



(12) BẢN MÔ TẢ SÁNG CHẾ THUỘC BẰNG ĐỘC QUYỀN SÁNG CHẾ

(19) Cộng hòa xã hội chủ nghĩa Việt Nam (VN)
CỤC SỞ HỮU TRÍ TUỆ

(11)



1-0033421

(51)⁷ G06F 21/32; H04L 29/06; G06K 9/00

(13) B

(21) 1-2018-05504

(22) 12/04/2017

(86) PCT/CN2017/080196 12/04/2017

(87) WO 2017/198014 23/11/2017

(30) 201610340549.4 19/05/2016 CN

(45) 26/09/2022 414

(43) 27/05/2019 374A

(73) Advanced New Technologies Co., Ltd. (KY)

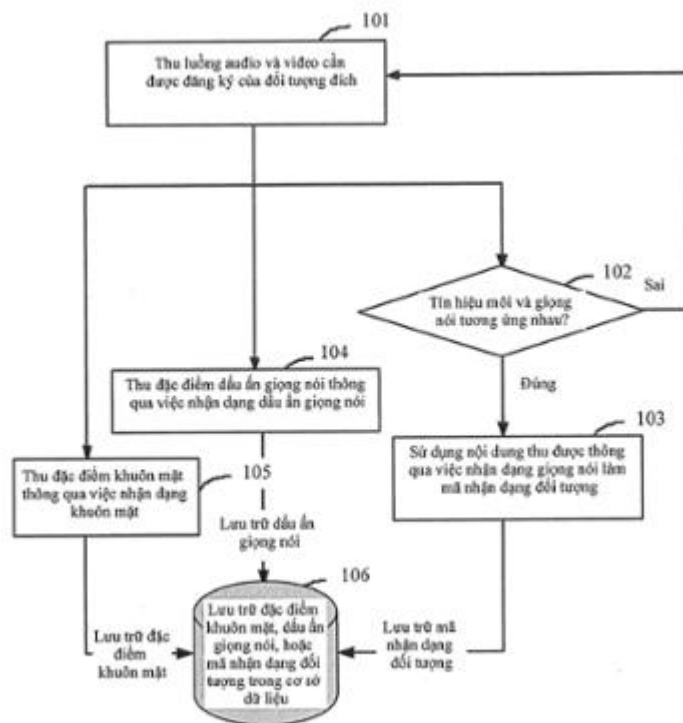
Cayman Corporate Centre, 27 Hospital Road, George Town, Grand Cayman KY1-9008, Cayman Islands

(72) LI, Peng (CN); SUN, Yipeng (CN); XIE, Yongxiang (CN); LI, Liang (CN).

(74) Công ty TNHH Tầm nhìn và Liên danh (VISION & ASSOCIATES CO.LTD.)

(54) PHƯƠNG PHÁP VÀ THIẾT BỊ XÁC THỰC CĂN CƯỚC

(57) Sáng chế đề cập đến phương pháp và thiết bị xác thực căn cước. Phương pháp này bao gồm các bước thu luồng audio và video đã thu thập được tạo ra bởi đối tượng đích cần được xác thực; xác định xem liệu tín hiệu môi và giọng nói trong luồng audio và video có tương ứng với nhau hay không, và nếu tín hiệu môi và giọng nói tương ứng với nhau, sử dụng nội dung giọng nói thu được bằng cách thực hiện nhận dạng giọng nói đối với luồng audio trong luồng audio và video làm mã nhận dạng đối tượng của đối tượng đích; thu đặc điểm sinh lý mẫu tương ứng với mã nhận dạng đối tượng từ thông tin đăng ký đối tượng, nếu thông tin đăng ký đối tượng được lưu trữ từ trước bao gồm mã nhận dạng đối tượng; thực hiện nhận dạng sinh lý đối với luồng audio và video để thu đặc điểm sinh lý của đối tượng đích; và so sánh đặc điểm sinh lý của đối tượng đích với đặc điểm sinh lý mẫu để thu kết quả so sánh, và nếu kết quả so sánh thỏa mãn điều kiện xác thực, xác định rằng đối tượng đích đã được xác thực. Sáng chế cải thiện hiệu quả và độ tin cậy của việc xác thực căn cước.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến các công nghệ mạng, và cụ thể là đề cập đến phương pháp và thiết bị xác thực căn cước.

Tình trạng kỹ thuật của sáng chế

Với sự phát triển của các công nghệ Internet, các dịch vụ dựa vào mạng được sử dụng rộng rãi bởi người sử dụng. Ví dụ, người sử dụng có thể gửi và nhận các thư điện tử bằng cách sử dụng hộp thư, mua sắm trực tuyến, hoặc thậm chí làm việc trực tuyến. Một số ứng dụng đòi hỏi yêu cầu bảo mật cao, và căn cước của người sử dụng cần được xác thực. Ví dụ, căn cước của người sử dụng cần được xác thực trước khi thanh toán được xác thực để mua sắm trực tuyến; hoặc người sử dụng có thể đăng nhập vào ứng dụng với yêu cầu bảo mật tương đối cao sau khi căn cước của người sử dụng đã được xác thực. Trong các công nghệ liên quan, có nhiều các phương pháp xác thực căn cước được sử dụng trên Internet, chẳng hạn như nhận dạng khuôn mặt và nhận dạng dấu ấn giọng nói. Tuy nhiên, các phương pháp xác thực được sử dụng nói chung tương đối phức tạp. Ví dụ, người sử dụng cần nhập ID của người sử dụng và sau đó xác minh dấu ấn giọng nói để nhận dạng dấu ấn giọng nói. Ngoài ra, các phương pháp xác thực hiện tại có độ tin cậy tương đối thấp. Ví dụ, kẻ tấn công có thể thu được nhận dạng khuôn mặt bằng cách sử dụng video tương tự hoặc bản ghi. Ngay cả nếu việc xác thực được thực hiện dựa vào hai phương pháp xác thực kết hợp, ví dụ, dựa vào khuôn mặt và dấu ấn giọng nói, do các phương pháp xác thực này tương đối độc lập với nhau, nên kẻ tấn công có thể phá vỡ thông qua các quy trình xác thực. Các nhược điểm của các phương pháp xác thực căn cước được mô tả trên có thể tạo ra một số rủi ro để bảo mật ứng dụng.

Bản chất kỹ thuật của sáng chế

Trên cơ sở nêu trên, sáng chế đề xuất phương pháp và thiết bị xác thực căn cước, để tăng cường hiệu quả và độ tin cậy của việc xác thực căn cước.

Cụ thể, sáng chế được thực hiện bằng các sử dụng các giải pháp kỹ thuật sau.

Theo khía cạnh thứ nhất, phương pháp xác thực căn cước được đề xuất, và

phương pháp này bao gồm các bước: thu luồng audio và video đã thu thập, trong đó luồng audio và video được tạo ra bởi đối tượng đích cần được xác thực; xác định xem liệu tín hiệu môi và giọng nói trong luồng audio và video có tương ứng với nhau hay không, và nếu tín hiệu môi và giọng nói tương ứng với nhau, sử dụng nội dung giọng nói thu được bằng cách thực hiện nhận dạng giọng nói đối với luồng audio của luồng audio và video làm mã nhận dạng đối tượng của đối tượng đích; thu đặc điểm sinh lý mẫu tương ứng với mã nhận dạng đối tượng từ thông tin đăng ký đối tượng, nếu thông tin đăng ký đối tượng được lưu trữ từ trước bao gồm mã nhận dạng đối tượng; thực hiện nhận dạng sinh lý đối với luồng audio và video để thu đặc điểm sinh lý của đối tượng đích; và so sánh đặc điểm sinh lý của đối tượng đích với đặc điểm sinh lý mẫu để thu kết quả so sánh, và nếu kết quả so sánh thỏa mãn điều kiện xác thực, thì xác định rằng đối tượng đích đã được xác thực.

Theo khía cạnh thứ hai, thiết bị xác thực căn cước được đề xuất, và thiết bị này bao gồm môđun thu thông tin, được tạo cấu hình để thu luồng audio và video đã thu thập, trong đó luồng audio và video được tạo ra bởi đối tượng đích cần được xác thực; môđun xác định mã nhận dạng, được tạo cấu hình để xác định xem liệu tín hiệu môi và giọng nói trong luồng audio và video có tương ứng với nhau không, và nếu tín hiệu môi và giọng nói tương ứng với nhau, sử dụng nội dung giọng nói thu được bằng cách thực hiện nhận dạng giọng nói đối với luồng audio trong luồng audio và video làm mã nhận dạng đối tượng của đối tượng đích; môđun quản lý thông tin, được tạo cấu hình để thu đặc điểm sinh lý mẫu tương ứng với mã nhận dạng đối tượng từ thông tin đăng ký đối tượng, nếu thông tin đăng ký đối tượng được lưu trữ từ trước bao gồm mã nhận dạng đối tượng; môđun nhận dạng đặc điểm, được tạo cấu hình để thực hiện nhận dạng sinh lý đối với luồng audio và video để thu đặc điểm sinh lý của đối tượng đích; và môđun xử lý xác thực, được tạo cấu hình để so sánh đặc điểm sinh lý của đối tượng đích với đặc điểm sinh lý mẫu để thu kết quả so sánh, và nếu kết quả so sánh thỏa mãn điều kiện xác thực thì xác định rằng đối tượng đích đã được xác thực.

