được hoặc kết thúc đầu vào thoại thứ nhất 633, bộ xử lý 410 có thể tính thời gian EPD và có thể xác định xem có phát hiện được đầu vào thoại mới trước khi hết thời gian EDP hay không. Khi không phát hiện được đầu vào thoại mới trước khi hết thời gian EPD và thời gian EDP đã hết, thì bộ xử lý 410 có thể xác định xem từ cuối cùng có phải là từ định sẵn hay không. Ví dụ, khi từ cuối cùng (ví dụ, and) tương ứng với từ định sẵn (ví dụ, liên từ) dựa trên thông tin được lưu trữ trong bộ nhớ 470, thì bộ xử lý 410 có thể kéo dài thời gian EPD và có thể chờ đầu vào thoại từ người sử dụng. Bộ xử lý 410 có thể nhận đầu vào thoại thứ hai 639 từ người sử dụng

trong thời gian chờ đầu vào thoại (ví dụ, thời gian EPD). Ví dụ, bộ xử lý 410 có thể nhận đầu vào thoại thứ hai 639 trong khi đang hiển thị giao diện người sử dụng thứ hai 635 bao gồm thông tin văn bản 637 giống hoặc tương tự với thông tin văn bản của giao diện người sử dụng thứ nhất 630 trên màn hiển thị 460. Theo cách khác, bộ xử lý 410 có thể hiển thị thông tin văn bản 641 tương ứng với đầu vào thoại thứ nhất 633 và đầu vào thoại thứ hai 639 trên giao diện người sử dụng thứ hai 635 cùng lúc với việc nhận đầu vào thoại thứ hai 639.

Bộ xử lý 410 có thể hiển thị giao diện người sử dụng thứ ba 600 bao gồm thông tin văn bản 641 tương ứng với đầu vào thoại thứ nhất 633 và đầu vào thoại thứ hai 639 trên màn hiển thị 460. Khi không phát hiện được đầu vào thoại mới sau khi đầu vào thoại thứ hai 639, thời gian EPD đã hết, và từ cuối cùng (ví dụ, “speaker”) không tương ứng với từ định sẵn, bộ xử lý 410 có thể xử lý nhận dạng giọng nói tương ứng với thông tin văn bản 641. Ví dụ, bộ xử lý 410 có thể truyền thông tin văn bản 641 đến máy chủ xử lý giọng nói và nhận lệnh “Turn on the living room light and play the music through the kitchen speaker.” từ máy chủ xử lý giọng nói, và có thể điều khiển để bật đèn của phòng khách và phát nhạc thông qua loa của phòng bếp. Khi lệnh thoại được hoàn thành theo quy trình xử lý nhận dạng giọng nói, bộ xử lý 410 có thể hiển thị giao diện người sử dụng thứ tư 645 bao gồm tin nhắn hoàn thành lệnh 647 trên màn hiển thị 460. Khi quy trình xử lý nhận dạng giọng nói được hoàn thành, bộ xử lý 410 có thể kết xuất thông tin thoại tương ứng với tin nhắn hoàn thành lệnh 647 thông qua loa 450.

Fig.6C là hình vẽ minh họa một ví dụ về việc kéo dài thời gian EPD theo việc phát hiện từ rỗng và lệnh chờ.

Dựa vào Fig.6C, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể nhận lệnh thoại về việc gửi tin nhắn từ người sử dụng. Ví dụ, bộ xử lý 410 có thể nhận đầu vào thoại thứ nhất 653 như “Send message to wife honey I’ll be...uh...well...”. Bộ xử lý 410 có thể hiển thị giao diện người sử dụng thứ nhất 650 bao gồm thông tin văn bản 651 tương ứng với đầu vào thoại thứ nhất 653 trên màn hiển thị 460. Theo các phương án khác nhau, bộ xử lý 410 có thể cung cấp một phần hoặc toàn bộ đầu vào thoại thứ nhất 653 dưới dạng thông tin văn bản 651.

Khi phát hiện được đầu vào thoại thứ nhất 653 hoặc kết thúc đầu vào thoại thứ

nhất 653, bộ xử lý 410 có thể tính thời gian EPD và có thể xác định xem có phát hiện được đầu vào thoại mới trước khi hết thời gian EPD hay không. Khi không phát hiện được đầu vào thoại mới trước khi hết thời gian EPD và thời gian EDP đã hết, bộ xử lý 410 có thể xác định xem từ cuối cùng có phải là từ định sẵn hay không. Ví dụ, khi từ cuối cùng (ví dụ, “uh”, “well”) tương ứng với từ định sẵn (ví dụ, từ rỗng) dựa trên thông tin được lưu trữ trong bộ nhớ 470, bộ xử lý 410 có thể kéo dài thời gian EPD và có thể chờ đầu vào thoại từ người sử dụng. Bộ xử lý 410 có thể nhận đầu vào thoại thứ hai 659 từ người sử dụng trong thời gian chờ đầu vào thoại (ví dụ, thời gian EPD). Ví dụ, bộ xử lý 410 có thể hiển thị giao diện người sử dụng thứ hai 655 bao gồm thông tin văn bản 657 giống hoặc tương tự với thông tin văn bản của giao diện người sử dụng thứ nhất 650 trên màn hiển thị 460. Theo cách khác, bộ xử lý 410 có thể hiển thị thông tin văn bản (ví dụ, “Honey I’ll be wait”) tương ứng với đầu vào thoại thứ nhất 653 và đầu vào thoại thứ hai 659 trên giao diện người sử dụng thứ hai 655 cùng lúc với việc nhận đầu vào thoại thứ hai 659.

Khi không phát hiện được đầu vào thoại mới sau khi đầu vào thoại thứ hai 659 và thời gian EPD đã hết, bộ xử lý 410 có thể xác định xem từ cuối cùng (ví dụ, “wait”) có tương ứng với từ định sẵn (ví dụ, lệnh chờ) hay không. Theo cách khác, bộ xử lý 410 có thể xác định xem đầu vào thoại thứ hai 659 nhận được trước khi hết thời gian EPD có tương ứng với từ định sẵn hay không. Ví dụ, khi từ cuối cùng (ví dụ, “wait”) tương ứng với từ định sẵn (ví dụ, lệnh chờ) dựa trên thông tin được lưu trữ trong bộ nhớ 470, thì bộ xử lý 410 có thể kéo dài thời gian EPD và có thể chờ đầu vào thoại từ người sử dụng. Bộ xử lý 410 có thể nhận đầu vào thoại thứ ba 663 từ người sử dụng trong thời gian chờ đầu vào thoại (ví dụ, thời gian EPD). Ví dụ, bộ xử lý 410 có thể hiển thị thông tin văn bản 661 tương ứng với đầu vào thoại thứ nhất 651 và đầu vào thoại thứ ba 663 trên giao diện người sử dụng thứ ba 660 cùng lúc với việc nhận đầu vào thoại thứ ba 663.

Bộ xử lý 410 có thể hiển thị giao diện người sử dụng thứ tư 665 bao gồm thông tin văn bản 667 tương ứng với đầu vào thoại thứ nhất 651 và đầu vào thoại thứ ba 663 trên màn hiển thị 460. Khi không phát hiện được đầu vào thoại mới sau khi đầu vào thoại thứ ba 663, thời gian EPD đã hết, và từ cuối cùng (ví dụ, “today”) không tương ứng với từ định sẵn, thì bộ xử lý 410 có thể xử lý nhận dạng giọng nói tương ứng với thông tin văn bản 667. Khi việc gửi tin nhắn được hoàn thành theo quy trình

xử lý nhận dạng giọng nói mặc dù không được minh họa, bộ xử lý 410 có thể hiển thị giao diện người sử dụng thứ năm (ví dụ, giao diện người sử dụng thứ tư 625) bao gồm tin nhắn hoàn thành việc gửi trên màn hiển thị 460. Khi quy trình xử lý nhận dạng giọng nói được hoàn thành, bộ xử lý 410 có thể kết xuất thông tin thoại tương ứng với tin nhắn hoàn thành việc gửi 667 thông qua loa 450.

Fig.7 là lưu đồ minh họa phương pháp thu thập từ cá nhân trong thiết bị điện tử theo các phương án khác nhau.

Dựa vào Fig.7, ở bước 701, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể nhận dạng lỗi lệnh thoại. Lỗi lệnh thoại có thể bao gồm trường hợp trong đó lệnh thoại bị dừng lại (hoặc bị hủy) bởi người sử dụng hoặc trường hợp trong đó bộ xử lý 410 không xử lý nhận dạng giọng nói. Khi lệnh thoại bị lỗi, người sử dụng có thể nhập lại lệnh thoại (ví dụ, người sử dụng có thể nói hoặc nhập giọng nói).

Ở bước 703, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể xử lý nhận dạng giọng nói. Ví dụ, bộ xử lý 410 (ví dụ, bộ nhận dạng giọng nói 411) có thể nhận đầu vào thoại từ người sử dụng theo lỗi lệnh thoại và có thể xử lý đầu vào thoại.

Ở bước 705, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể xác định xem việc nhận dạng giọng nói có thành công hay không. Ví dụ, bộ xử lý 400 (ví dụ, bộ nhận dạng giọng nói 411) có thể xác định xem chức năng tương ứng với đầu vào thoại đã phát hiện được ở bước 703 có được thực hiện đúng hay không.

Khi việc nhận dạng giọng nói thành công, thì bộ xử lý 410 có thể thực hiện bước 707, và khi việc nhận dạng giọng nói không thành công, thì bộ xử lý 410 có thể quay trở lại bước 701.

Ở bước 707, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể cung cấp dịch vụ (hoặc chức năng) theo giọng nói đã nhận dạng được. Bước 707 có thể được bao gồm trong bước 703 và được thực hiện. Trong trường hợp này, bước 707 có thể được lược bỏ. Tức là, khi chức năng theo dữ liệu nhận dạng giọng nói được cung cấp ở bước 703, thì bước 707 có thể được lược bỏ.

Ở bước 709, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể xác định xem dữ liệu nhận dạng giọng nói này có tương tự với dữ liệu nhận dạng giọng nói trước đó hay

không. Ví dụ, người sử dụng có thể nhập lệnh thoại giống hoặc tương tự với lệnh thoại trước đó do lỗi nhận dạng giọng nói ở bước 701. Bộ xử lý 410 có thể thu thập từ cá nhân liên quan đến lệnh thoại được nhập vào trong thời gian định sẵn sau khi lỗi lệnh thoại bằng cách so sánh lệnh thoại trước đó (ví dụ, dữ liệu nhận dạng giọng nói trước đó) và lệnh thoại hiện tại (ví dụ, dữ liệu nhận dạng giọng nói hiện tại). Ví dụ, bộ xử lý 410 có thể xác định mức độ tương tự giữa dữ liệu nhận dạng giọng nói (ví dụ, dữ liệu nhận dạng giọng nói trước đó) ở bước 701 và dữ liệu nhận dạng giọng nói được nhận dạng ở bước 707. Khi mức độ tương tự giữa dữ liệu nhận dạng giọng nói trước đó và dữ liệu nhận dạng giọng nói hiện tại lớn hơn hoặc bằng tỷ lệ định sẵn (ví dụ, 70%), thì bộ xử lý 410 có thể xác định rằng dữ liệu nhận dạng giọng nói trước đó và dữ liệu nhận dạng giọng nói hiện tại là tương tự với nhau.

Khi dữ liệu nhận dạng giọng nói hiện tại tương tự với dữ liệu nhận dạng giọng nói trước đó, thì bộ xử lý 41 có thể thực hiện bước 711, và khi dữ liệu nhận dạng giọng nói hiện tại không tương tự với dữ liệu nhận dạng giọng nói trước đó, thì bộ xử lý 410 có thể kết thúc việc xử lý.

Ở bước 711, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể phân tích thông tin văn bản. Ví dụ, bộ xử lý 410 có thể phân tích văn bản của phần khác nhau giữa dữ liệu nhận dạng giọng nói trước đó và dữ liệu nhận dạng giọng nói hiện tại. Việc phân tích văn bản có thể được thực hiện dựa trên thông tin văn bản thu được bằng cách thực hiện chuyển đổi giọng nói thành văn bản.

Ở bước 713, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể xác định xem thông tin văn bản đã phân tích có được phát hiện với số lần đã chọn hoặc nhiều hơn hay không. Bộ xử lý 410 có thể lưu trữ tạm thời thông tin văn bản đã phân tích trong bộ nhớ 470 hoặc có thể lưu trữ số lần phát hiện được thông tin văn bản này trong bộ nhớ 470. Bộ xử lý 410 có thể xác định xem số lần phát hiện được thông tin văn bản đã phân tích có lớn hơn hoặc bằng số lần đã chọn (ví dụ, 5 lần hoặc 10 lần) hay không.

Khi phát hiện được thông tin văn bản đã phân tích với số lần đã chọn hoặc nhiều hơn, thì bộ xử lý 410 có thể thực hiện bước 715, và khi không phát hiện được thông tin văn bản đã phân tích với số lần đã chọn hoặc nhiều hơn, thì bộ xử lý 410 có thể kết thúc việc xử lý. Ví dụ, khi không phát hiện được thông tin văn bản đã phân tích với số lần đã chọn hoặc nhiều hơn, thì bộ xử lý 410 có thể tăng số lần phát hiện được

thông tin văn bản đã phân tích trong bộ nhớ 470 thêm 1, và có thể kết thúc việc xử lý.

Ở bước 715, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể lưu trữ thông tin văn bản trong cơ sở dữ liệu (Database, DB) cá nhân. Ví dụ, khi phát hiện được thông tin văn bản đã phân tích với số lần đã chọn hoặc nhiều hơn, thì bộ xử lý 410 có thể lưu trữ thông tin văn bản đã phát hiện được trong bộ nhớ 470 dưới dạng từ cá nhân. Trong trường hợp này, bộ xử lý 410 có thể tham chiếu đến các từ cá nhân được lưu trữ trong bộ nhớ 470 khi xác định xem từ cuối cùng có tương ứng với từ định sẵn hay không, để có thể phân tích chính xác hơn ý định kết thúc giọng nói của người sử dụng. Các từ cá nhân được thu thập theo cách này có thể được thu thập và quản lý theo từng người sử dụng như được thể hiện trong bảng 2.

Fig.8A đến Fig.8C là các hình vẽ minh họa các ví dụ về việc thu thập từ cá nhân trong thiết bị điện tử theo các phương án khác nhau.

Fig.8A là hình vẽ minh họa một ví dụ về việc thu thập từ cá nhân khi lệnh thoại bị hủy theo yêu cầu của người sử dụng.

Dựa vào Fig.8A, thiết bị điện tử 820 có thể nhận đầu vào thoại thứ nhất 811 từ người sử dụng 810, và có thể nhận dạng lệnh thoại thứ nhất 821. Tuy nhiên, mặc dù người sử dụng 810 không hoàn thành đầu vào thoại, nhưng thiết bị điện tử 820 có thể xác định rằng đầu vào thoại được hoàn thành và có thể nhận dạng lệnh thoại thứ nhất 821 dựa trên giọng nói được nhập vào. Trong trường hợp này, người sử dụng 810 có thể nói đầu vào yêu cầu hủy thoại 813. Thiết bị điện tử 820 có thể nhận đầu vào yêu cầu hủy thoại 813 và có thể xử lý lệnh hủy 823. Người sử dụng có thể nói lại đầu vào thoại thứ hai 815 tương tự với đầu vào thoại thứ nhất 811 sau khi nói đầu vào yêu cầu hủy thoại 813.

Khi nhận được đầu vào thoại thứ hai 815 giống hoặc tương tự với đầu vào thoại thứ nhất 811 sau khi lệnh hủy 823 được thực hiện, thiết bị điện tử 820 có thể xử lý nhận dạng giọng nói đối với đầu vào thoại thứ hai 815. Ví dụ, thiết bị điện tử 820 có thể gửi tin nhắn theo đầu vào thoại thứ hai 815. Sau khi gửi tin nhắn, thiết bị điện tử 820 có thể xác định mức độ tương tự giữa đầu vào thoại thứ nhất 811 và đầu vào thoại thứ hai 815, và khi mức độ tương tự là cao, thiết bị điện tử 820 có thể phân tích thông tin văn bản thay đổi (ví dụ, “you know”) giữa đầu vào thoại thứ nhất 811 và đầu vào thoại thứ hai 815, và có thể lưu trữ thông tin văn bản này trong DB cá nhân 830. Sau

đó, thiết bị điện tử 820 có thể kéo dài thời gian EPD khi từ cuối cùng tương ứng với từ cá nhân dựa trên các từ cá nhân được lưu trữ trong DB cá nhân 830.