Theo phương pháp và thiết bị xác thực căn cước được đề xuất theo sáng chế, mã nhận dạng người sử dụng thu được thông qua việc nhận dạng luồng audio và video trong khi xác thực người sử dụng, và đặc điểm khuôn mặt và đặc điểm dấu ấn giọng nói cũng có thể được xác minh bằng cách sử dụng cùng luồng audio và video. Nó giúp đơn giản hóa thao tác của người sử dụng, tăng cường hiệu quả xác thực, duy trì mẫu

xác thực một-một, đảm bảo nhận dạng chính xác. Ngoài ra, trong phương pháp này, sự thống nhất giữa tín hiệu môi và giọng nói được xác định để đảm bảo rằng đối tượng đích là đối tượng thực thay vì đoạn ghi video giả của kẻ tấn công, nhờ đó tăng cường tính bảo mật và độ tin cậy xác thực.

Mô tả vắn tắt các hình vẽ

FIG. 1 minh họa thủ tục đăng ký căn cước, theo một phương án ví dụ của sáng chế;

FIG. 2 minh họa thủ tục xác định tính thống nhất giữa tín hiệu môi và giọng nói, theo một phương án ví dụ của sáng chế;

FIG. 3 minh họa thủ tục nhận dạng đặc điểm khuôn mặt, theo một phương án ví dụ của sáng chế;

FIG. 4 minh họa thủ tục nhận dạng đặc điểm dấu ấn giọng nói, theo một phương án ví dụ của sáng chế;

FIG. 5 minh họa thủ tục xác thực căn cước, theo một phương án ví dụ của sáng chế;

FIG. 6 là hình vẽ sơ đồ cấu trúc của thiết bị xác thực căn cước, theo một phương án ví dụ của sáng chế; và

FIG. 7 là hình vẽ sơ đồ cấu trúc của thiết bị xác thực căn cước, theo một phương án ví dụ của sáng chế.

Mô tả chi tiết sáng chế

Các phương án ví dụ được mô tả chi tiết ở đây, và các ví dụ của các phương án ví dụ được biểu diễn trong các hình vẽ kèm theo. Khi phần mô tả sau liên quan đến các hình vẽ kèm theo, trừ khi được chỉ ra khác đi, các số chỉ dẫn giống nhau trong các hình vẽ kèm theo biểu diễn các thành phần giống hoặc tương tự nhau. Các phương án được mô tả trong các phương án ví dụ không đại diện cho tất cả các phương án theo sáng chế. Thay vào đó, chúng chỉ là các ví dụ về các thiết bị và phương pháp phù hợp với một số khía cạnh của sáng chế được mô tả chi tiết trong các điểm yêu cầu bảo hộ.

Phương pháp xác thực căn cước được đề xuất trong một phương án của sáng chế có thể được áp dụng cho việc xác thực nhận dạng Internet. Ví dụ, người sử dụng

có thể đăng nhập vào ứng dụng mạng sau khi đã được xác thực bằng cách sử dụng phương pháp này để đảm bảo bảo mật cho ứng dụng.

Việc sử dụng ứng dụng với yêu cầu bảo mật tương đối cao sau đây là một ví dụ. Giả sử rằng ứng dụng có thể chạy trên thiết bị thông minh của người sử dụng, ví dụ, điện thoại thông minh hoặc máy tính bảng thông minh. Khi người sử dụng cần đăng nhập vào ứng dụng trên thiết bị thông minh, luồng audio và video có thể được thu thập bằng cách sử dụng camera và micrô trên thiết bị thông minh. Ví dụ, người sử dụng có thể đọc ID ứng dụng của người sử dụng vào camera và micrô của điện thoại di động. ID ứng dụng có thể là số tài khoản "123456" mà được đăng ký bởi người sử dụng trong ứng dụng. Sau khi người sử dụng đọc ID ứng dụng, điện thoại di động có thể thu thập luồng audio và video của người sử dụng, bao gồm hình ảnh video và giọng đọc.

Theo phương pháp xác thực căn cước trong phương án này của sáng chế, luồng audio và video được thu thập có thể được xử lý. Trước khi xác thực, người sử dụng cần thực hiện thủ tục đăng ký căn cước để thực hiện xác thực căn cước sau đó. Thủ tục đăng ký cũng được thực hiện dựa vào luồng audio và video được thu thập. Phần sau đây mô tả riêng rẽ thủ tục đăng ký căn cước và thủ tục xác thực căn cước. Ngoài ra, phương án này không áp đặt giới hạn về thiết bị thực hiện đăng ký căn cước hoặc quy trình xác thực trong các ứng dụng thực. Ví dụ, sau khi luồng audio và video của người sử dụng được thu thập, điện thoại thông minh có thể truyền luồng audio và video đến máy chủ phía sau của ứng dụng, hoặc một phần của quy trình có thể được thực hiện trên phía phần mềm khách của điện thoại thông minh, và các phần khác của quy trình có thể được thực hiện trên phía máy chủ; hoặc các phương pháp khác có thể được sử dụng.

Đăng ký căn cước

Trong phương pháp của phương án này, khi người sử dụng thực hiện đăng ký căn cước, hai kiểu thông tin có thể được đưa vào. Một kiểu thông tin là mã nhận dạng đối tượng. Ví dụ, khi người sử dụng đăng nhập vào ứng dụng, người sử dụng có thể được xem là đối tượng đích. Khi người sử dụng đăng ký ứng dụng, thông tin được sử dụng để phân biệt người sử dụng với người sử dụng khác là mã nhận dạng đối tượng trong ứng dụng, ví dụ, có thể là số tài khoản 123456 của người sử dụng trong ứng dụng, và số tài khoản 123456 là mã nhận dạng đối tượng của đối tượng đích. Kiểu

thông tin khác là thông tin sinh lý mà có thể nhận dạng duy nhất người sử dụng, ví dụ, đặc điểm dấu ấn giọng nói của người sử dụng hoặc đặc điểm khuôn mặt của người sử dụng. Thông thường, dấu ấn giọng nói và khuôn mặt của những người khác nhau là khác nhau, và thông tin sinh lý mà nhận dạng mỗi người sử dụng có thể được xem là đặc điểm sinh lý mẫu.

Mỗi quan hệ ánh xạ được thiết lập giữa hai kiểu thông tin: mã nhận dạng đối tượng và đặc điểm sinh lý mẫu và mỗi quan hệ ánh xạ được lưu trữ. Mã nhận dạng đối tượng được lưu trữ tương ứng và đặc điểm sinh lý mẫu của đối tượng đích có thể được xem là "thông tin đăng ký đối tượng". Ví dụ, người sử dụng Xiao Zhang có thể lưu trữ thông tin đăng ký đối tượng của người sử dụng là "123456—đặc điểm sinh lý mẫu A". Đối với mã nhận dạng chính xác của người sử dụng, thông tin sinh lý được bao gồm trong đặc điểm sinh lý mẫu được sử dụng trong ví dụ này có thể bao gồm ít nhất hai kiểu thông tin sinh lý, ví dụ, khuôn mặt và dấu ấn giọng nói.

FIG. 1 minh họa thủ tục đăng ký căn cước ví dụ. Thủ tục này bao gồm các bước sau:

Bước 101: Thu luồng audio và video cần được đăng ký của đối tượng đích.

Ví dụ, người sử dụng đăng ký ứng dụng, và người sử dụng có thể đọc số tài khoản "123456" của ứng dụng vào thiết bị thông minh của người sử dụng, chẳng hạn như điện thoại di động. Trong ví dụ này, người sử dụng thực hiện đăng ký có thể được xem là đối tượng đích; và camera và micrô của thiết bị thông minh có thể thu thập luồng audio và video được tạo ra khi người sử dụng đọc số tài khoản. Luồng audio và video được thu thập trong quá trình đăng ký có thể được xem là luồng audio và video cần được đăng ký, bao gồm luồng audio và luồng video. Luồng audio là giọng đọc của người sử dụng, và luồng video là hình ảnh video đọc của người sử dụng.