Thiết bị điện tử 820 theo các phương án khác nhau có thể thiết lập thời gian EPD khác nhau theo đặc trưng của người sử dụng (ví dụ, tốc độ nói, thói quen nói). Ví dụ, khi phát hiện được từ cụ thể (ví dụ, từ rỗng, liên từ, lệnh chờ) dưới dạng kết quả của việc phân tích lịch sử nói của người sử dụng (ví dụ, thói quen, tốc độ) trong giai đoạn định sẵn, thiết bị điện tử 820 có thể nhận dạng tổng thời gian EPD trung bình được kéo dài dưới dạng N+M giây. Trong trường hợp này, khi phát hiện được đầu vào là từ cụ thể dựa trên đặc trưng của người sử dụng, thiết bị điện tử 820 có thể kéo dài thời gian EPD đến N+M giây.

Fig.8B là hình vẽ minh họa một ví dụ về việc thu thập từ cá nhân khi lệnh thoại bị lỗi.

Dựa vào Fig.8B, thiết bị điện tử 820 có thể nhận đầu vào thoại thứ nhất 817 từ người sử dụng 810 và có thể kết xuất phản hồi 825 liên quan đến lỗi nhận dạng giọng nói. Ví dụ, khi thiết bị điện tử 820 không tìm thấy chức năng cần thực hiện để phản hồi lại đầu vào thoại thứ nhất 817 nhận được hoặc không thể chuyển đổi thông tin thoại được nhập vào thành thông tin văn bản, thì thiết bị điện tử 820 có thể kết xuất phản hồi lỗi nhận dạng 825. Ví dụ, thiết bị điện tử 820 có thể không tìm thấy chức năng tương ứng với đầu vào thoại thứ nhất 817 do từ rỗng (ví dụ, “I mean”) được bao gồm trong đầu vào thoại thứ nhất 817. Trong trường hợp này, người sử dụng 810 có thể nói lại đầu vào thoại thứ hai 819 tương tự với đầu vào thoại thứ nhất 817 theo cách chính xác hơn. Ví dụ, người sử dụng 810 có thể nói đầu vào thoại không bao gồm từ rỗng từ đầu vào thoại thứ nhất 817 làm đầu vào thoại thứ hai 819 để nhập lệnh thoại chính xác hơn.

Khi nhận được đầu vào thoại thứ hai 819 sau khi phản hồi lỗi nhận dạng 825 được kết xuất, thì thiết bị điện tử 820 có thể xử lý nhận dạng giọng nói đối với đầu vào thoại thứ hai 819. Ví dụ, thiết bị điện tử 820 có thể gửi tin nhắn (hoặc thư điện tử) theo đầu vào thoại thứ hai 819 (ví dụ, gửi tệp tin tới người nhận, Jack). Sau khi việc gửi tin nhắn (hoặc thư điện tử) được hoàn thành, thì thiết bị điện tử 820 có thể xác định mức độ tương tự giữa đầu vào thoại thứ nhất 817 và đầu vào thoại thứ hai 819, và khi mức độ tương tự là cao, thiết bị điện tử 820 có thể phân tích thông tin văn bản thay đổi (ví

dụ, “I mean”) giữa đầu vào thoại thứ nhất 817 và đầu vào thoại thứ hai 819, và có thể lưu trữ thông tin văn bản này trong DB cá nhân 830. Thiết bị điện tử 820 có thể kéo dài thời gian EPD khi từ cuối cùng tương ứng với từ cá nhân dựa trên các từ cá nhân được lưu trữ trong DB cá nhân 830.

Fig.8C là hình vẽ minh họa một ví dụ về việc phân tích từ cá nhân trong thiết bị điện tử theo các phương án khác nhau.

Dựa vào Fig.8C, thiết bị điện tử 820 có thể lưu trữ từ cá nhân trong DB cá nhân 830 bằng cách thực hiện bước so sánh hai lệnh thoại 850, bước thu thập lệnh thoại bị lỗi 860, bước trích xuất từ 870, và bước tính số lần phát hiện được từ 890. Ví dụ, bước so sánh 850 có thể là bước so sánh lệnh thoại thứ nhất (ví dụ, lệnh thoại trước đó) và lệnh thoại thứ hai (ví dụ, lệnh thoại được nhận dạng lại). Thiết bị điện tử 820 có thể phân tích thông tin văn bản giữa hai lệnh thoại này ở bước so sánh 850. Bước thu thập 860 có thể là bước phân tích thông tin văn bản liên quan đến lệnh thoại bị lỗi. Thiết bị điện tử 820 có thể thực hiện bước trích xuất từ vô nghĩa lặp lại giữa các từ có trong lệnh thoại (bước 870) thông qua bước so sánh 850 và bước thu thập 860. Bước tính 890 có thể là bước xác định xem số lần mà từ lặp lại được phát hiện có lớn hơn hoặc bằng số lần định sẵn hay không.

Khi nhận được đầu vào thoại thứ nhất 841 từ người sử dụng 810 sau khi DB cá nhân 830 được tạo ra, thiết bị điện tử 820 có thể xác định xem đầu vào thoại thứ nhất 841 có bao gồm từ cá nhân được lưu trữ trong DB cá nhân 830 hay không. Khi đầu vào thoại thứ nhất 841 bao gồm từ cá nhân, thiết bị điện tử 820 có thể kéo dài thời gian EPD và có thể chờ nhận đầu vào thoại. Thiết bị điện tử 820 có thể nhận đầu vào thoại thứ hai 843 trong khi chờ nhận đầu vào thoại. Thiết bị điện tử 820 có thể xử lý chức năng nhận dạng giọng nói 845 dựa trên thông tin văn bản bao gồm đầu vào thoại thứ nhất 841 và đầu vào thoại thứ hai 843.

Thiết bị điện tử 820 theo các phương án khác nhau tạo ra DB cá nhân 830 dựa trên đặc điểm cá nhân như ngữ điệu, thói quen, cách nói, v.v., của người sử dụng, để có thể xác định chính xác hơn ý định kết thúc giọng nói của người sử dụng.

Fig.9 là lưu đồ minh họa phương pháp cảm biến sự kết thúc giọng nói trong thiết bị điện tử theo các phương án khác nhau.

Dựa vào Fig.9, ở bước 901, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể nhận đầu vào thoại. Bộ xử lý 410 (ví dụ, bộ phận nhập giọng nói 411) có thể phát hiện xem liệu âm thanh có được nhập vào từ micrô 420 hay không. Bước 901 tương tự hoặc giống với bước 501 trên Fig.5, và do đó, mô tả chi tiết về bước này được lược bỏ.

Ở bước 903, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể tính thời gian EPD. Ví dụ, thời gian EPD có thể là thời gian chờ để xác định rằng đầu vào thoại đã kết thúc, và có thể chỉ báo thời gian chờ đầu vào hoặc thời gian chờ ghi âm micrô. Bước 903 tương tự hoặc giống với bước 503 trên Fig.5, và do đó, mô tả chi tiết về bước này được lược bỏ.

Ở bước 905, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể xác định xem thời gian EPD đã hết chưa. Bộ xử lý 410 (ví dụ, bộ kiểm soát thời gian EPD 413) có thể tính thời gian EPD và sau đó có thể tiếp tục xác định xem thời gian EPD đã hết chưa.

Khi chưa hết thời gian EPD, bộ xử lý 410 có thể thực hiện bước 915, và khi hết thời gian EPD, bộ xử lý 410 có thể thực hiện bước 907.

Khi chưa hết thời gian EPD, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể xác định xem có phát hiện được đầu vào thoại ở bước 915 hay không. Bộ xử lý 410 (ví dụ, bộ nhận dạng giọng nói 411) có thể xác định xem có phát hiện được đầu vào thoại mới trước khi hết thời gian EPD hay không.

Khi phát hiện được đầu vào thoại, bộ xử lý 410 có thể thực hiện bước 917, và khi không phát hiện được đầu vào thoại, bộ xử lý 410 có thể quay trở lại bước 905.

Khi phát hiện được đầu vào thoại, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể kéo dài thời gian EPD ở bước 917. Bộ xử lý 410 (ví dụ, bộ kiểm soát thời gian EPD 413) có thể tính lại thời gian EPD khi phát hiện được đầu vào thoại trong thời gian EPD. Bộ xử lý 410 có thể tính thời gian EPD từ thời điểm kết thúc đầu vào thoại đã phát hiện được.

Khi hết thời gian EPD, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể xác định xem từ cuối cùng có phải là từ được chọn hay không ở bước 907. Từ cuối cùng có thể là toàn bộ hoặc một phần của thông tin văn bản tương ứng với thông tin thoại được nhập vào cuối cùng bởi người sử dụng. Ví dụ, khi không phát hiện được đầu vào

thoại sau khi từ cuối cùng và thời gian EPD đã hết, bộ xử lý 410 (ví dụ, bộ xác định ngôn ngữ 415) có thể xác định xem từ cuối cùng có phải là từ định sẵn hay không.

Khi từ cuối cùng là từ đã chọn, bộ xử lý 410 có thể thực hiện bước 917, và khi từ cuối cùng không phải là từ đã chọn, bộ xử lý 410 có thể thực hiện bước 909.

Ở bước 909, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể xác định xem có phát hiện được cử chỉ định sẵn hay không. Bộ xử lý 410 (ví dụ, bộ xác định cử chỉ 417) có thể xác định xem cử chỉ của người sử dụng đã phát hiện được từ camera 430 hoặc bộ cảm biến (ví dụ, bộ cảm biến chuyển động) có phải là cử chỉ định sẵn hay không. Người sử dụng có thể tạo ra cử chỉ cụ thể trong khi nghĩ về từ mà người sử dụng muốn nói vì người sử dụng không nhớ ra từ chính xác trong khi nhập giọng nói. Ví dụ, người sử dụng có thể nhướn lông mày hoặc chuyển động tay trong khi nói. Ngoài ra, người sử dụng có thể nhìn lên và nhìn sang phải hoặc có thể nghiêng đầu trong khi nghĩ. Theo cách khác, khi người sử dụng kết thúc (hoặc dừng) nói, người sử dụng có thể không chuyển động tay. Bộ xử lý 410 có thể xác định xem có phát hiện được cử chỉ định sẵn bao gồm cử chỉ thông dụng hoặc cử chỉ cá nhân hay không. Ví dụ, bộ xử lý 410 có thể xác định xem cử chỉ được phát hiện sau khi hết thời gian EPD hoặc trước khi hết thời gian EPD có phải là cử chỉ định sẵn hay không.

Mặc dù Fig.9 mô tả rằng bước 907 được thực hiện trước tiên và sau đó bước 909 được thực hiện, nhưng bước 909 có thể được thực hiện trước tiên và sau đó bước 907 có thể được thực hiện, hoặc bước 907 và bước 909 có thể được thực hiện đồng thời. Đây là một thay đổi thiết kế đơn giản và thứ tự của các bước không bị giới hạn bởi các hình vẽ.

Khi phát hiện được cử chỉ định sẵn, bộ xử lý 410 có thể thực hiện bước 917, và khi không phát hiện được cử chỉ định sẵn, bộ xử lý 410 có thể thực hiện bước 911.

Ở bước 911, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể xác định rằng đầu vào thoại đã được hoàn thành. Ví dụ, khi hết thời gian EPD, từ cuối cùng không phải là từ đã chọn, và cử chỉ phát hiện được không phải là cử chỉ định sẵn, bộ xử lý 410 (ví dụ, bộ nhận dạng giọng nói 411) có thể xác định rằng người sử dụng có ý định kết thúc đầu vào thoại. Bước 911 tương tự hoặc giống với bước 509 trên Fig.5, và do đó, mô tả chi tiết về bước này được lược bỏ.

Ở bước 913, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể xử lý nhận dạng giọng nói. Bộ xử lý 410 (ví dụ, bộ nhận dạng giọng nói 411) có thể cung cấp dịch vụ theo đầu vào thoại từ người sử dụng. Bước 913 tương tự hoặc giống với bước 511 trên Fig.5, và do đó, mô tả chi tiết về bước này được lược bỏ.

Theo các phương án khác nhau, bộ xử lý 410 có thể xác định thời gian kéo dài EPD khi phát hiện được đầu vào thoại, thời gian kéo dài EPD khi phát hiện được từ định sẵn, hoặc thời gian kéo dài EPD khi phát hiện được cử chỉ định sẵn là bằng hoặc khác nhau. Bộ xử lý 410 có thể thiết lập thời gian kéo dài EPD bằng hoặc khác nhau dựa trên thiết lập của người sử dụng hoặc thiết lập của thiết bị điện tử 400.

Fig.10 là lưu đồ khác minh họa phương pháp cảm biến sự kết thúc giọng nói trong thiết bị điện tử theo các phương án khác nhau.

Dựa vào Fig.10, ở bước 1001, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể nhận đầu vào thoại. Bộ xử lý 410 (hoặc bộ nhận dạng giọng nói 411) có thể phát hiện xem liệu âm thanh có được nhập vào từ micrô 420 hay không. Bước 1001 tương tự hoặc giống với bước 501 trên Fig.5, và do đó, mô tả chi tiết về bước này được lược bỏ.

Ở bước 1003, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể tính thời gian EPD. Ví dụ, thời gian EPD là thời gian chờ để xác định rằng đầu vào thoại đã kết thúc, và có thể chỉ báo thời gian chờ đầu vào hoặc thời gian chờ ghi âm micrô. Bước 1003 tương tự hoặc giống với bước 503 trên Fig.5, và do đó, mô tả chi tiết về bước này được lược bỏ.

Ở bước 1005, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể xác định xem thời gian EPD đã hết chưa. Bộ xử lý 410 (ví dụ, bộ kiểm soát thời gian EPD 413) có thể tính thời gian EPD và sau đó có thể tiếp tục xác định xem thời gian EPD đã hết chưa.

Khi chưa hết thời gian EPD, bộ xử lý 410 có thể thực hiện bước 1013, và khi hết thời gian EPD, bộ xử lý 410 có thể thực hiện bước 1007.

Khi chưa hết thời gian EPD, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể xác định xem có phát hiện được đầu vào thoại ở bước 1013 hay không. Bộ xử lý 410 (ví dụ, bộ nhận dạng giọng nói 411) có thể xác định xem có phát hiện được đầu vào

thoại mới trước khi hết thời gian EPD hay không.

Khi phát hiện được đầu vào thoại, bộ xử lý 410 có thể thực hiện bước 1015, và khi không phát hiện được đầu vào thoại, bộ xử lý 410 có thể quay trở lại bước 1005.

Khi phát hiện được đầu vào thoại, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể kéo dài thời gian EPD ở bước 1005. Khi phát hiện được đầu vào thoại trong thời gian EPD, bộ xử lý 410 (ví dụ, bộ kiểm soát thời gian EPD 413) có thể tính lại thời gian EPD.

Khi hết thời gian EPD, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể xác định xem câu theo đầu vào thoại có được hoàn thành ở bước 1007 hay không. Trong trường hợp câu hoàn thành về mặt ngữ pháp, thì khả năng cao là người sử dụng hoàn thành giọng nói, nhưng trong trường hợp câu không hoàn thành, thì khả năng cao là người sử dụng sẽ tiếp tục nói. Bộ xử lý 410 có thể xác định xem có kéo dài thời gian EPD dựa trên việc câu có được hoàn thành hay không. Để đạt được điều này, bộ xử lý 410 có thể phân tích thông tin văn bản tương ứng với thông tin thoại được nhập vào ở bước 1001 hoặc 1013, và có thể xác định xem câu có được hoàn thành hay không dựa trên thông tin văn bản đã phân tích. Theo các phương án khác nhau, bộ xử lý 410 có thể xác định xem câu không được hoàn thành hay không và có cần đến từ thiết yếu để hoàn thành câu hay không. Khi không cần đến từ thiết yếu, bộ xử lý 410 có thể xác định rằng đầu vào thoại được hoàn thành mặc dù câu không được hoàn thành. Theo cách khác, khi cần đến từ thiết yếu, bộ xử lý 410 có thể kéo dài thời gian EPD để nhận đầu vào thoại bổ sung.

Theo các phương án khác nhau, bộ xử lý 410 có thể thu thập thông tin ngữ cảnh của thiết bị điện tử 400 và có thể xác định xem có thực hiện bước 1007 để xác định xem câu có được hoàn thành hay không, dựa trên thông tin ngữ cảnh này. Ví dụ, câu có thể không được hoàn thành, nhưng có thể là không có vấn đề gì trong việc thực hiện chức năng tương ứng với lệnh thoại. Theo đó, bộ xử lý 410 có thể xác định rằng đầu vào thoại được hoàn thành ngay cả khi câu không được hoàn thành dựa trên loại lệnh thoại.

Khi câu được hoàn thành, bộ xử lý 410 có thể thực hiện bước 1009, và khi câu không được hoàn thành, bộ xử lý 410 có thể quay trở lại bước 1015.