Sau khi luồng audio và video trong bước này được thu, ba khía cạnh sau đây của quy trình có thể được thực hiện để hoàn tất đăng ký cho người sử dụng. Một cách chi tiết, vẫn có thể tham chiếu đến FIG. 1.

Theo một khía cạnh, quy trình là như sau: Bước 102: Xác định xem liệu tín hiệu môi và giọng nói trong luồng audio và video được đăng ký có tương ứng với nhau hay không.

Sự thống nhất ở đây chỉ ra mối quan hệ ánh xạ giữa cử động môi và chuyển động được chỉ ra bởi giọng nói. Ví dụ, giả sử rằng giọng nói là "the weather today is sunny", trong đó giọng nói được tạo ra ở tốc độ thấp, nhưng cử động môi tương ứng với "the weather today is sunny" được tạo ra ở tốc độ cao. Rõ ràng là giọng nói và cử động môi không tương ứng với nhau: khi cử động môi đã dừng lại (nội dung được đọc hoàn toàn), giọng nói tiếp tục (...today is sunny). Điều này có thể gặp phải trong nhiều tình huống, chẳng hạn như khi kẻ tấn công thử nhập ID của người sử dụng và phát hiện khuôn mặt, kẻ tấn công có thể tấn công việc phát hiện khuôn mặt bằng cách sử dụng bản ghi video trước đó của người sử dụng (người sử dụng bị tấn công), và kẻ tấn công đọc ID của người sử dụng để nhận dạng việc tấn công ID nội dung giọng nói. Như vậy, kẻ tấn công có thể bị xác thực. Tuy nhiên, trong cuộc tấn công này, tín hiệu môi và giọng nói thường không thống nhất, và có thể xác định được rằng người đọc không phải là người sử dụng.

Như được thể hiện trong bước 102, nếu xác định được rằng tín hiệu môi và giọng nói trong luồng audio và video cần được đăng ký là không thống nhất, lỗi đăng ký có thể được thông báo cho người sử dụng. Theo cách khác, như được thể hiện trên FIG. 1, đến bước 101 để thu thập lại luồng audio và video để ngăn lỗi xử lý.

Nếu không, nếu xác định được rằng tín hiệu môi và giọng nói trong luồng audio và video được đăng ký có tương ứng với nhau, thực hiện bước 103: sử dụng nội dung giọng nói thu được bằng cách thực hiện nhận dạng giọng nói đối với luồng audio trong luồng audio và video được thu thập, làm mã nhận dạng đối tượng của đối tượng đích. Việc nhận dạng giọng nói sử dụng các công nghệ máy tính để tự động nhận dạng nội dung giọng nói của người, nghĩa là quy trình nhận dạng chuyển đổi giọng nói thành nội dung. Ví dụ, sau khi nhận dạng giọng nói được thực hiện đối với phần audio "123456" được đọc bởi người sử dụng để thực hiện đăng ký, nội dung giọng nói thu được của luồng audio là "123456", và nội dung thu được thông qua việc nhận dạng có thể được sử dụng là mã nhận dạng của người sử dụng, nghĩa là ID của người sử dụng.

Việc nhận dạng giọng nói trước đó của luồng audio có thể được thực hiện sau khi xác định được rằng tín hiệu môi và giọng nói là tương ứng, để thu mã nhận dạng đối tượng, hoặc có thể được thực hiện trong quy trình xác định xem liệu tín hiệu môi và giọng nói có tương ứng hay không, để thu mã nhận dạng đối tượng.

Theo một khía cạnh khác, việc nhận dạng sinh lý được thực hiện đối với luồng audio và video cần được đăng ký, để thu đặc điểm sinh lý mẫu của luồng audio và video cần được đăng ký. Trong ví dụ này, đặc điểm sinh lý bao gồm đặc điểm khuôn mặt và đặc điểm dấu ấn giọng nói, nhưng không bị giới hạn ở hai đặc điểm này, giả sử rằng đặc điểm ví dụ có thể nhận dạng duy nhất người sử dụng, và có thể được sử dụng để phân biệt giữa các đặc điểm sinh lý của những người sử dụng khác nhau. Theo khía cạnh này, như được thể hiện trong bước 104, việc nhận dạng dấu ấn giọng nói có thể được thực hiện đối với luồng audio trong luồng audio và video cần được đăng ký, để thu đặc điểm dấu ấn giọng nói của đối tượng đích.

Theo một khía cạnh khác, việc phát hiện khuôn mặt được thực hiện đối với luồng video trong luồng audio và video cần được đăng ký, để thu đặc điểm khuôn mặt của đối tượng đích.

Trong thủ tục đăng ký, đặc điểm khuôn mặt được phát hiện có thể được xem là đặc điểm khuôn mặt mẫu và được sử dụng là tiêu chí trong quy trình xác thực tiếp theo. Tương tự, đặc điểm dấu ấn giọng nói được phát hiện cũng có thể được xem là đặc điểm dấu ấn giọng nói mẫu, và đặc điểm dấu ấn giọng nói mẫu và đặc điểm khuôn mặt mẫu có thể được gọi chung là đặc điểm sinh lý mẫu.

Trong phương án này, đặc điểm sinh lý mẫu và mã nhận dạng đối tượng của đối tượng đích cũng được xem là thông tin đăng ký đối tượng. Sau khi xác định được rằng dữ liệu trong thông tin đăng ký đối tượng hoàn tất, trong bước 106, mã nhận dạng đối tượng của đối tượng đích và đặc điểm sinh lý mẫu tương ứng được lưu trữ trong cơ sở dữ liệu làm thông tin đăng ký đối tượng.

Ngoài ra, thứ tự thực thi ba khía cạnh được thể hiện trên FIG. 1 không bị giới hạn. Ví dụ, sau khi thu luồng audio và video cần được đăng ký trong bước 101, ba khía cạnh này có thể được thực thi song song. Nếu tín hiệu môi và giọng nói là không thống nhất, thì đặc điểm dấu ấn giọng nói được nhận dạng và đặc điểm khuôn mặt được nhận dạng có thể không được lưu trữ. Theo cách khác, sau khi xác định được rằng tín hiệu môi và giọng nói là tương ứng, thì việc phát hiện và nhận dạng sau đó được thực hiện đối với đặc điểm dấu ấn giọng nói và đặc điểm khuôn mặt.

FIG. 2 minh họa thủ tục xác định tính thống nhất giữa tín hiệu môi và giọng nói trên FIG. 1. Thủ tục này có thể bao gồm các bước sau:

Bước 201: Thực hiện phát hiện điểm cuối dựa vào luồng audio trong luồng audio và video cần được đăng ký. Trong bước này, thời điểm bắt đầu và thời điểm kết thúc của luồng audio có thể được phát hiện trong luồng audio liên tục.

Bước 202: Trích xuất các đặc điểm giọng nói liên tục dựa vào luồng audio, trong đó các đặc điểm bao gồm nhưng không giới hạn ở đặc điểm MFCC và đặc điểm LPCC. Các đặc điểm được trích xuất trong bước này có thể được sử dụng để nhận dạng giọng nói.

Bước 203: Xác định âm tiết giọng nói và thời điểm tương ứng trong luồng audio. Trong bước này, mỗi âm tiết trong luồng audio có thể được xác định dựa vào các đặc điểm giọng nói được trích xuất trong bước 202, và thời điểm xuất hiện tương ứng và thời điểm biến mất tương ứng của âm tiết có thể được xác định. Phương pháp nhận dạng giọng nói bao gồm nhưng không bị giới hạn ở các phương pháp như mẫu Markov ẩn (HMM), mạng neuron sâu (DNN), và mẫu bộ nhớ dài-ngắn (LSTM).

Bước 204: Phát hiện vị trí của môi dựa vào luồng video trong luồng audio và video cần được đăng ký. Trong bước này, các vị trí của môi có thể được phát hiện từ hình ảnh video.

Bước 205: Xác định chất lượng của hình ảnh môi được phát hiện. Ví dụ, các thông số chẳng hạn như độ nét và mức độ bộc lộ vị trí của môi có thể được xác định. Nếu độ nét thấp hoặc mức độ bộc lộ vị trí là quá cao, thì xác định được rằng hình ảnh không đạt chất lượng; và trong trường hợp này, thu thập lại luồng audio và video cần được đăng ký. Nếu hình ảnh đạt chất lượng, tiến đến bước 206 để thực hiện nhận dạng tín hiệu môi.

Bước 206: Trích xuất các đặc điểm liên tục của môi. Trong bước này, một đặc điểm có thể được trích xuất từ các hình ảnh môi liên tục, và đặc điểm bao gồm nhưng không bị giới hạn ở thông tin mô tả hình ảnh cục bộ như điểm ảnh, LBP, Gabor, SIFT, hoặc Surf.

Bước 207: Xác định âm tiết tín hiệu môi và thời điểm tương ứng trong luồng video. Trong bước này, âm tiết tín hiệu môi có thể được xác định bằng cách sử dụng các phương pháp chẳng hạn như mẫu Markov ẩn (HMM), và mẫu bộ nhớ dài-ngắn (LSTM). Thời điểm tương ứng với âm tiết tín hiệu môi trong chuỗi thời gian của video cũng được xác định bằng cách sử dụng mẫu trong quá trình nhận dạng tín hiệu môi.

Bước 208: Xác định xem liệu cả âm tiết tín hiệu môi và âm tiết giọng nói có tương ứng với thời điểm tương ứng hay không. Ví dụ, trong bước này, thông tin thời điểm của âm tiết của giọng nói có thể được so sánh với thông tin thời điểm của âm tiết tín hiệu môi. Nếu kết quả so sánh chỉ ra tương ứng, thì có thể coi luồng audio được tạo ra bởi người thực; trong trường hợp này, tiến đến bước 209. Nếu kết quả so sánh chỉ ra không tương ứng, thì đoán được rằng luồng audio là hành vi tấn công; và trong trường hợp này, quay trở lại thủ tục đăng ký. Trong phương án này, phương pháp phát hiện tính không tương ứng giữa cả âm tiết tín hiệu môi và âm tiết giọng nói và thời điểm tương ứng được thực hiện chi tiết hơn, và do đó tạo ra độ chính xác cao hơn trong việc xác định giọng nói của người thực.

Bước 209: Thực hiện nhận dạng giọng nói đối với các đặc điểm giọng nói được trích xuất trong bước 202, để thu ID của người sử dụng, nghĩa là mã nhận dạng đối tượng. Phương pháp nhận dạng giọng nói bao gồm nhưng không bị giới hạn ở các phương pháp chẳng hạn như mẫu Markov ẩn (HMM), mạng neuron sâu (DNN), và mẫu bộ nhớ dài-ngắn (LSTM).

Ngoài ra, trong ví dụ được thể hiện trên FIG. 2, việc nhận dạng giọng nói của luồng audio có thể được thực hiện trong bước 209 sau khi xác định được rằng tín hiệu môi và giọng nói là tương ứng. Theo cách khác, khi thời điểm âm tiết trong luồng audio được xác định trong bước 203, ID của người sử dụng thu được bằng cách thực hiện nhận dạng giọng nói dựa vào đặc điểm giọng nói. Trong trường hợp này, sau khi xác định được rằng tín hiệu môi và giọng nói là tương ứng trong bước 208, ID của người sử dụng thu được thông qua việc nhận dạng có thể được sử dụng trực tiếp làm mã nhận dạng đối tượng.

FIG. 3 minh họa thủ tục nhận dạng đặc điểm khuôn mặt trên FIG. 1. Thủ tục này có thể bao gồm các bước sau:

Bước 301: Phát hiện các hình ảnh khuôn mặt dựa vào luồng video trong luồng audio và video cần được đăng ký. Trong bước này, hình ảnh khung video có thể được trích xuất từ luồng video trong luồng audio và video, và có thể phát hiện xem liệu khuôn mặt có xuất hiện trong hình ảnh khung video hay không. Nếu có, tiến đến bước 302. Nếu không, quay trở lại thủ tục xác định.

Bước 302: Phát hiện chất lượng của các hình ảnh khuôn mặt. Trong bước này,

việc phát hiện điểm của đặc điểm khuôn mặt có thể được thực hiện trên khuôn mặt được phát hiện trong bước 301, và các góc của khuôn mặt theo cả hai hướng ngang và dọc có thể được xác định dựa vào kết quả của việc phát hiện điểm của đặc điểm này. Nếu cả hai góc không vượt quá các góc nghiêng nhất định, yêu cầu chất lượng được thỏa mãn. Nếu không, thì yêu cầu chất lượng không được thỏa mãn. Ngoài ra, độ nét, mức độ bộc lộ, v.v., của vùng khuôn mặt được xác định, cũng cần nằm trong ngưỡng nhất định. Nếu các hình ảnh khuôn mặt đạt chất lượng tốt, đặc điểm khuôn mặt có thể được nhận dạng tốt hơn.

Bước 303: Đối với các hình ảnh khuôn mặt thỏa mãn yêu cầu chất lượng, trích xuất vector đặc trưng từ các hình ảnh khuôn mặt, trong đó vector đặc trưng bao gồm nhưng không giới hạn ở mẫu nhị phân cục bộ (LBP), Gabor, mạng nơ-ron tích chập, v.v..

Bước 304: Kết hợp nhiều vector đặc trưng khuôn mặt được trích xuất trong bước 303, để tạo thành đặc điểm khuôn mặt duy nhất của người sử dụng, nghĩa là đặc điểm khuôn mặt mẫu.

FIG. 4 minh họa thủ tục nhận dạng đặc điểm dấu ấn giọng nói trên FIG. 1. Thủ tục này có thể bao gồm các bước sau.

Bước 401: Thu luồng audio trong luồng audio và video cần được đăng ký.

Trong ví dụ này, việc nhận dạng đặc điểm dấu ấn giọng nói có thể được thực hiện dựa vào luồng audio trong luồng audio và video cần được đăng ký.

Bước 402: Xác định rằng chất lượng audio của luồng audio thỏa mãn tiêu chí chất lượng.

Trong bước này, chất lượng audio có thể được xác định. Chất lượng của luồng audio được thu thập càng tốt dẫn đến hiệu quả thực hiện nhận dạng dấu ấn giọng nói đối với audio càng tốt. Do đó, trước khi việc nhận dạng dấu ấn giọng nói tiếp theo được thực hiện, chất lượng của luồng audio có thể được xác định trước tiên. Ví dụ, thông tin chẳng hạn như cường độ tín hiệu và tỷ lệ tín hiệu-nhiều của giọng nói trong luồng audio có thể được tính để xác định xem liệu giọng nói thỏa mãn tiêu chí chất lượng hay không. Ví dụ, tiêu chí chất lượng có thể là tỷ lệ tín hiệu-nhiều nằm trong phạm vi nhất định, hoặc có thể là cường độ tín hiệu của giọng nói lớn hơn ngưỡng

cường độ. Nếu luồng audio đạt chất lượng, tiến đến bước 403. Nếu không, thu thập lại luồng audio và video cần được đăng ký.

Bước 403: Trích xuất vector đặc trưng dấu ấn giọng nói từ luồng audio.

Trong ví dụ này, có thể có nhiều luồng audio và video cần được đăng ký. Ví dụ, người sử dụng có thể đọc ID của người sử dụng hai lần, và hai luồng audio và video tương ứng được thu thập. Trong bước này, vector đặc trưng dấu ấn giọng nói của luồng audio trong mỗi luồng audio và video có thể được trích xuất. Vector đặc trưng có thể được trích xuất theo nhiều các phương pháp, và chi tiết được bỏ qua cho đơn giản. Ví dụ, hệ số cepstral theo tần số Mel (MFCC) của tham số đặc điểm giọng nói có thể được trích xuất từ tín hiệu giọng nói của luồng audio, và sau đó vector đặc trưng được tính bằng cách sử dụng phương pháp chẳng hạn như i-vector (thuật toán nhận dạng người nói) hoặc hoặc phân tích phân biệt tuyến tính theo xác suất (PLDA, nghĩa là, thuật toán bù kênh để nhận dạng dấu ấn giọng nói).

Bước 404: Xác định xem liệu vector đặc trưng dấu ấn giọng nói của nhiều luồng audio có tương ứng với nhau hay không.

Ví dụ, khi người sử dụng đọc ID của người sử dụng ít nhất hai lần trong quá trình đăng ký, thì có ít nhất hai luồng audio tương ứng được thu thập. Để đảm bảo rằng sự chênh lệch giữa đặc điểm dấu ấn giọng nói của nhiều luồng audio là không quá lớn, việc xác định tính tương thích dấu ấn giọng nói có thể được thực hiện giữa nhiều luồng audio. Ví dụ, điểm số tương tự giữa nhiều luồng audio có thể được tính dựa vào vector đặc trưng dấu ấn giọng nói được trích xuất từ mỗi luồng audio trong bước 403.