Ở bước 1009, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể xác định rằng đầu vào thoại đã được hoàn thành. Ví dụ, khi hết thời gian EPD, từ cuối cùng không phải là từ đã chọn, và cử chỉ phát hiện được không phải là cử chỉ định sẵn, thì bộ xử lý 410 (ví dụ, bộ nhận dạng giọng nói 411) có thể xác định rằng người sử dụng định kết thúc đầu vào thoại. Bước 1009 tương tự hoặc giống với bước 509 trên Fig.5, và do đó, mô tả chi tiết về bước này được lược bỏ.

Ở bước 1011, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể xử lý nhận dạng giọng nói. Bộ xử lý 410 (ví dụ, bộ nhận dạng giọng nói 411) có thể cung cấp dịch vụ theo đầu vào thoại từ người sử dụng. Bước 1011 tương tự hoặc giống với bước 511 trên Fig.5, và mô tả chi tiết về bước này được lược bỏ.

Theo các phương án khác nhau, bộ xử lý 410 có thể xác định thời gian kéo dài EPD khi phát hiện được đầu vào thoại hoặc thời gian kéo dài EPD khi câu không được hoàn thành là bằng hoặc khác nhau. Bộ xử lý 410 có thể thiết lập thời gian kéo dài EPD bằng hoặc khác nhau dựa trên thiết lập của người sử dụng hoặc thiết lập của thiết bị điện tử 400.

Fig.11 là hình vẽ minh họa một ví dụ về việc xác định xem câu có được hoàn thành hay không trong thiết bị điện tử theo các phương án khác nhau.

Dựa vào Fig.11, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể xác định xem câu có được hoàn thành hay không, dựa trên loại lệnh thoại (ví dụ, lệnh gửi ảnh). Bộ xử lý 410 có thể hiểu lệnh thoại 1110 liên quan đến việc gửi ảnh là một câu hoàn chỉnh khi tất cả trong số động từ 111 (“send”), đối tượng 113 (“this photo”) và người nhận 1115 (Jack) đều được nhập vào. Bộ xử lý 410 có thể xử lý thông tin thoại được nhập vào từ người sử dụng, và có thể nhận dạng lệnh thoại 1110. Ví dụ, bộ xử lý 410 có thể xóa bỏ từ vô nghĩa (ví dụ, giới từ) trong số các từ được bao gồm trong thông tin thoại (ví dụ, “send this photo to Jack”).

Khi nhận được lệnh thoại thứ nhất 1120 chỉ bao gồm động từ 1121 và đối tượng 1123, bộ xử lý 410 có thể xác định rằng câu này là câu không hoàn chỉnh. Trong trường hợp này, khi nhận được lệnh thoại thứ nhất 1120, bộ xử lý 410 có thể kéo dài thời gian EPD. Theo cách khác, khi lệnh thoại thứ hai 1130 chỉ bao gồm động từ 1131 và đối tượng 1133, bộ xử lý 410 có thể xác định rằng câu này là câu không hoàn chỉnh. Trong trường hợp này, khi phát hiện thêm các từ bổ sung (“to, well”) ngoài động từ

1131 và đối tượng 1133, nhưng không tìm thấy người nhận liên quan đến các từ bổ sung này, bộ xử lý 410 có thể xác định rằng câu này là câu không hoàn chỉnh. Khi nhận được lệnh thoại thứ hai 1130, bộ xử lý 410 có thể kéo dài thời gian EPD. Khi nhận được lệnh thoại thứ nhất 1120 hoặc lệnh thoại thứ hai 1130, bộ xử lý 410 có thể kéo dài thời gian EPD, và do đó có thể nhận lệnh thoại thứ ba 1140. Lệnh thoại thứ ba 1140 có thể được cấu tạo dưới dạng câu hoàn chỉnh bao gồm tất cả động từ 1141, đối tượng 1143 và người nhận 1145. Bộ xử lý 410 còn nhận đầu vào thoại tương ứng với người nhận 1145 trong thời gian kéo dài EPD, để có thể thu được lệnh thoại thứ ba 1140 hoàn chỉnh.

Fig.12 là lưu đồ minh họa phương pháp cảm biến sự kết thúc giọng nói bằng cách sử dụng ý định kết thúc của người sử dụng theo các phương án khác nhau.

Dựa vào Fig.12, ở bước 1201, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể nhận đầu vào thoại. Bộ xử lý 410 (ví dụ, bộ nhận dạng giọng nói 411) có thể phát hiện xem liệu âm thanh có được nhập vào từ micrô 420 hay không. Bước 1201 tương tự hoặc giống với bước 501 trên Fig.5, và do đó, mô tả chi tiết về bước này được lược bỏ.

Ở bước 1203, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể tính thời gian EPD. Ví dụ, thời gian EPD có thể là thời gian chờ để xác định rằng đầu vào thoại đã kết thúc, và có thể chỉ báo thời gian chờ đầu vào hoặc thời gian chờ ghi âm micrô. Bước 1203 tương tự hoặc giống với bước 503 trên Fig.5, và do đó, mô tả chi tiết về bước này được lược bỏ.

Ở bước 1205, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể phân tích ý định kết thúc giọng nói của người sử dụng (hoặc ý định tiếp tục nói). Ví dụ, bộ xử lý 410 có thể phân tích ý định kết thúc giọng nói của người sử dụng, dựa trên ít nhất một trong số thông tin ngữ cảnh của thiết bị điện tử 400, thông tin đặc trưng của người sử dụng, việc có phát hiện được đầu vào thoại bổ sung hay không, việc có phát hiện được từ định sẵn hay không, việc có phát hiện được cử chỉ định sẵn hay không, hoặc việc có hoàn thành câu hay không. Ví dụ, thiết bị điện tử 400 có thể xác định thời gian EPD dựa trên thông tin ngữ cảnh của thiết bị điện tử 400 hoặc thông tin đặc trưng của người sử dụng, và có thể xác định rằng người sử dụng có ý định tiếp tục nói (ví dụ, không có ý định kết thúc giọng nói) khi phát hiện được ít nhất một trong số đầu vào thoại bổ sung, từ định sẵn, cử chỉ định sẵn hoặc câu không hoàn chỉnh trong thời gian EPD.

Theo các phương án khác nhau, bộ xử lý 410 có thể gán điểm khác nhau cho ít nhất một trong số thời gian phát hiện sự im lặng, việc có phát hiện được từ định sẵn hay không, việc có phát hiện được cử chỉ định sẵn hay không, hoặc việc có hoàn thành câu hay không. Theo các phương án khác nhau, bộ xử lý 410 có thể gán giá trị trọng số khác nhau cho ít nhất một yếu tố trong số thời gian phát hiện sự im lặng, việc có phát hiện được từ định sẵn hay không, việc có phát hiện được cử chỉ định sẵn hay không, hoặc việc có hoàn thành câu hay không, và có thể tính toán tổng số điểm.

Bảng 3

	Âm thanh	Yếu tố lời nói				Yếu tố không phải lời nói	
Yếu tố	Thời gian im lặng	Từ rỗng	Liên từ	Lệnh chờ	Hoàn thành câu	Mắt	Cử chỉ bàn tay
	Trọng số A	Trọng số B	Trọng số C	Trọng số D	Trọng số E	Trọng số F	Trọng số G
Tín hiệu	7 giây	Phát hiện được	Không xác định	Không xác định	Hoàn thành	Lên trên bên phải	Chuyển động
Điểm	2 điểm	4 điểm	0 điểm	0 điểm	1 điểm	4 điểm	4 điểm

Tham chiếu đến bảng 3, bộ xử lý 410 có thể gán điểm cao hơn khi thời gian phát hiện sự im lặng trong đó không âm thanh nào được nhập tăng lên. Khi phát hiện được ngôn ngữ định sẵn (ví dụ, từ rỗng, liên từ, lệnh chờ), bộ xử lý 410 có thể gán điểm đã thiết lập (ví dụ, 4 điểm). Khi phát hiện được ngôn ngữ định sẵn dư thừa (ví dụ, phát hiện được từ rỗng và phát hiện được liên từ), bộ xử lý 410 có thể gán điểm một cách tương ứng. Khi câu được hoàn thành, bộ xử lý 410 có thể gán điểm thứ nhất (ví dụ, 1 điểm), và khi câu không được hoàn thành, bộ xử lý 410 có thể gán điểm thứ hai (ví dụ, 4 điểm). Khi phát hiện được cử chỉ định sẵn, bộ xử lý 410 có thể gán điểm đã thiết lập (ví dụ, 4 điểm). Khi phát hiện được gấp đôi các cử chỉ đã thiết lập khác nhau, bộ xử lý 410 có thể gán điểm một cách tương ứng. Bộ xử lý 410 có thể tính toán chỉ số ý định kết thúc giọng nói của người sử dụng bằng cách nhân điểm đã gán tương ứng với giá trị trọng số.

Ở bước 1207, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể xác định xem chỉ số ý định này có vượt quá giá trị ngưỡng (ví dụ, 3 điểm) hay không. Khi chỉ số ý định là cao, bộ xử lý 410 có thể xác định rằng không có ý định kết thúc (ví dụ, có ý định tiếp tục nói), và khi chỉ số ý định là thấp, bộ xử lý 410 có thể xác định rằng có ý

định kết thúc (ví dụ, có ý định kết thúc giọng nói). Theo cách khác, điều ngược lại có thể xảy ra.

Khi chỉ số ý định vượt quá giá trị ngưỡng, bộ xử lý 410 có thể thực hiện bước 1213, và khi chỉ số ý định không vượt quá giá trị ngưỡng, bộ xử lý 410 có thể thực hiện bước 1209.

Ở bước 1213, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể kéo dài thời gian EPD. Theo các phương án khác nhau, bộ xử lý 410 có thể thay đổi thời gian kéo dài EPD dựa trên chỉ số ý định. Ví dụ, khi chỉ số ý định là cao, thì bộ xử lý 410 có thể thiết lập thời gian kéo dài EPD dài hơn (ví dụ, 5 giây), và khi chỉ số ý định là thấp, thì bộ xử lý 410 có thể thiết lập thời gian kéo dài EPD ngắn hơn (ví dụ, 3 giây). Theo cách khác, bộ xử lý 410 có thể cố định thời gian kéo dài EPD bất kể chỉ số ý định.

Khi chỉ số ý định không vượt quá giá trị ngưỡng, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể xác định rằng đầu vào thoại được hoàn thành ở bước 1209. Ví dụ, khi hết thời gian EPD và chỉ số ý định là thấp (ví dụ, ý định tiếp tục nói là thấp), bộ xử lý 410 (ví dụ, bộ nhận dạng giọng nói 411) có thể xác định rằng người sử dụng có ý định kết thúc đầu vào thoại. Bước 1209 tương tự hoặc giống với bước 509 trên Fig.5, và do đó, mô tả chi tiết về bước này được lược bỏ.

Ở bước 1211, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể xử lý nhận dạng giọng nói. Bộ xử lý 400 (ví dụ, bộ nhận dạng giọng nói 411) có thể cung cấp dịch vụ theo đầu vào thoại từ người sử dụng. Bước 1211 tương tự hoặc giống với bước 511 trên Fig.5, và do đó, mô tả chi tiết về bước này được lược bỏ.

Fig.13 là lưu đồ minh họa phương pháp kiểm soát thời gian EPD theo thông tin ngữ cảnh theo các phương án khác nhau.

Dựa vào Fig.13, ở bước 1301, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể nhận yêu cầu về dịch vụ thoại. Yêu cầu về dịch vụ thoại có thể được tạo ra bởi người sử dụng. Người sử dụng có thể yêu cầu dịch vụ thoại khi ứng dụng cụ thể đang được thực hiện, màn hình chính đang được hiển thị, hoặc khi màn hình thị 460 đang được tắt. Ví dụ, khi phát hiện được lệnh gọi thoại, nút cụ thể được chọn, chức năng cụ thể (ví dụ, chức năng nhận dạng giọng nói) được chọn bởi người sử dụng, hoặc khi phát hiện được cử chỉ/khuôn mặt/mắt cụ thể, bộ xử lý 410 có thể xác định rằng dịch vụ

thoại được yêu cầu.

Ở bước 1303, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể thu thập (hoặc thu nhận) thông tin ngữ cảnh. Thông tin ngữ cảnh có thể chỉ thông tin trạng thái của thiết bị điện tử 400. Bộ xử lý 410 có thể phân tích, thu thập hoặc thu nhận thông tin ngữ cảnh của thiết bị điện tử 400 tại thời điểm yêu cầu dịch vụ thoại. Ví dụ, bộ xử lý 410 có thể nhận dạng ít nhất một trong số ứng dụng đang được thực hiện, trạng thái hiển thị của màn hình 460 và trạng thái mạng của bộ truyền thông 440 tại thời điểm yêu cầu dịch vụ thoại.

Ở bước 1305, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể xác định (hoặc thiết lập) thời gian EPD theo thông tin ngữ cảnh. Thời gian EPD là để xác định sự kết thúc giọng nói của người sử dụng, và có thể là ngắn hoặc dài tùy theo thông tin ngữ cảnh.

Bảng 4 thể hiện thời gian EPD theo từng phân loại và từng yếu tố.

Bảng 4

Phân loại	Yếu tố	Thời gian EPD
Phương pháp tương tác	Đánh thức bằng cảm ứng chạm (cơ bản)	Trung bình
	Đánh thức bằng giọng nói	Trung bình
	Ấn nút để nói	Không có EPD cho đến khi nhả
Miền chức năng	Tin nhắn, lịch trình	Dài
	Gọi điện, bật đèn	Ngắn
	Danh sách mua sắm, điều khiển âm nhạc	Không có EPD cho đến khi người sử dụng nói “that’s it” hoặc “over”
Bước đối thoại	Bước bắt đầu đối thoại (ví dụ, “Send a message!”)	Ngắn
	Bước hỏi tiếp theo (ví dụ, “Who should I send the message?”)	Trung bình
	Bước ra lệnh (ví dụ, “Enter the text.”)	Dài
Người sử dụng cá nhân	Người sử dụng A	Dài
	Người sử dụng B	Trung bình
	Người sử dụng C	Ngắn

Dựa vào Fig.4, bộ xử lý 410 có thể thiết lập hoặc có thể không thiết lập thời

gian EPD theo phương pháp tương tác mà yêu cầu dịch vụ thoại. Theo cách khác, bộ xử lý 410 có thể thiết lập thời gian EPD cho lệnh thoại (hoặc ứng dụng, chức năng) khá phức tạp, như tin nhắn, lịch trình (hoặc lịch), v.v., là ngắn, có thể thiết lập thời gian EPD cho lệnh thoại đơn giản như gọi điện, bật đèn, v.v., là ngắn, và đối với lệnh thoại đòi hỏi phải nhận đầu vào liên tục như danh sách mua sắm, điều khiển âm nhạc, v.v., có thể không thiết lập thời gian EPD hoặc có thể thiết lập thời gian EPD dài hơn so với cho lệnh thoại phức tạp. Ngoài ra, bộ xử lý 410 có thể thiết lập thời gian EPD khác nhau theo giai đoạn đối thoại hoặc đặc điểm cá nhân. Ví dụ, vì tốc độ nói thay đổi theo người sử dụng, nên bộ xử lý 410 có thể điều chỉnh thời gian EPD ngắn hoặc dài bằng cách xem xét tốc độ nói của người sử dụng.

Ở bước 1307, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể nhận đầu vào thoại. Bộ xử lý 410 (ví dụ, bộ nhận dạng giọng nói 411) có thể phát hiện xem liệu âm thanh có được nhập vào từ micrô 420 hay không. Bước 1307 tương tự hoặc giống với bước 501 trên Fig.5, và do đó, mô tả chi tiết về bước này được lược bỏ.

Ở bước 1309, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể xác định xem thời gian EPD đã hết chưa. Khi nhận được đầu vào thoại, bộ xử lý 410 có thể tính thời gian EPD đã xác định được ở bước 1305, và có thể xác định xem thời gian EPD đã tính được đã hết chưa.

Khi hết thời gian EPD, bộ xử lý 410 có thể thực hiện bước 1311, và khi chưa hết thời gian EPD, bộ xử lý 410 có thể thực hiện bước 1309.

Khi hết thời gian EPD, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể xác định để kéo dài thời gian EPD ở bước 1311. Ví dụ, bộ xử lý 410 có thể xác định xem giọng nói của người sử dụng có kết thúc hay không theo các phương pháp khác nhau được bao gồm trong các lưu đồ được mô tả ở trên, và có thể xác định để kéo dài thời gian EPD dưới dạng kết quả của việc xác định.

Ở bước 1313, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể nhận dạng ngữ cảnh hiện tại. Thông tin ngữ cảnh hiện tại có thể là giống hoặc khác với thông tin ngữ cảnh thu được ở bước 1303. Bộ xử lý 410 có thể nhận dạng thông tin ngữ cảnh hiện tại một lần nữa trước khi kéo dài thời gian EPD.