Nếu điểm số tương tự nằm trong ngưỡng điểm số nhất định, thì chỉ ra rằng các luồng audio thỏa mãn yêu cầu về tính tương tự; và trong trường hợp này, tiến đến bước 405. Nếu không, thì chỉ ra rằng có sự chênh lệch lớn giữa nhiều audio được nhập vào bởi người sử dụng, và người sử dụng thực hiện đăng ký có thể được lệnh đọc lại ID của người sử dụng, nghĩa là, để thu thập lại luồng audio.

Bước 405: Tạo ra đặc điểm dấu ấn giọng nói mẫu dựa vào vector đặc trưng dấu ấn giọng nói của nhiều luồng audio.

Trong bước này, tổng trọng số có thể được thực hiện trên vector đặc trưng dấu ấn giọng nói được trích xuất từ các luồng audio trong bước trước đó, để thu đặc điểm

dấu ấn giọng nói mẫu.

Sau khi thủ tục đăng ký trước đó hoàn thành, thông tin đăng ký đối tượng của đối tượng đích đã được lưu trữ trong cơ sở dữ liệu. Thông tin đăng ký đối tượng có thể bao gồm mã nhận dạng đối tượng và đặc điểm sinh lý mẫu tương ứng. Đặc điểm sinh lý mẫu có thể bao gồm đặc điểm dấu ấn giọng nói mẫu và đặc điểm khuôn mặt mẫu, và quy trình xác thực căn cước của đối tượng có thể được thực hiện sau đây dựa vào thông tin đăng ký đối tượng.

Xác thực căn cước

FIG. 5 minh họa thủ tục xác thực căn cước ví dụ. Trong thủ tục này, đặc điểm sinh lý được sử dụng để xác thực được mô tả bằng cách sử dụng sự kết hợp của đặc điểm khuôn mặt và đặc điểm dấu ấn giọng nói làm ví dụ. Ngoài ra, các đặc điểm sinh lý có thể được so sánh sau khi xác định được rằng đối tượng đích được xác thực là đối tượng thực thay vì video. Như được thể hiện trên FIG. 5, thủ tục xác thực bao gồm các bước sau:

Bước 501: Thu luồng audio và video đã thu thập, trong đó luồng audio và video được tạo ra bởi đối tượng đích cần được xác thực.

Ví dụ, người sử dụng chỉ có thể đăng nhập vào ứng dụng với yêu cầu bảo mật tương đối cao sau khi căn cước của người sử dụng đã được xác thực bởi ứng dụng. Trong bước này, người sử dụng có thể bắt đầu ứng dụng trên thiết bị thông minh, ví dụ, điện thoại thông minh của người sử dụng, và người sử dụng có thể thu thập luồng audio và video cần được xác thực bằng cách sử dụng camera và micrô của điện thoại thông minh. Luồng audio và video có thể được tạo ra khi người sử dụng đọc ID ứng dụng của người sử dụng.

Bước 502: Xác định xem liệu tín hiệu môi và giọng nói trong luồng audio và video có tương ứng với nhau không.

Trong ví dụ này, trước tiên có thể xác định xem liệu tín hiệu môi và giọng nói trong luồng audio và video có tương ứng với nhau không. Đối với thủ tục xác định tính tương thích cụ thể, có thể tham chiếu đến FIG. 2, và chi tiết được bỏ qua cho đơn giản.

Nếu tín hiệu môi và giọng nói là tương ứng, thì chỉ ra rằng đối tượng đích được xác thực là đối tượng thực thay vì video, v.v.. Trong trường hợp này, tiến đến bước 503.

Nếu không, quay trở lại bước 501 để thực hiện thu thập lại.

Bước 503: Thực hiện nhận dạng giọng nói đối với luồng audio trong luồng audio và video để thu được nội dung giọng nói của luồng audio. Ví dụ, nội dung giọng nói thu được thông qua việc nhận dạng có thể là ID của người sử dụng "123456".

Bước 504: Sử dụng nội dung giọng nói làm mã nhận dạng đối tượng của đối tượng đích, và xác định xem liệu thông tin đăng ký đối tượng được lưu trữ từ trước có bao gồm mã nhận dạng đối tượng hay không.

Ví dụ, nếu thông tin đăng ký đối tượng được lưu trữ từ trước bao gồm mã nhận dạng đối tượng, đặc điểm sinh lý mẫu tương ứng với mã nhận dạng đối tượng có thể thu được trong thông tin đăng ký đối tượng, ví dụ, đặc điểm khuôn mặt mẫu và đặc điểm dấu ấn giọng nói mẫu. Việc nhận dạng sinh lý còn được thực hiện đối với luồng audio và video cần được xác thực, để thu đặc điểm sinh lý của đối tượng đích, để so sánh đặc điểm sinh lý với đặc điểm sinh lý mẫu. Nếu thông tin đăng ký đối tượng được lưu trữ từ trước không bao gồm mã nhận dạng đối tượng, người sử dụng có thể được thông báo rằng người sử dụng đã không được thực hiện đăng ký.

Bước 505: Thực hiện nhận dạng dấu ấn giọng nói đối với luồng audio và video để thu đặc điểm dấu ấn giọng nói của đối tượng đích. Đối với việc trích xuất đặc điểm dấu ấn giọng nói trong bước này, có thể tham chiếu đến FIG. 4.

Bước 506: Thực hiện nhận dạng khuôn mặt đối với luồng audio và video để thu đặc điểm khuôn mặt của đối tượng đích.

Sau đó đặc điểm sinh lý của đối tượng đích có thể được so sánh với đặc điểm sinh lý mẫu để thu kết quả so sánh, và nếu kết quả so sánh thỏa mãn điều kiện xác thực, xác định được rằng đối tượng đích đã được xác thực. Ví dụ, bước 507 đến bước 509 được đưa vào.

Bước 507: So sánh đặc điểm dấu ấn giọng nói của đối tượng đích với đặc điểm dấu ấn giọng nói mẫu để thu điểm số so sánh dấu ấn giọng nói.

Bước 508: So sánh đặc điểm khuôn mặt của đối tượng đích với đặc điểm khuôn mặt mẫu để thu điểm số so sánh khuôn mặt.

Bước 509: Xác định xem liệu điểm số so sánh dấu ấn giọng nói và điểm số so khớp khuôn mặt có thỏa mãn điều kiện xác thực hay không.

Ví dụ, xác định được rằng đối tượng đích đã được xác thực nếu điểm số so sánh dấu ấn giọng nói và điểm số so sánh khuôn mặt thỏa mãn ít nhất một trong số: điểm số so sánh dấu ấn giọng nói lớn hơn ngưỡng điểm số dấu ấn giọng nói, và điểm số so sánh khuôn mặt lớn hơn ngưỡng điểm số khuôn mặt; hoặc tích của điểm số so sánh dấu ấn giọng nói và điểm số so sánh khuôn mặt lớn hơn ngưỡng tích số tương ứng; hoặc tổng trọng số của điểm số so sánh dấu ấn giọng nói và điểm số so sánh khuôn mặt lớn hơn ngưỡng trọng số tương ứng.

Nếu xác định được rằng điểm số so sánh dấu ấn giọng nói và điểm số so sánh khuôn mặt thỏa mãn điều kiện xác thực trong bước này, xác định được rằng đối tượng đích đã được xác thực. Nếu không, xác định được rằng đối tượng đích bị lỗi xác thực.

Ngoài ra, trong ví dụ xác thực căn cước này, tương tự với thủ tục đăng ký căn cước trước đó, việc nhận dạng giọng nói có thể được thực hiện đối với luồng audio để thu ID của người sử dụng sau khi xác định được rằng tín hiệu môi và giọng nói là tương ứng, hoặc ID của người sử dụng có thể thu được khi thời điểm âm tiết trong luồng audio được xác định. Trong ví dụ trước đó, ID của người sử dụng được xác định sau khi xác định được rằng tín hiệu môi và giọng nói là tương ứng.

Theo phương pháp xác thực căn cước trong phương án này của sáng chế, luồng audio và video cần được tạo ra chỉ một lần trong khi xác thực người sử dụng. Ví dụ, người sử dụng cần đọc ID của người sử dụng chỉ một lần. Trong phương pháp này, ID của người sử dụng có thể thu được bằng cách thực hiện nhận dạng giọng nói đối với luồng audio và video, và đặc điểm khuôn mặt và đặc điểm dấu ấn giọng nói cũng có thể được xác minh bằng cách sử dụng cùng luồng audio và video. Nó giúp đơn giản hóa thao tác của người sử dụng, tăng cường hiệu quả xác thực, duy trì mẫu xác thực một-một, đảm bảo nhận dạng chính xác. Nói cách khác, đặc điểm sinh lý được nhận dạng chỉ được so sánh với đặc điểm tương ứng với mã nhận dạng đối tượng trong cơ sở dữ liệu, nhờ đó đảm bảo tính chính xác nhận dạng. Ngoài ra, trong phương pháp này, sự thống nhất giữa tín hiệu môi và giọng nói được xác định để đảm bảo rằng đối tượng đích là đối tượng thực thay vì đoạn ghi video giả của kẻ tấn công, nhờ đó tăng cường tính bảo mật và độ tin cậy xác thực. Trong phương pháp này, ID của người sử dụng và đặc điểm sinh lý thu được thông qua việc nhận dạng thu được dựa vào cùng luồng audio và video. Ở một mức độ nhất định, luồng audio và video giả của kẻ tấn công có thể được nhận dạng.

Để thực hiện phương pháp xác thực căn cước, một phương án của sáng chế còn đề xuất thiết bị xác thực căn cước. Như được thể hiện trên FIG. 6, thiết bị có thể bao gồm môđun thu thông tin 61, môđun xác định mã nhận dạng 62, môđun quản lý thông tin 63, môđun nhận dạng đặc điểm 64, và môđun xử lý xác thực 65.

Môđun thu thông tin 61 được tạo cấu hình để thu luồng audio và video đã thu thập, trong đó luồng audio và video được tạo ra bởi đối tượng đích cần được xác thực.

Môđun xác định mã nhận dạng 62 được tạo cấu hình để xác định xem liệu tín hiệu môi và giọng nói trong luồng audio và video có tương ứng với nhau không, và nếu tín hiệu môi và giọng nói tương ứng với nhau, thì sử dụng nội dung giọng nói thu được bằng cách thực hiện nhận dạng giọng nói đối với luồng audio trong luồng audio và video làm mã nhận dạng đối tượng của đối tượng đích.

Môđun quản lý thông tin 63 được tạo cấu hình để thu đặc điểm sinh lý mẫu tương ứng với mã nhận dạng đối tượng từ thông tin đăng ký đối tượng, nếu thông tin đăng ký đối tượng được lưu trữ từ trước bao gồm mã nhận dạng đối tượng.

Môđun nhận dạng đặc điểm 64 được tạo cấu hình để thực hiện nhận dạng sinh lý đối với luồng audio và video để thu đặc điểm sinh lý của đối tượng đích.

Môđun xử lý xác thực 65 được tạo cấu hình để so sánh đặc điểm sinh lý của đối tượng đích với đặc điểm sinh lý mẫu để thu kết quả so sánh, và nếu kết quả so sánh thỏa mãn điều kiện xác thực thì xác định rằng đối tượng đích đã được xác thực.

Dựa vào FIG. 7, theo một ví dụ, môđun nhận dạng đặc điểm 64 có thể bao gồm môđun con nhận dạng dấu ấn giọng nói 641 và môđun con nhận dạng khuôn mặt 642.

Môđun con nhận dạng dấu ấn giọng nói 641 được tạo cấu hình để thực hiện nhận dạng dấu ấn giọng nói đối với luồng audio và video để thu đặc điểm dấu ấn giọng nói của đối tượng đích.

Môđun con nhận dạng khuôn mặt 642 được tạo cấu hình để thực hiện nhận dạng khuôn mặt đối với luồng audio và video để thu đặc điểm khuôn mặt của đối tượng đích.

Môđun xử lý xác thực 65 được tạo cấu hình để so sánh đặc điểm dấu ấn giọng nói của đối tượng đích với đặc điểm dấu ấn giọng nói mẫu để thu điểm số so sánh dấu ấn giọng nói, và so sánh đặc điểm khuôn mặt của đối tượng đích với đặc điểm khuôn

mặt mẫu để thu điểm số so sánh khuôn mặt; và nếu điểm số so sánh dấu ấn giọng nói và điểm số so sánh khuôn mặt thỏa mãn điều kiện xác thực thì xác định rằng đối tượng đích đã được xác thực.

Theo một ví dụ, xác định được rằng đối tượng đích đã được xác thực nếu điểm số so sánh dấu ấn giọng nói và điểm số so sánh khuôn mặt thỏa mãn ít nhất một trong số: điểm số so sánh dấu ấn giọng nói lớn hơn ngưỡng điểm số dấu ấn giọng nói, và điểm số so sánh khuôn mặt lớn hơn ngưỡng điểm số khuôn mặt; hoặc tích của điểm số so sánh dấu ấn giọng nói và điểm số so sánh khuôn mặt lớn hơn ngưỡng tích số tương ứng; hoặc tổng trọng số của điểm số so sánh dấu ấn giọng nói và điểm số so sánh khuôn mặt lớn hơn ngưỡng trọng số tương ứng.

Theo một ví dụ, như được thể hiện trên FIG. 7, môđun xác định mã nhận dạng 62 có thể bao gồm: môđun con nhận dạng âm tiết 621, được tạo cấu hình để xác định âm tiết giọng nói và thời điểm tương ứng trong luồng audio trong luồng audio và video, và xác định âm tiết tín hiệu môi và thời điểm tương ứng trong luồng video trong luồng audio và video; và môđun con xác định tính tương ứng 622, được tạo cấu hình để xác định rằng tín hiệu môi và giọng nói là tương ứng nếu cả âm tiết giọng nói và âm tiết tín hiệu môi tương ứng với nhau tại thời điểm tương ứng.

Theo một ví dụ, môđun thu thông tin 61 còn được tạo cấu hình để thu luồng audio và video cần được đăng ký của đối tượng đích.

Môđun xác định mã nhận dạng 62 còn được tạo cấu hình để sử dụng nội dung giọng nói thu được bằng cách thực hiện nhận dạng giọng nói đối với luồng audio trong luồng audio và video làm mã nhận dạng đối tượng của đối tượng đích, khi tín hiệu môi và giọng nói trong luồng audio và video được đăng ký tương ứng với nhau.

Môđun nhận dạng đặc điểm 64 còn được tạo cấu hình để thực hiện nhận dạng sinh lý đối với luồng audio và video cần được đăng ký, để thu đặc điểm sinh lý mẫu của luồng audio và video cần được đăng ký.

Môđun quản lý thông tin 63 còn được tạo cấu hình để lưu trữ tương ứng mã nhận dạng đối tượng của đối tượng đích và đặc điểm sinh lý mẫu tương ứng trong thông tin đăng ký đối tượng.

Các phần mô tả trên chỉ là các phương án ví dụ của sáng chế, và không nhằm

giới hạn sáng chế. Tất cả các cải biến, các thay thế tương đương, và cải tiến, v.v., được thực hiện trong mục đích và nguyên tắc của sáng chế sẽ nằm trong phạm vi bảo hộ của sáng chế.

YÊU CẦU BẢO HỘ

1. Phương pháp xác thực căn cước, trong đó phương pháp này bao gồm các bước:

thu luồng audio và video, trong đó luồng audio và video được tạo ra bởi đối tượng đích cần được xác thực (501);

xác định xem liệu tín hiệu môi và giọng nói trong luồng audio và video có tương ứng với nhau hay không (502),

để đáp lại việc xác định rằng tín hiệu môi và giọng nói tương ứng với nhau, sử dụng nội dung giọng nói thu được bằng cách thực hiện nhận dạng giọng nói (503) đối với luồng audio trong luồng audio và video làm mã nhận dạng đối tượng của đối tượng đích (504);

xác định xem liệu thông tin đăng ký đối tượng được lưu trữ từ trước có bao gồm mã nhận dạng đối tượng hay không;

để đáp lại việc xác định rằng thông tin đăng ký đối tượng được lưu trữ từ trước bao gồm mã nhận dạng đối tượng, thu đặc điểm sinh lý mẫu tương ứng với mã nhận dạng đối tượng từ thông tin đăng ký đối tượng được lưu trữ từ trước;

thực hiện nhận dạng sinh lý đối với luồng audio và video để thu đặc điểm sinh lý của đối tượng đích (505, 506);

so sánh đặc điểm sinh lý của đối tượng đích với đặc điểm sinh lý mẫu để thu kết quả so sánh (507, 508);

xác định xem liệu kết quả so sánh có thỏa mãn điều kiện xác thực hay không; và

để đáp lại việc xác định rằng kết quả so sánh thỏa mãn điều kiện xác thực, xác định rằng đối tượng đích đã được xác thực.

2. Phương pháp theo điểm 1, trong đó bước xác định xem liệu tín hiệu môi và giọng nói trong luồng audio và video có tương ứng với nhau hay không, bước sử dụng nội dung giọng nói thu được bằng cách thực hiện nhận dạng giọng nói đối với luồng audio trong luồng audio và video làm mã nhận dạng đối tượng của đối tượng đích, và bước thu đặc điểm sinh lý mẫu tương ứng với mã nhận dạng đối tượng từ thông tin đăng ký đối tượng được thực thi song song.