Ở bước 1315, thiết bị điện tử 400 (ví dụ, bộ xử lý 410) có thể thay đổi thời

gian kéo dài EPD theo ngữ cảnh. Thời gian kéo dài EPD có thể là ngắn hoặc dài theo thông tin ngữ cảnh theo cách giống hoặc tương tự với thời gian EPD. Ví dụ, bộ xử lý 410 có thể thiết lập thời gian kéo dài EPD (ví dụ, 3 giây) khi người nhận tin nhắn được nhập vào, và thời gian kéo dài EPD (ví dụ, 10 giây) khi nội dung tin nhắn được nhập vào là khác nhau. Tức là, thời gian kéo dài EPD có thể là khác nhau theo ngữ cảnh hiện tại ngay cả khi cùng một ứng dụng tin nhắn được thực hiện.

Fig.14 là hình vẽ minh họa ví dụ về thiết bị điện tử theo các phương án khác nhau.

Dựa vào Fig.14, theo các phương án khác nhau, thiết bị điện tử 1400 có thể bao gồm các thiết bị khác nhau gồm có các bộ phận liên quan đến việc nhận dạng giọng nói và micrô. Theo các phương án khác nhau, thiết bị điện tử 1400 có thể bao gồm các loại thiết bị khác nhau có khả năng nhận giọng nói được nói bởi người sử dụng thông qua micrô và cung cấp dịch vụ nhận dạng giọng nói. Ví dụ, thiết bị điện tử 1400 có thể bao gồm điện thoại thông minh 1410, máy tính 1420 (ví dụ, máy tính cá nhân, máy tính dạng notebook, v.v.), loa Bluetooth 1430, thiết bị kết nối vạn vật (IoT) 1440, vô tuyến 1450 hoặc tủ lạnh 1460. Thiết bị điện tử 1400 và phương pháp dịch vụ nhận dạng giọng nói của nó theo các phương án khác nhau có thể được thực hiện bằng các thiết bị khác nhau, như PC dạng bảng, thiết bị thông tin giải trí trong xe, máy giặt, máy điều hòa không khí, v.v., ngoài các cấu hình được mô tả ở trên.

Theo các phương án khác nhau, các thiết bị điện tử 1410, 1420, 1430, 1440, 1450, 1460 có thể bao gồm các danh sách miền tương ứng mà có thể thực hiện được của chúng. Ví dụ, các thiết bị điện tử 1410, 1420, 1430, 1440, 1450, 1460 có thể thiết lập các danh sách miền dựa trên chức năng mà các thiết bị này có thể thực hiện. Theo các phương án khác nhau, miền này có thể biểu thị chức năng hoặc tác vụ mà có thể thực hiện được bằng mỗi thiết bị 1410, 1420, 1430, 1440, 1450, 1460 bằng cách sử dụng dịch vụ nhận dạng giọng nói. Theo một phương án, trong trường hợp điện thoại, miền bao gồm gọi điện, tin nhắn, âm nhạc, v.v., có thể được thiết lập. Theo một phương án, trong trường hợp TV, miền bao gồm kênh, âm lượng, âm nhạc, v.v., có thể được thiết lập. Theo một phương án, trong trường hợp tủ lạnh, miền bao gồm công thức, kiểm soát nhiệt độ, thời tiết, v.v. có thể được thiết lập.

Bảng 5 thể hiện các ý tưởng có thể áp dụng được với mỗi miền liên quan đến

thiết bị.

Bảng 5

Thiết bị	Miền	Ví dụ về giọng nói	Ý tưởng có thể áp dụng được
Điện thoại	Chia sẻ nội dung	“Send this photo...um...to my mom.”	Phát hiện từ rỗng thông dụng
	Tin nhắn	“After a while, send the message I will be about... 10 minutes late.”	Phát hiện từ rỗng cá nhân
	Lịch trình	“Arrange a meeting at 1 o'clock tomorrow, and arrange a meeting at 10 o'clock in the morning the day after tomorrow.”	Phát hiện liên từ
	Đặt chỗ ở nhà hàng	Công cụ hỗ trợ: “How many people will you need the reservation for?” Người sử dụng: (nghĩ trong khi nhìn lên và sang phải) “There will be five...no...six!”	Phát hiện mắt
	...	...	...
TV	Tìm kiếm nội dung	“Search the action movie...umm, what was it? that... starring Brad Pitt.”	Phát hiện từ rỗng và trạng từ liên kết
	...	...	...
Tủ lạnh	Mua thực phẩm	“Order onion, and also, three potatoes, and also....one apple” (tạo cử chỉ bàn tay)	Phát hiện cử chỉ và trạng từ liên kết
	...	...	...
Loa	Tìm kiếm âm nhạc	“Play the album of Radiohead...what is it?...uh...the album having creep.”	Phát hiện từ rỗng
	...	...	...
...	...	...	...

Dựa vào Fig.5, các thiết bị điện tử 1410, 1420, 1430, 1440, 1450, 1460 có thể phân tích mẫu giọng nói trong mỗi miền, và có thể thiết lập ngôn ngữ hoặc phương pháp khác để xác định ý định kết thúc giọng nói của người sử dụng dựa trên mẫu giọng nói. Ví dụ, điện thoại thông minh 1410 có thể bao gồm mẫu giọng nói theo miền (ví dụ, chia sẻ nội dung, tin nhắn, lịch trình, đặt chỗ ở nhà hàng, v.v.), và có thể thiết lập ngôn ngữ hoặc phương pháp khác nhau để xác định ý định kết thúc giọng nói của người sử dụng dựa trên mẫu giọng nói. Ví dụ, trong trường hợp lệnh thoại liên quan

đến gửi tin nhắn, điện thoại thông minh 1410 có thể xác định xem từ rỗng có được bao gồm trong thông tin thoại hay không. Trong trường hợp lệnh thoại liên quan đến lịch trình, điện thoại thông minh 1410 có thể xác định xem thông tin thoại có bao gồm liên từ hay không. Nhờ đó, có thể xác định chính xác hơn ý định kết thúc giọng nói của người sử dụng. Trong trường hợp lệnh thoại liên quan đến đặt chỗ ở nhà hàng, điện thoại thông minh 1410 có thể phát hiện mất trong số các cử chỉ của người sử dụng, và bằng cách xác định xem mất của người sử dụng có tương ứng với cử chỉ định sẵn hay không, có thể xác định chính xác hơn ý định kết thúc giọng nói của người sử dụng.

Theo các phương án khác nhau, phương pháp vận hành thiết bị điện tử có thể bao gồm các bước: tính thời gian EPD dựa trên đầu vào thoại; khi hết thời gian EPD, xác định xem từ cuối cùng của đầu vào thoại có tương ứng với từ định sẵn được lưu trữ trong bộ nhớ hay không; và khi từ cuối cùng tương ứng với từ định sẵn, thì kéo dài thời gian EPD và chờ nhận đầu vào thoại.

Từ định sẵn có thể bao gồm từ thông dụng và từ cá nhân, và phương pháp này có thể còn bao gồm các bước: xác định mức độ tương tự giữa lệnh thoại được nhận dạng sau khi lỗi lệnh thoại và lệnh thoại trước đó; và thu thập từ cá nhân dựa trên mức độ tương tự này.

Bước thu thập có thể bao gồm bước phân tích thông tin văn bản thay đổi giữa lệnh thoại này và lệnh thoại trước đó, và khi phát hiện được thông tin văn bản thay đổi với số lần định sẵn hoặc nhiều hơn, thì cập nhật thông tin văn bản bằng từ cá nhân.

Phương pháp này có thể còn bao gồm bước xác định xem có phát hiện được cử chỉ định sẵn khi hết thời gian EPD hay không, và khi cử chỉ định sẵn được phát hiện, thì kéo dài thời gian EPD.

Phương pháp này có thể còn bao gồm bước xác định xem câu theo đầu vào thoại có được hoàn thành khi hết thời gian EPD hay không, và khi xác định được rằng câu này không được hoàn thành, thì kéo dài thời gian EPD.

Phương pháp này có thể còn bao gồm bước xác định thời gian EPD hoặc thời gian kéo dài EPD, dựa trên thông tin ngữ cảnh của thiết bị điện tử và thông tin đặc trưng của người sử dụng.

Phương pháp này có thể còn bao gồm bước phân tích ý định kết thúc giọng nói của người sử dụng dựa trên ít nhất một trong số thông tin ngữ cảnh của thiết bị điện tử, thông tin đặc trưng của người sử dụng, việc có phát hiện được đầu vào thoại bổ sung hay không, việc có phát hiện được từ định sẵn hay không, việc có phát hiện được cử chỉ định sẵn hay không, hoặc việc có hoàn thành câu hay không.

Vật ghi đọc được bằng máy tính theo các phương án khác nhau có thể bao gồm chương trình để thực hiện các hoạt động: tính thời gian EPD dựa trên đầu vào thoại; khi hết thời gian EPD, xác định xem từ cuối cùng của đầu vào thoại có tương ứng với từ định sẵn được lưu trữ trong bộ nhớ hay không; và khi từ cuối cùng tương ứng với từ định sẵn, thì kéo dài thời gian EPD và chờ nhận đầu vào thoại.

Vật ghi đọc được bằng máy tính có thể bao gồm đĩa cứng, đĩa mềm, vật ghi từ tính (ví dụ, băng từ), vật ghi quang học (ví dụ, đĩa nén có bộ nhớ chỉ đọc (compact disc read only memory, CD-ROM) và đĩa đa năng kỹ thuật số (Digital video disk, DVD)), vật ghi quang từ (ví dụ, đĩa mềm quang học) hoặc bộ nhớ trong. Ngoài ra, lệnh có thể bao gồm mã được tạo ra bởi trình biên dịch hoặc mã có thể thực hiện được bởi trình thông dịch. Môđun hoặc môđun chương trình theo các phương án khác nhau có thể bao gồm một hoặc nhiều bộ phận trong số các bộ phận nêu trên, một số bộ phận có thể được lược bỏ, hoặc (các) bộ phận khác có thể được thêm vào. Theo các phương án khác nhau, các hoạt động được thực hiện bởi môđun, chương trình hoặc bộ phận khác có thể được thực hiện lần lượt, song song, lặp lại hoặc theo kinh nghiệm, hoặc một hoặc nhiều hoạt động trong số các hoạt động này có thể được thực hiện theo thứ tự khác hoặc được lược bỏ, hoặc một hoặc nhiều hoạt động khác có thể được thêm vào.

Ngoài ra, các phương án được mô tả trong bản mô tả này và các hình vẽ được gợi ý nhằm để giải thích về dễ hiểu các dấu hiệu kỹ thuật được mô tả, và không được dự định là làm giới hạn phạm vi của sáng chế. Do đó, phạm vi của sáng chế cần được hiểu là bao gồm tất cả các thay đổi hoặc các dạng cải biến thu được dựa vào ý tưởng kỹ thuật của sáng chế, ngoài các phương án được mô tả trong bản mô tả này.

YÊU CẦU BẢO HỘ

1. Thiết bị điện tử (101) bao gồm:

micrô (288);

camera (430) hoặc bộ cảm biến được sắp xếp để phát hiện cử chỉ của người sử dụng;

bộ nhớ (130); và

bộ xử lý (210) được kết nối chức năng với micrô (288) hoặc bộ nhớ (130),

trong đó bộ xử lý (210) được tạo cấu hình để:

tính thời gian phát hiện điểm cuối (End Point Detection, EPD) dựa trên đầu vào thoại (639),

xác định xem từ cuối cùng của đầu vào thoại (639) có tương ứng với từ định sẵn được lưu trữ trong bộ nhớ (130) khi hết thời gian EPD hay không, và

kéo dài thời gian EPD và chờ nhận đầu vào thoại (639) khi từ cuối cùng tương ứng với từ định sẵn,

đặc trưng ở chỗ bộ xử lý (210) còn được tạo cấu hình để:

xác định xem có phát hiện được cử chỉ định sẵn bởi camera hoặc bộ cảm biến khi hết thời gian EPD hay không, và

kéo dài thời gian EPD khi phát hiện được cử chỉ định sẵn.

2. Thiết bị điện tử (101) theo điểm 1, trong đó bộ xử lý (210) được tạo cấu hình để kéo dài thời gian EPD khi từ cuối cùng tương ứng với từ định sẵn bao gồm ít nhất một trong số từ rỗng, liên từ hoặc lệnh chờ.

3. Thiết bị điện tử (101) theo điểm 1, trong đó bộ xử lý (210) được tạo cấu hình để kéo dài thời gian EPD khi phát hiện được đầu vào thoại (639) bổ sung trước khi hết thời gian EPD.

4. Thiết bị điện tử (101) theo điểm 1, trong đó từ định sẵn bao gồm từ thông dụng và từ cá nhân, và

trong đó bộ xử lý (210) được tạo cấu hình để:

xác định mức độ tương tự giữa lệnh thoại (821) được nhận dạng sau lỗi lệnh thoại và lệnh thoại (821) trước đó, và

thu thập từ cá nhân dựa trên mức độ tương tự này.

5. Thiết bị điện tử (101) theo điểm 4, trong đó bộ xử lý (210) được tạo cấu hình để:

phân tích thông tin văn bản (611) được thay đổi giữa lệnh thoại (821) và lệnh thoại (821) trước đó, và

cập nhật thông tin văn bản (611) bằng từ cá nhân khi thông tin văn bản (611) được thay đổi được phát hiện với số lần định sẵn hoặc nhiều hơn.

6. Thiết bị điện tử (101) theo điểm 1, trong đó bộ xử lý (210) được tạo cấu hình để:

xác định xem câu theo đầu vào thoại (639) có được hoàn thành khi hết thời gian EPD hay không, và

kéo dài thời gian EPD khi xác định được rằng câu này không được hoàn thành.

7. Thiết bị điện tử (101) theo điểm 6, trong đó bộ xử lý (210) được tạo cấu hình để xác định xem có thực hiện bước xác định xem câu có được hoàn thành hay không, dựa trên loại lệnh thoại (821) theo đầu vào thoại (639).

8. Thiết bị điện tử (101) theo điểm 1, trong đó bộ xử lý (210) được tạo cấu hình để:

kéo dài thời gian EPD theo giá trị cố định, hoặc

thay đổi thời gian EPD đến giá trị tương ứng với sự nhận dạng ngữ cảnh, và kéo dài thời gian EPD theo giá trị đã thay đổi.

9. Thiết bị điện tử (101) theo điểm 1, trong đó bộ xử lý (210) được tạo cấu hình để xác định thời gian EPD hoặc thời gian kéo dài EPD, dựa trên thông tin ngữ cảnh của thiết bị điện tử và thông tin đặc trưng của người sử dụng.

10. Thiết bị điện tử (101) theo điểm 1, trong đó bộ xử lý (210) được tạo cấu hình để phân tích ý định kết thúc giọng nói của người sử dụng dựa trên ít nhất một trong số thông tin ngữ cảnh của thiết bị điện tử (101), thông tin đặc trưng của người sử dụng, việc có phát hiện được đầu vào thoại (639) bổ sung hay không, việc có phát hiện được từ định sẵn hay không, việc có phát hiện được cử chỉ định sẵn hay không, hoặc việc có hoàn thành câu hay không.

11. Thiết bị điện tử (101) theo điểm 10, trong đó bộ xử lý (210) được tạo cấu hình để:

tính toán chỉ số liên quan đến ý định kết thúc giọng nói của người sử dụng bằng cách gán giá trị trọng số hoặc điểm cho ít nhất một trong số thời gian phát hiện sự im lặng, việc có phát hiện được từ định sẵn hay không, việc có phát hiện được cử chỉ định sẵn hay không, hoặc việc có hoàn thành câu hay không, và

kéo dài thời gian EPD khi chỉ số này lớn hơn hoặc bằng chỉ số định sẵn.

12. Thiết bị điện tử theo điểm 11, trong đó bộ xử lý được tạo cấu hình để thay đổi thời gian kéo dài EPD dựa trên chỉ số.

13. Phương pháp vận hành thiết bị điện tử (101), trong đó phương pháp này bao gồm các bước:

tính thời gian EPD dựa trên đầu vào thoại (639);

khi hết thời gian EPD, xác định xem từ cuối cùng của đầu vào thoại (639) có tương ứng với từ định sẵn được lưu trữ trong bộ nhớ hay không; và

khi từ cuối cùng tương ứng với từ định sẵn, thì kéo dài thời gian EPD và chờ nhận đầu vào thoại (639),

được đặc trưng ở chỗ phương pháp này còn bao gồm các bước:

xác định xem có phát hiện được cử chỉ định sẵn khi hết thời gian EPD hay không; và

kéo dài thời gian EPD khi phát hiện được cử chỉ định sẵn.