3. Phương pháp theo điểm 1 hoặc 2, trong đó đặc điểm sinh lý bao gồm đặc điểm dấu ấn giọng nói và đặc điểm khuôn mặt, và đặc điểm sinh lý mẫu bao gồm đặc điểm khuôn mặt mẫu và đặc điểm dấu ấn giọng nói mẫu.

4. Phương pháp theo điểm 3, trong đó bước thực hiện nhận dạng sinh lý đối với luồng audio và video để thu đặc điểm sinh lý của đối tượng đích bao gồm các bước:

thực hiện nhận dạng dấu ấn giọng nói đối với luồng audio và video để thu đặc điểm dấu ấn giọng nói của đối tượng đích; và

thực hiện nhận dạng khuôn mặt đối với luồng audio và video để thu đặc điểm khuôn mặt của đối tượng đích.

5. Phương pháp theo điểm 4, trong đó bước so sánh đặc điểm sinh lý của đối tượng đích với đặc điểm sinh lý mẫu để thu kết quả so sánh, và nếu kết quả so sánh thỏa mãn điều kiện xác thực, thì xác định rằng đối tượng đích đã được xác thực, trong đó phương pháp này bao gồm các bước:

so sánh đặc điểm dấu ấn giọng nói của đối tượng đích với đặc điểm dấu ấn giọng nói mẫu để thu điểm số so sánh dấu ấn giọng nói;

so sánh đặc điểm khuôn mặt của đối tượng đích với đặc điểm khuôn mặt mẫu để thu điểm số so sánh khuôn mặt;

xác định xem liệu điểm số so sánh dấu ấn giọng nói và điểm số so sánh khuôn mặt có thỏa mãn điều kiện xác thực hay không; và

để đáp lại việc xác định rằng điểm số so sánh dấu ấn giọng nói và điểm số so sánh khuôn mặt thỏa mãn điều kiện xác thực, xác định rằng đối tượng đích đã được xác thực.

6. Phương pháp theo theo điểm bất kỳ trong số các điểm từ 3 đến 5, trong đó xác định được rằng đối tượng đích đã được xác thực nếu điểm số so sánh dấu ấn giọng nói và điểm số so sánh khuôn mặt thỏa mãn ít nhất một trong số:

điểm số so sánh dấu ấn giọng nói lớn hơn ngưỡng điểm số dấu ấn giọng nói, và điểm số so sánh khuôn mặt lớn hơn ngưỡng điểm số khuôn mặt; hoặc

tích của điểm số so sánh dấu ấn giọng nói và điểm số so sánh khuôn mặt lớn hơn ngưỡng tích số tương ứng; hoặc

tổng trọng số của điểm số so sánh dấu ấn giọng nói và điểm số so sánh khuôn mặt lớn hơn ngưỡng trọng số tương ứng.

7. Phương pháp theo điểm 1 hoặc 2, trong đó bước xác định xem liệu tín hiệu môi và giọng nói trong luồng audio và video có tương ứng với nhau hay không, và nếu tín hiệu môi và giọng nói tương ứng với nhau, sử dụng nội dung giọng nói thu được bằng cách thực hiện nhận dạng giọng nói đối với luồng audio trong luồng audio và video làm mã nhận dạng đối tượng của đối tượng đích bao gồm:

xác định âm tiết giọng nói và thời điểm tương ứng trong luồng audio trong luồng audio và video;

xác định âm tiết tín hiệu môi và thời điểm tương ứng trong luồng video trong luồng audio và video; và

xác định rằng tín hiệu môi và giọng nói là tương ứng nếu cả âm tiết giọng nói và âm tiết tín hiệu môi tương ứng với nhau tại thời điểm tương ứng.

8. Phương pháp theo điểm 1 hoặc 2, trong đó, trước khi thu luồng audio và video đã thu thập, phương pháp này còn bao gồm các bước:

thu luồng audio và video cần được đăng ký của đối tượng đích;

để đáp lại việc xác định rằng tín hiệu môi và giọng nói trong luồng audio và video được đăng ký tương ứng với nhau, sử dụng nội dung giọng nói thu được bằng cách thực hiện nhận dạng giọng nói đối với luồng audio trong luồng audio và video làm mã nhận dạng đối tượng của đối tượng đích;

thực hiện nhận dạng sinh lý đối với luồng audio và video cần được đăng ký, để thu đặc điểm sinh lý mẫu của luồng audio và video cần được đăng ký; và

lưu trữ mã nhận dạng đối tượng của đối tượng đích và đặc điểm sinh lý mẫu tương ứng trong thông tin đăng ký đối tượng.

9. Phương pháp theo điểm 1 hoặc 2, trong đó đối tượng đích cần được xác thực bao gồm người sử dụng của thiết bị người sử dụng, luồng audio bao gồm giọng đọc của người sử dụng, luồng video bao gồm hình ảnh video đọc của người sử dụng, và luồng audio và video được thu thập bao gồm luồng audio và video được tạo ra khi người sử dụng đăng ký ứng dụng.

10. Thiết bị xác thực căn cước, trong đó thiết bị này bao gồm nhiều mô đun được tạo cấu hình để thực hiện các hoạt động theo phương pháp theo điểm bất kỳ trong số các điểm từ 1 đến 9.