14. Phương pháp theo điểm 13, trong đó từ định sẵn bao gồm từ thông dụng và từ cá nhân, và

trong đó phương pháp này còn bao gồm các bước:

xác định mức độ tương tự giữa lệnh thoại (821) được nhận dạng sau khi lỗi lệnh thoại và lệnh thoại trước đó; và

thu thập từ cá nhân dựa trên mức độ tương tự này.

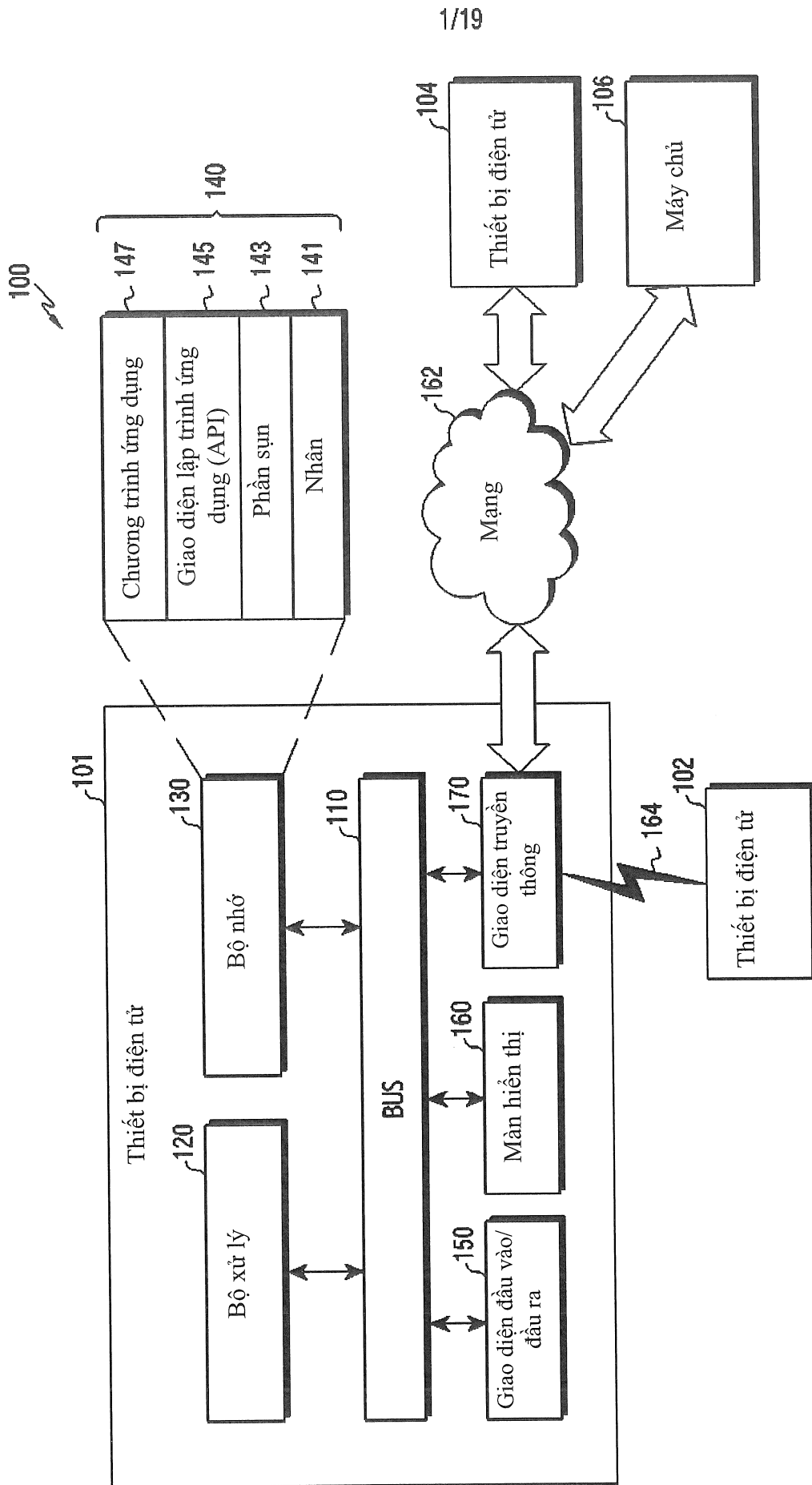


FIG.1

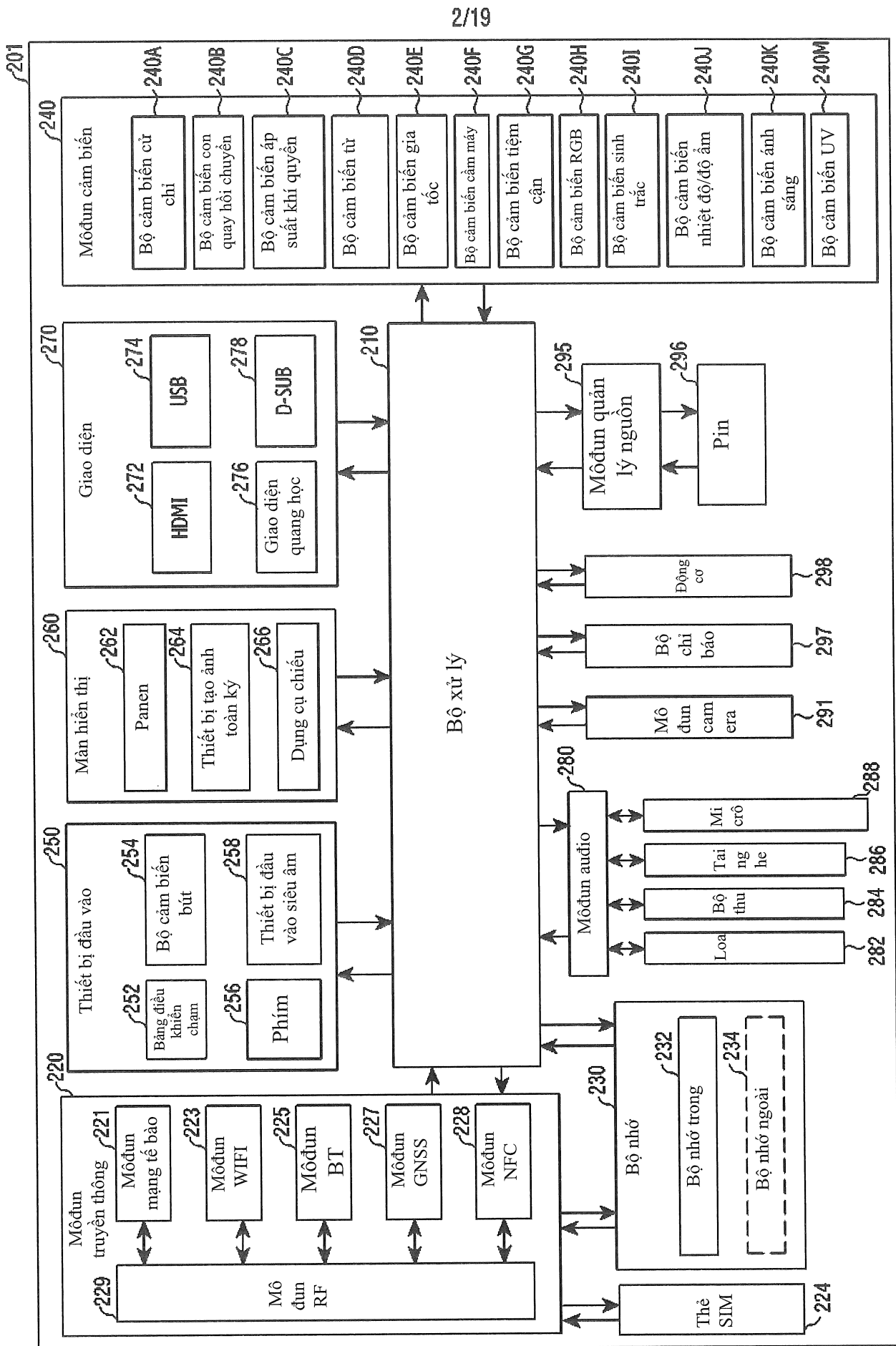
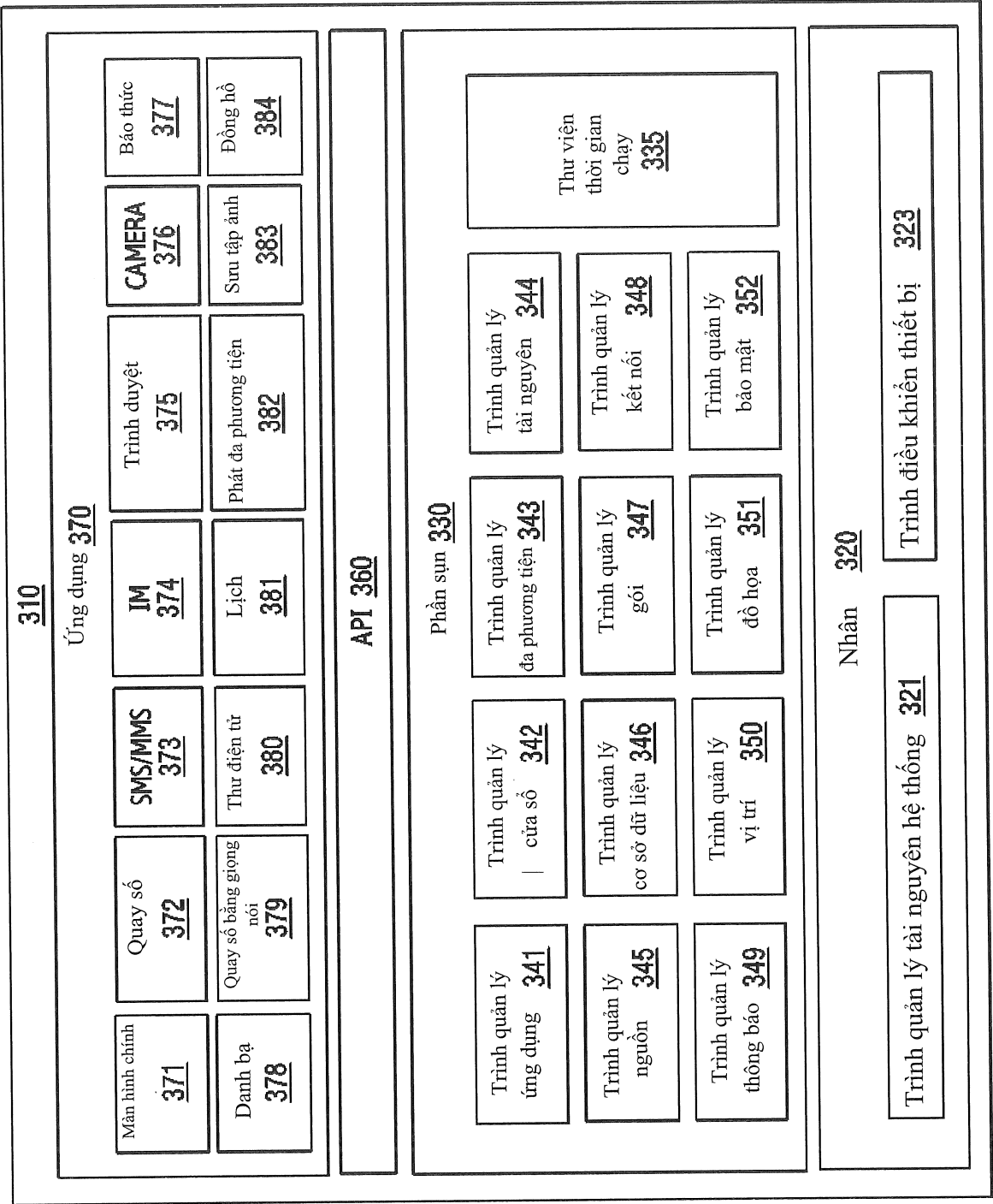


FIG.2



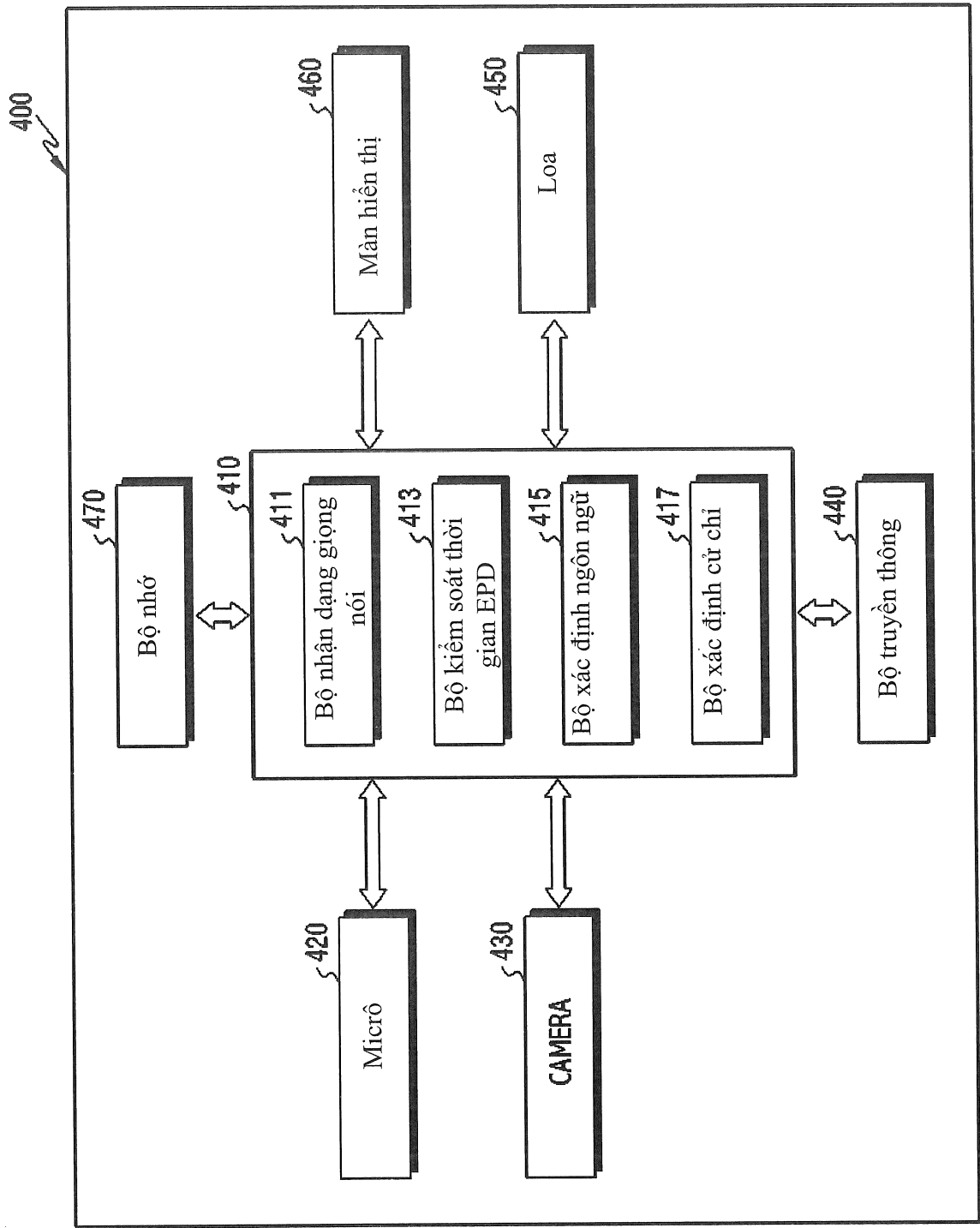


FIG.4

5/19

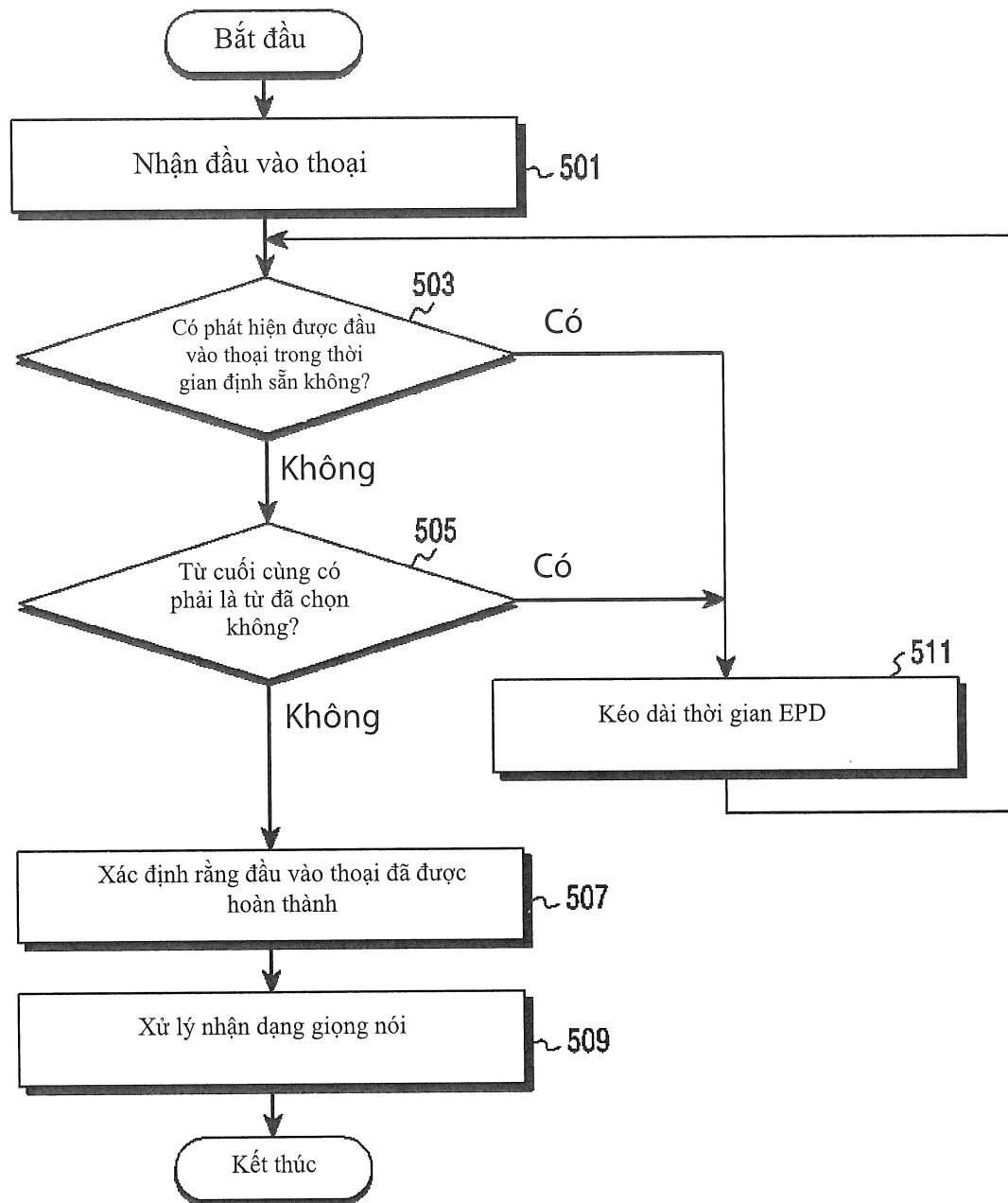


FIG.5A

6/19

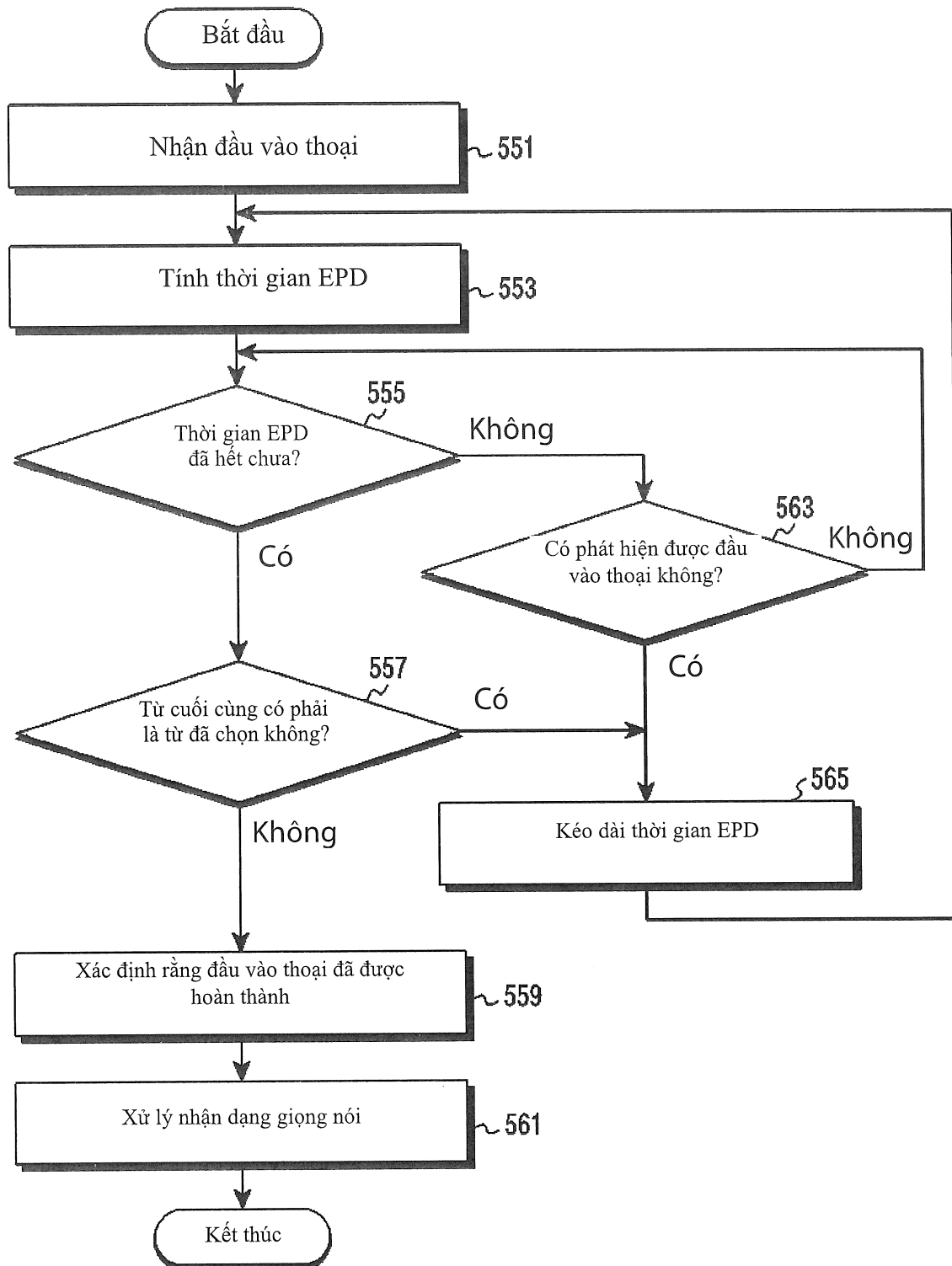


FIG.5B

7/19

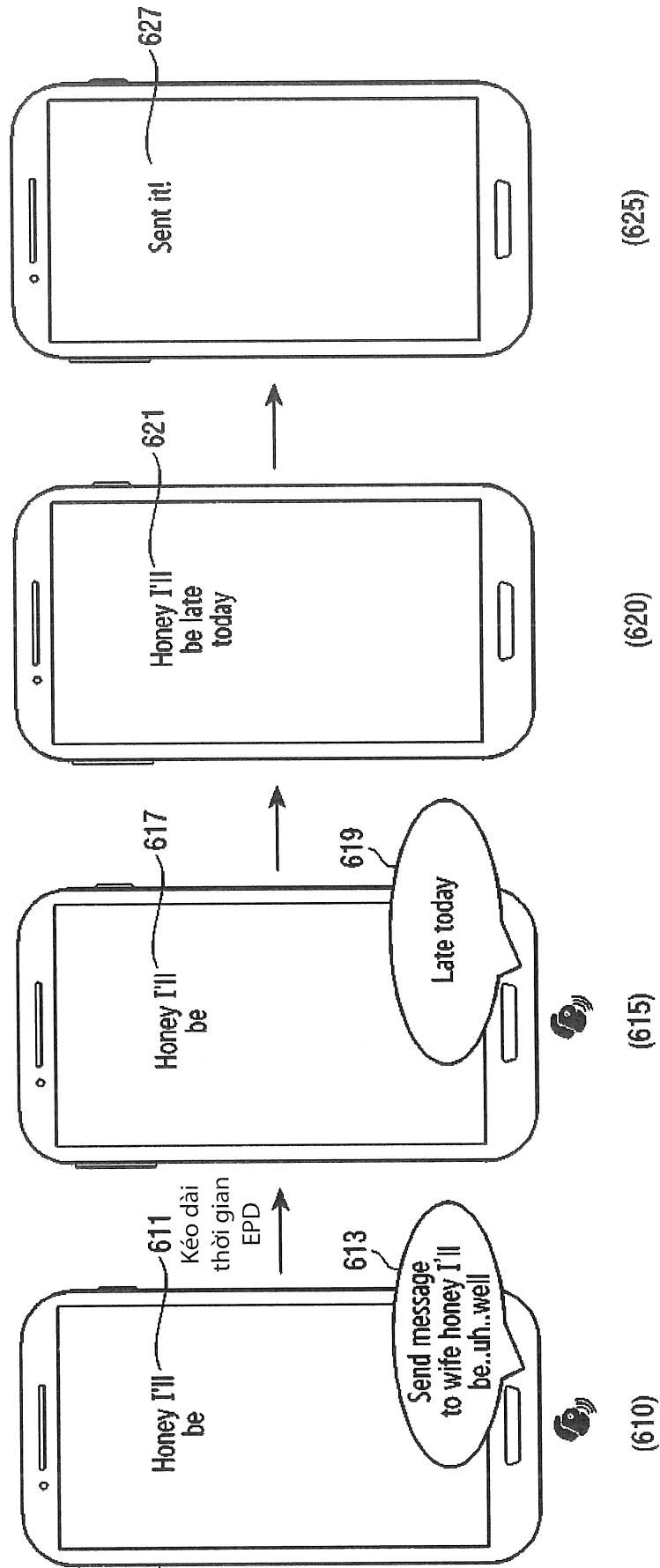


FIG.6A

8/19

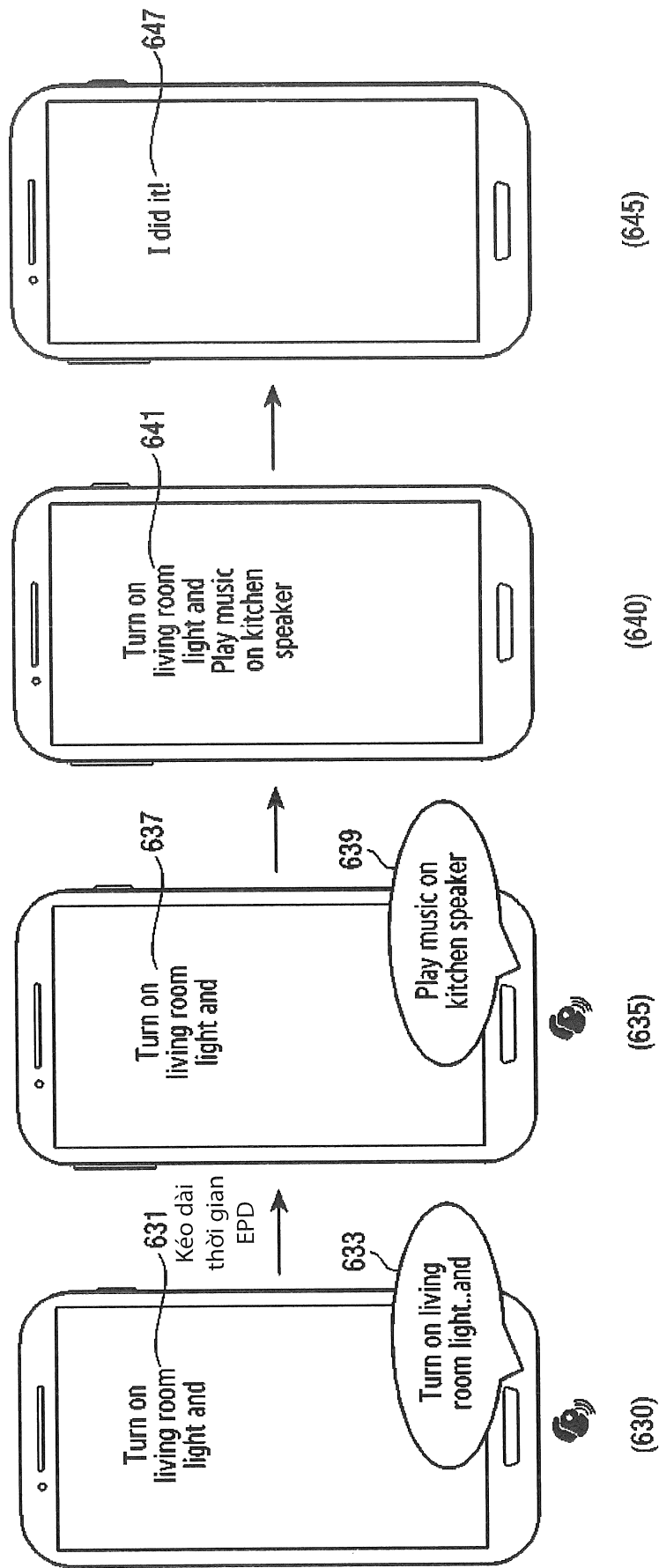


FIG.6B

9/19

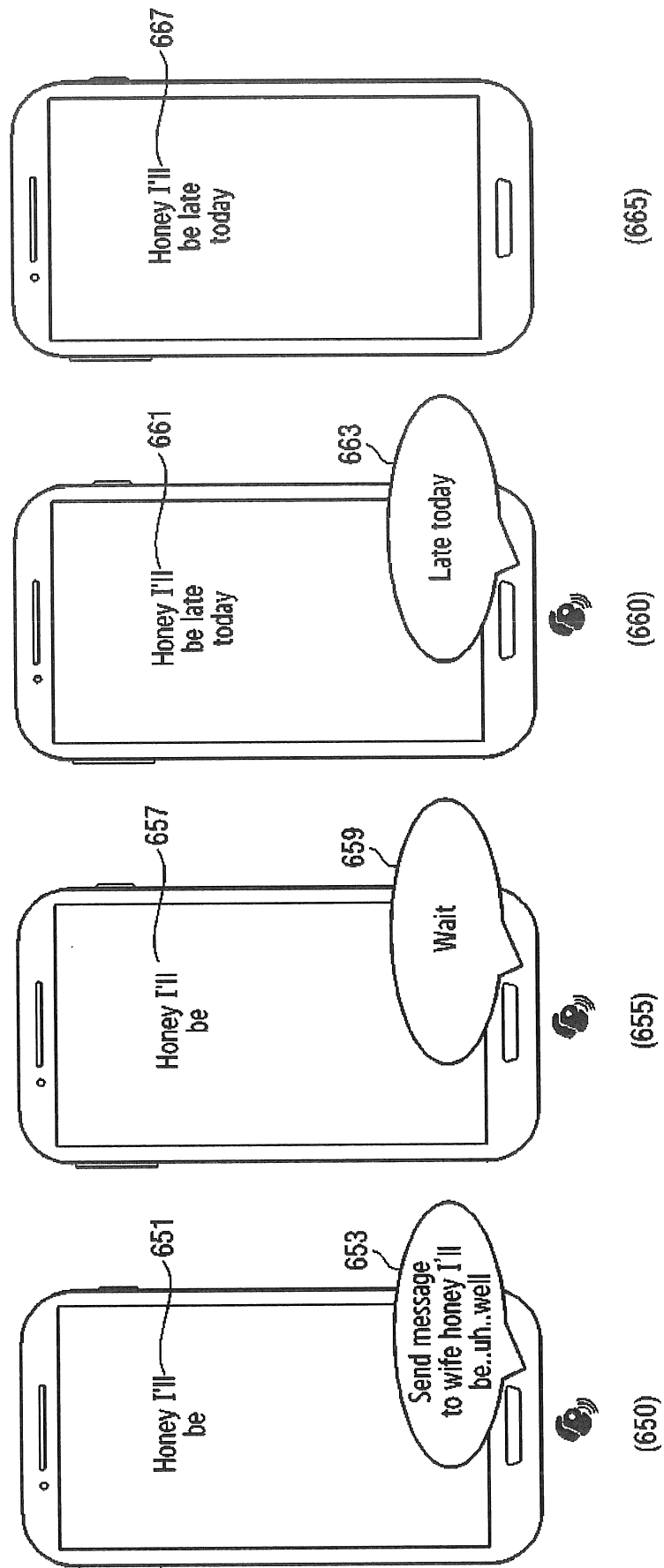


FIG.6C

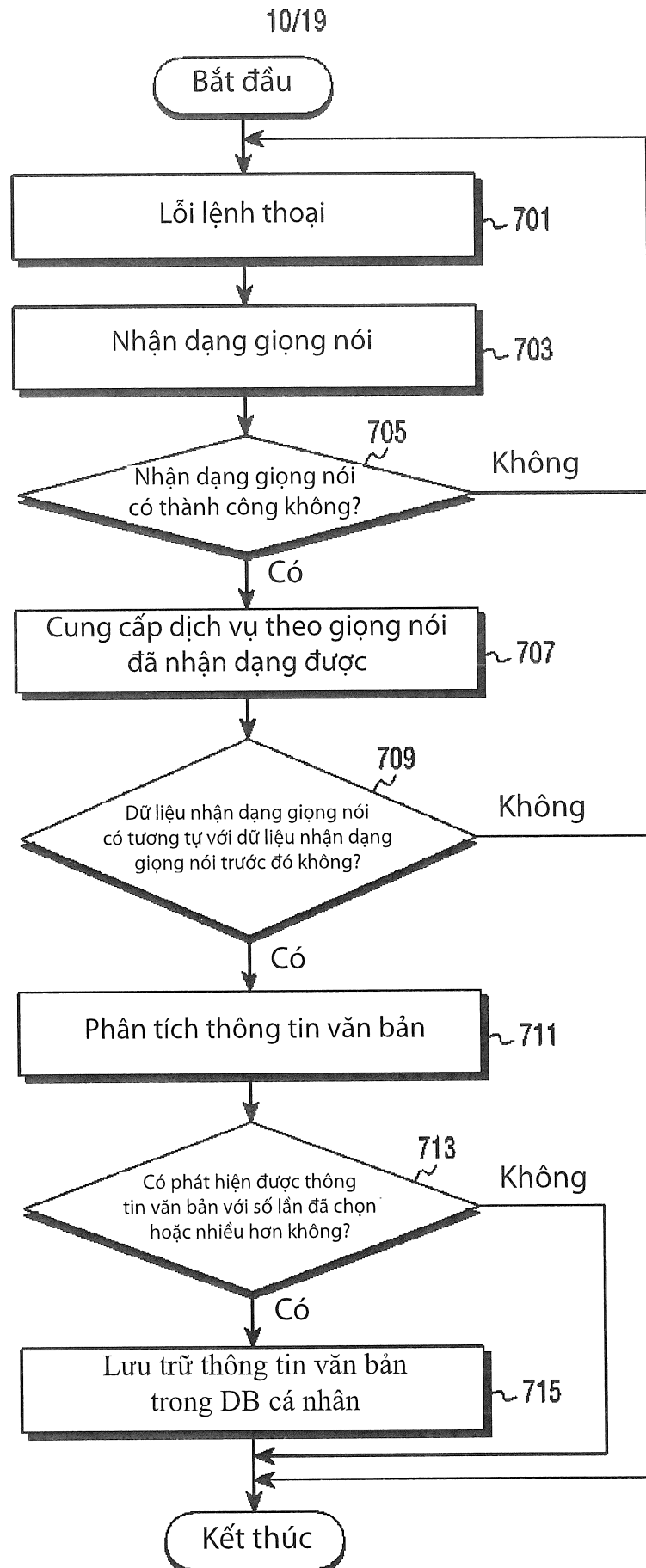


FIG. 7

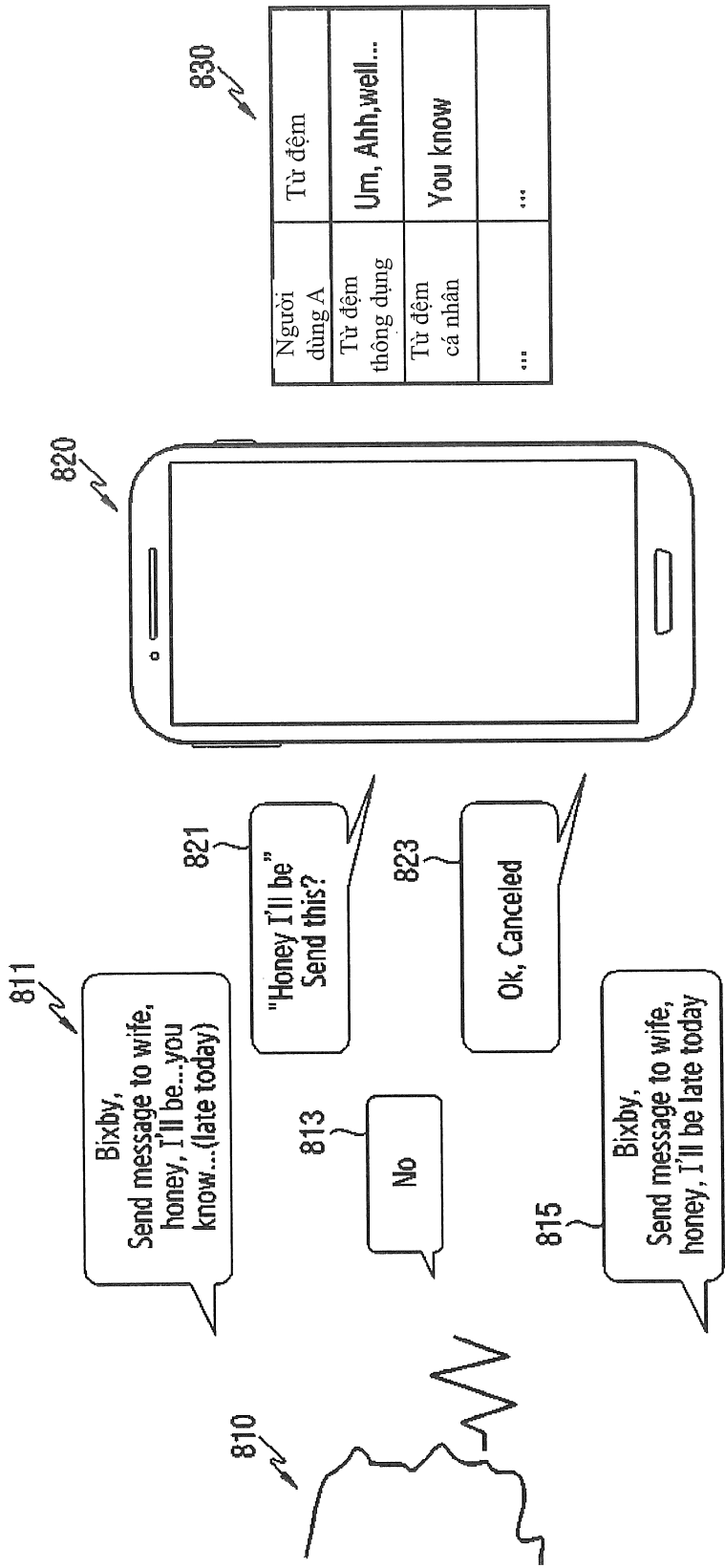


FIG.8A

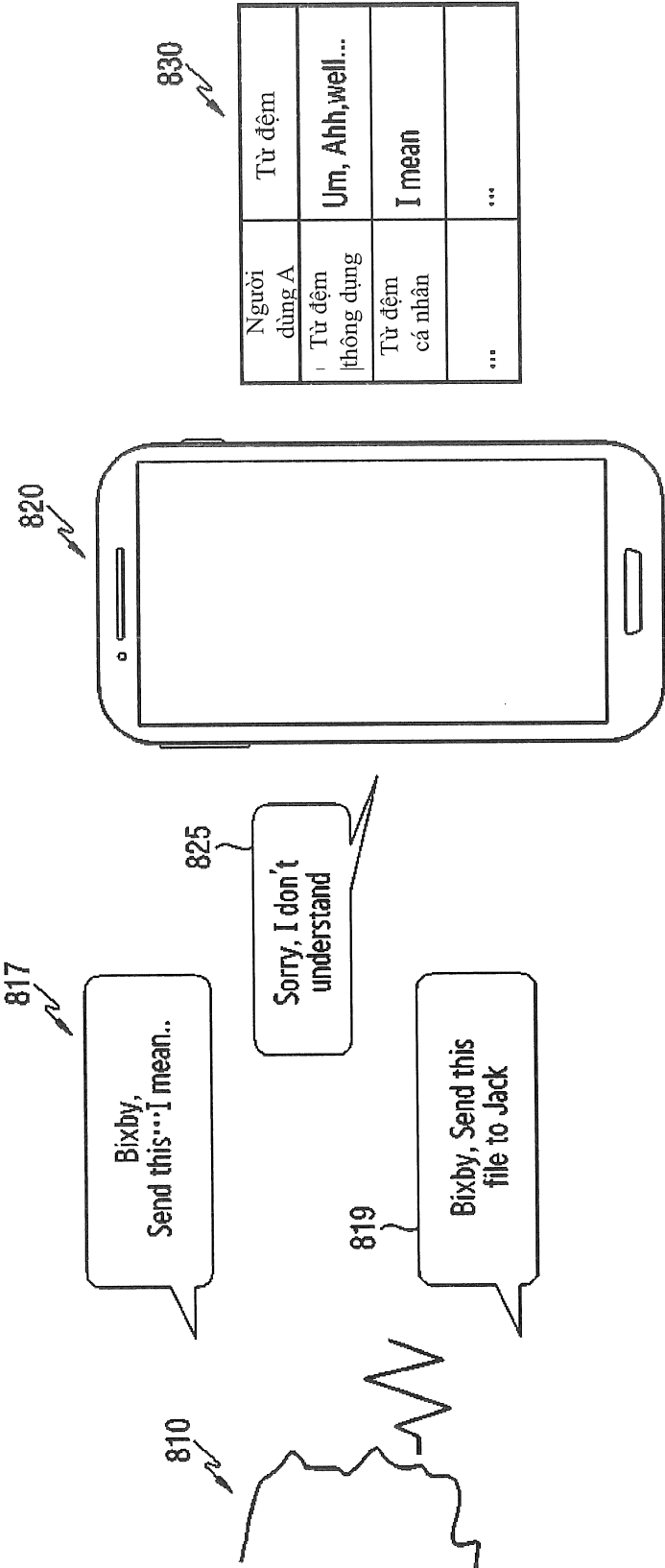


FIG.8B

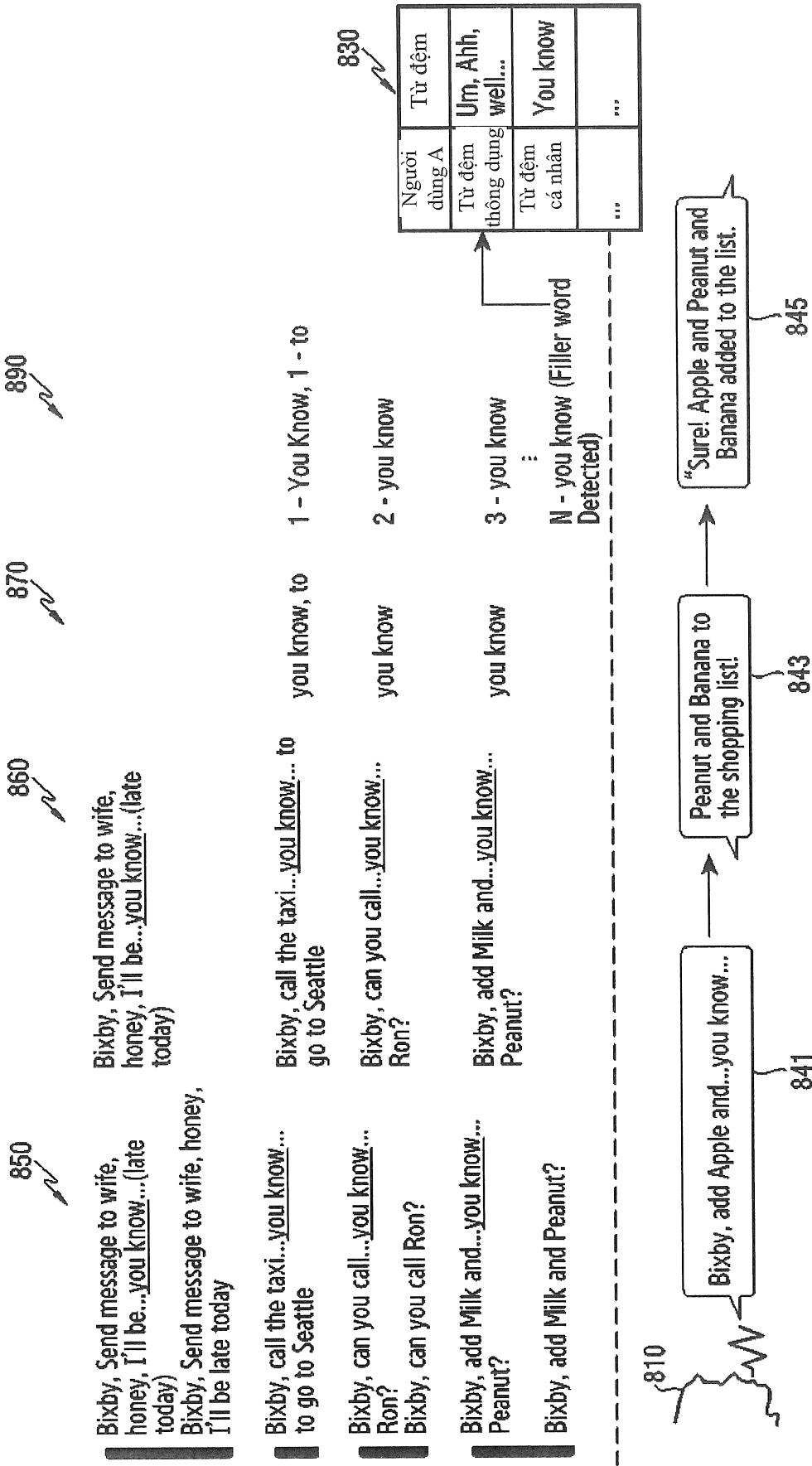


FIG.8C

14/19

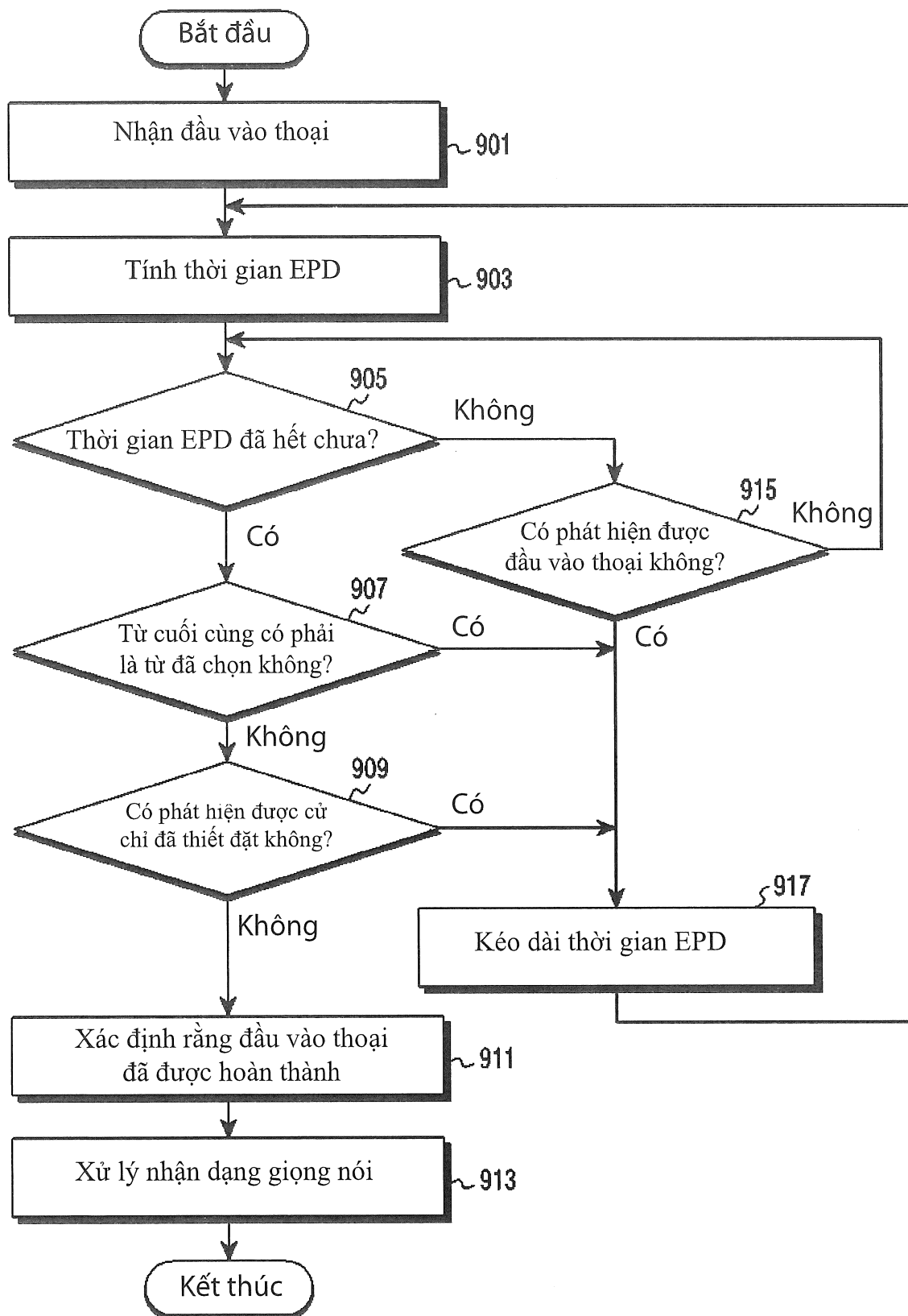


FIG.9

15/19

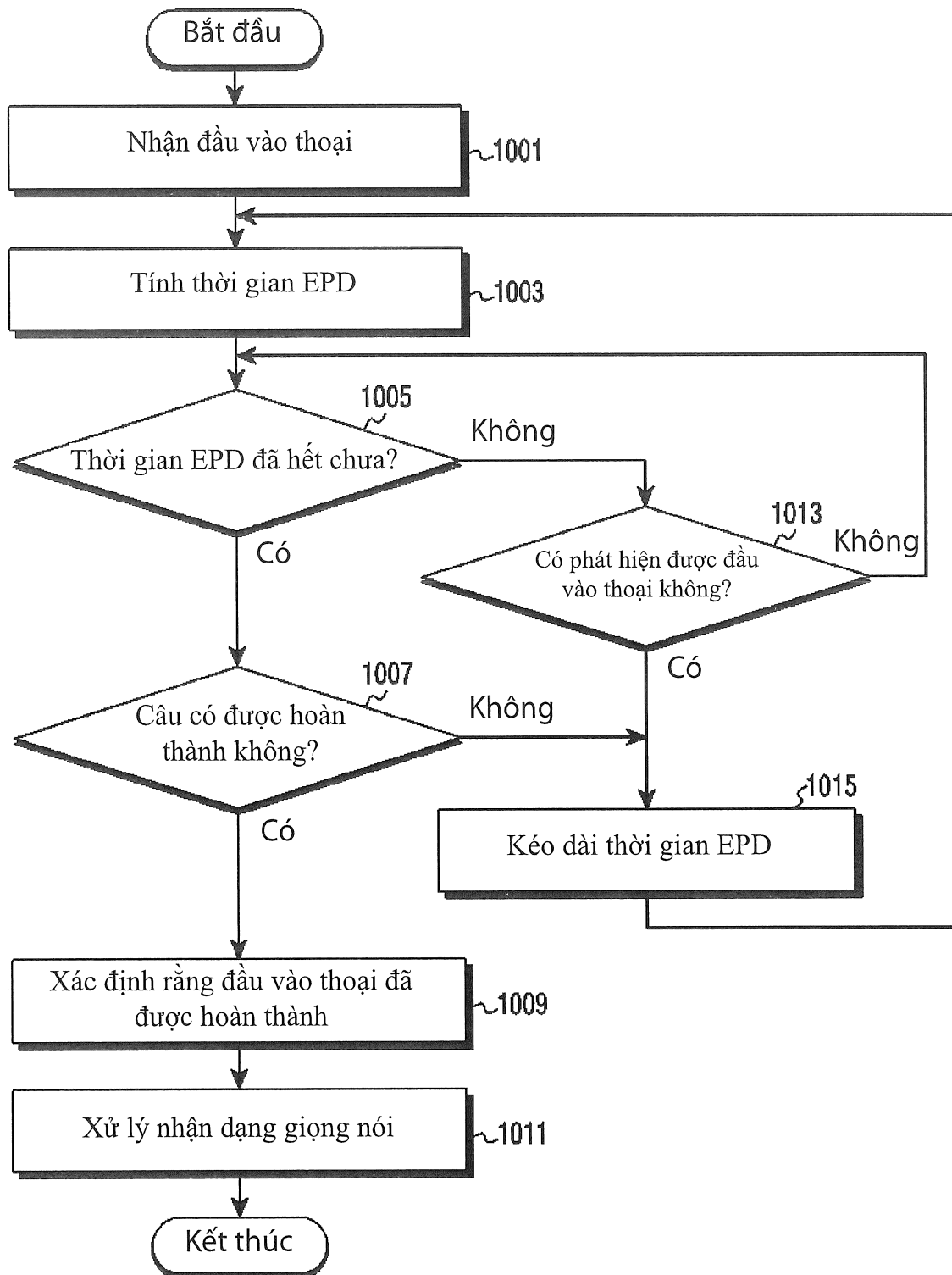


FIG.10

16/19

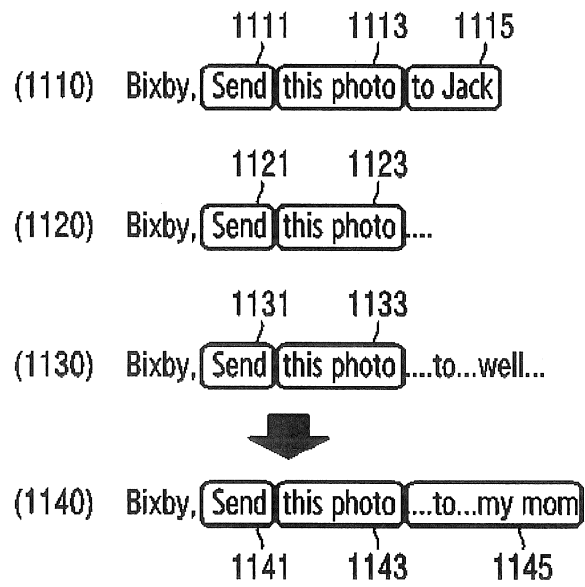


FIG.11

17/19

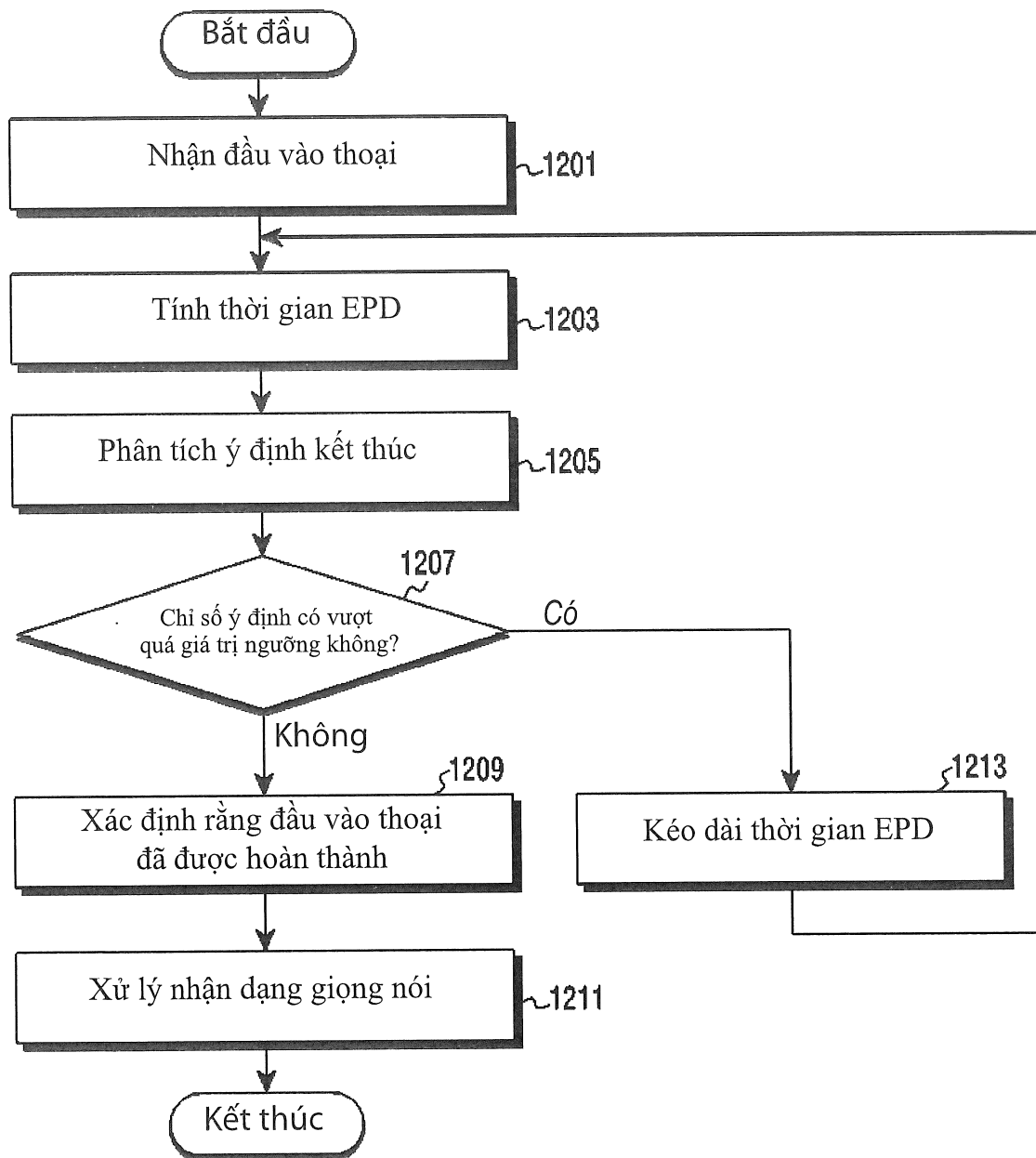


FIG.12

18/19

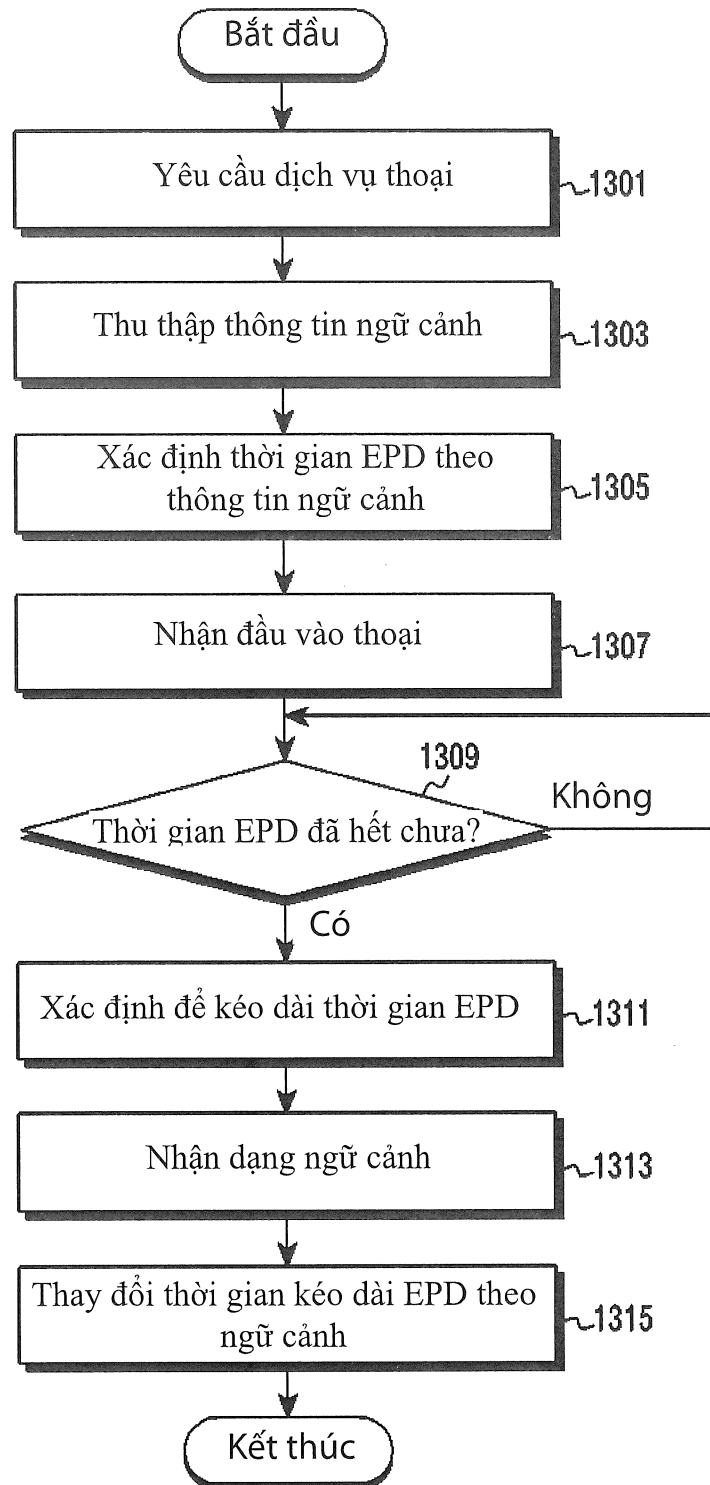


FIG.13

19/19

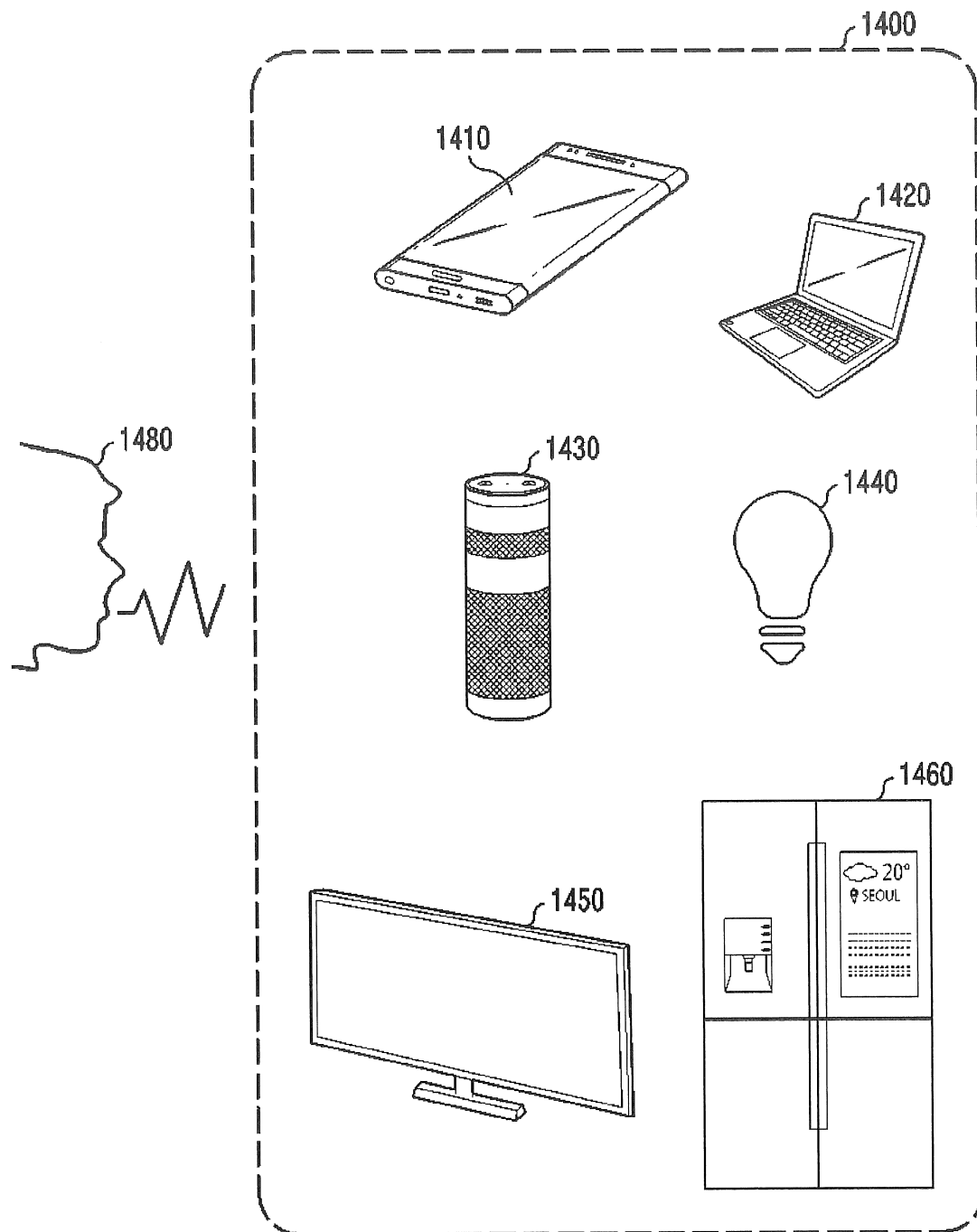


FIG. 14