1/7

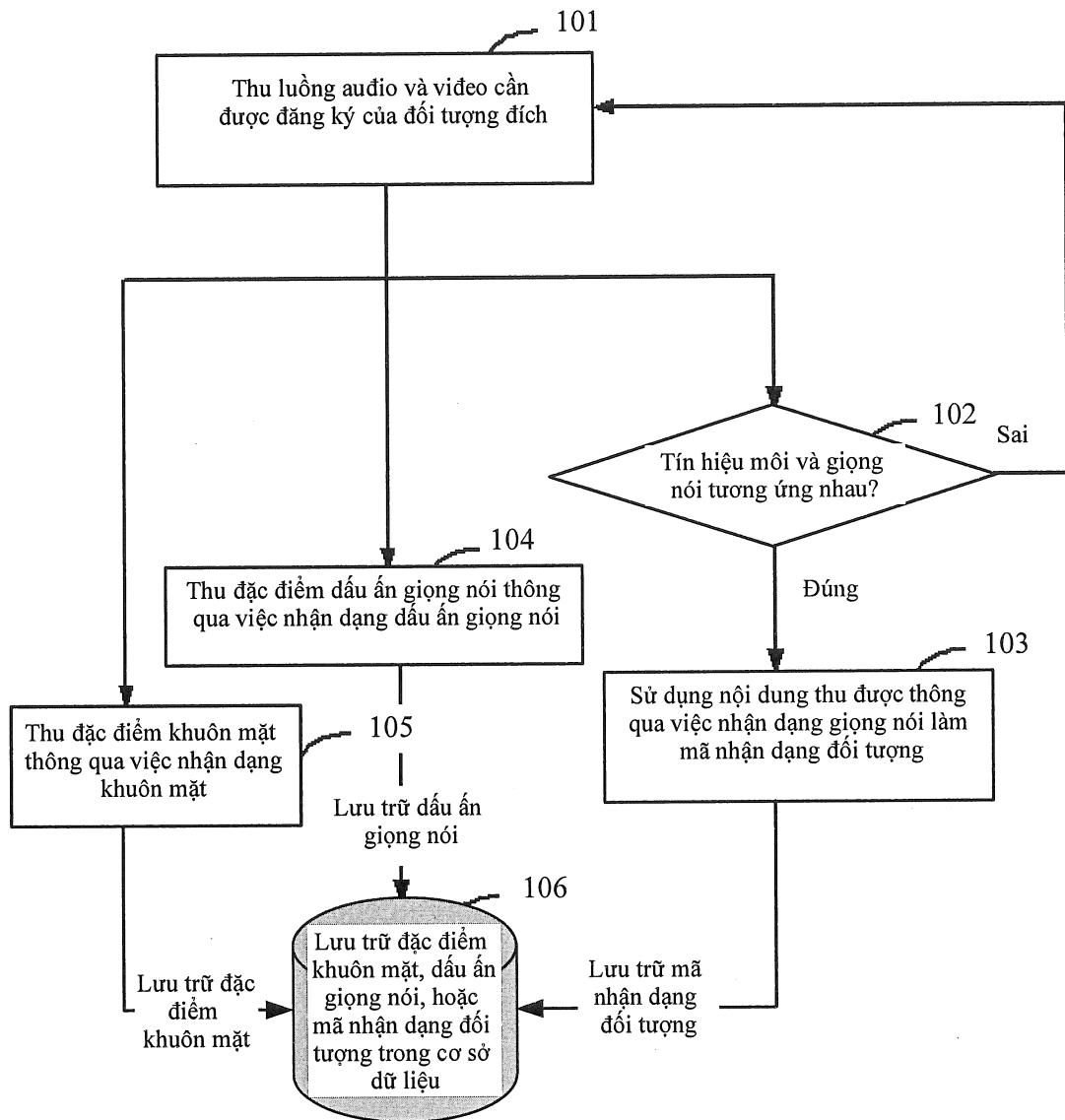


FIG. 1

2/7

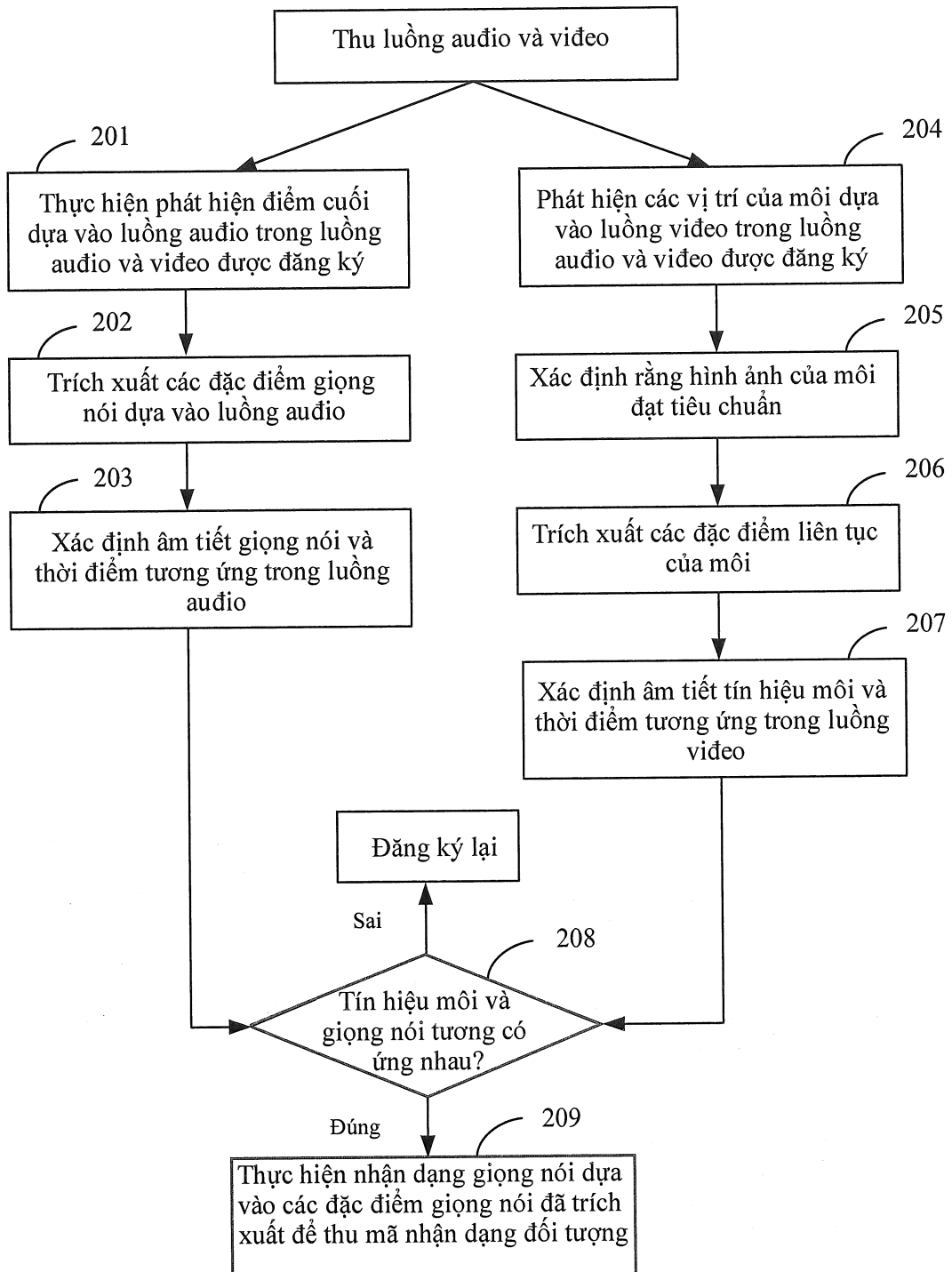


FIG. 2

3/7

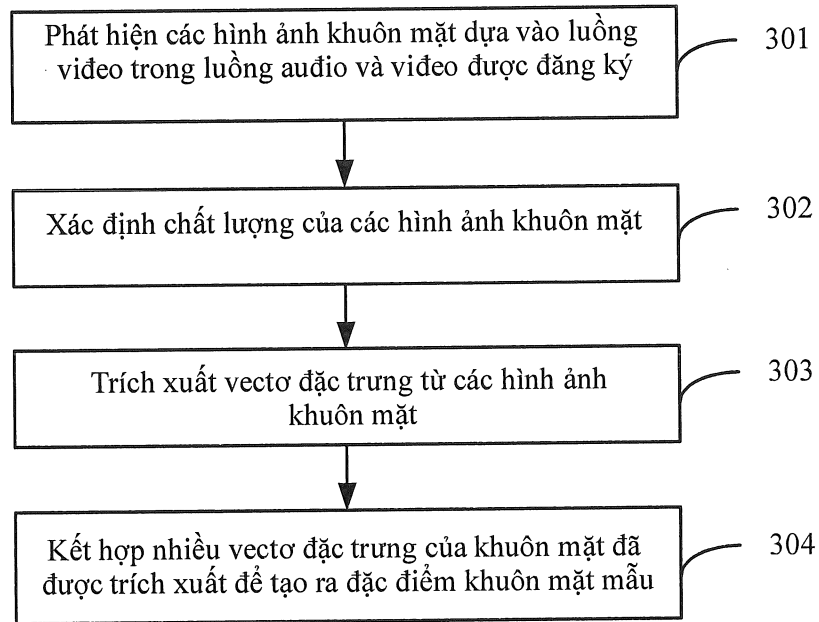


FIG. 3

4/7

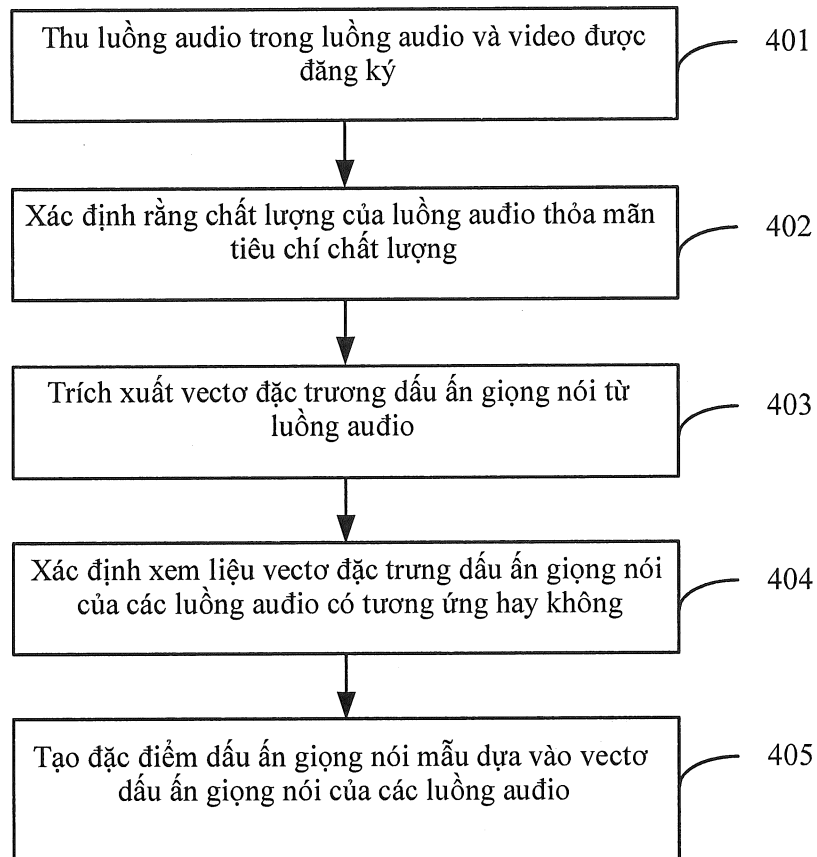


FIG. 4

5/7

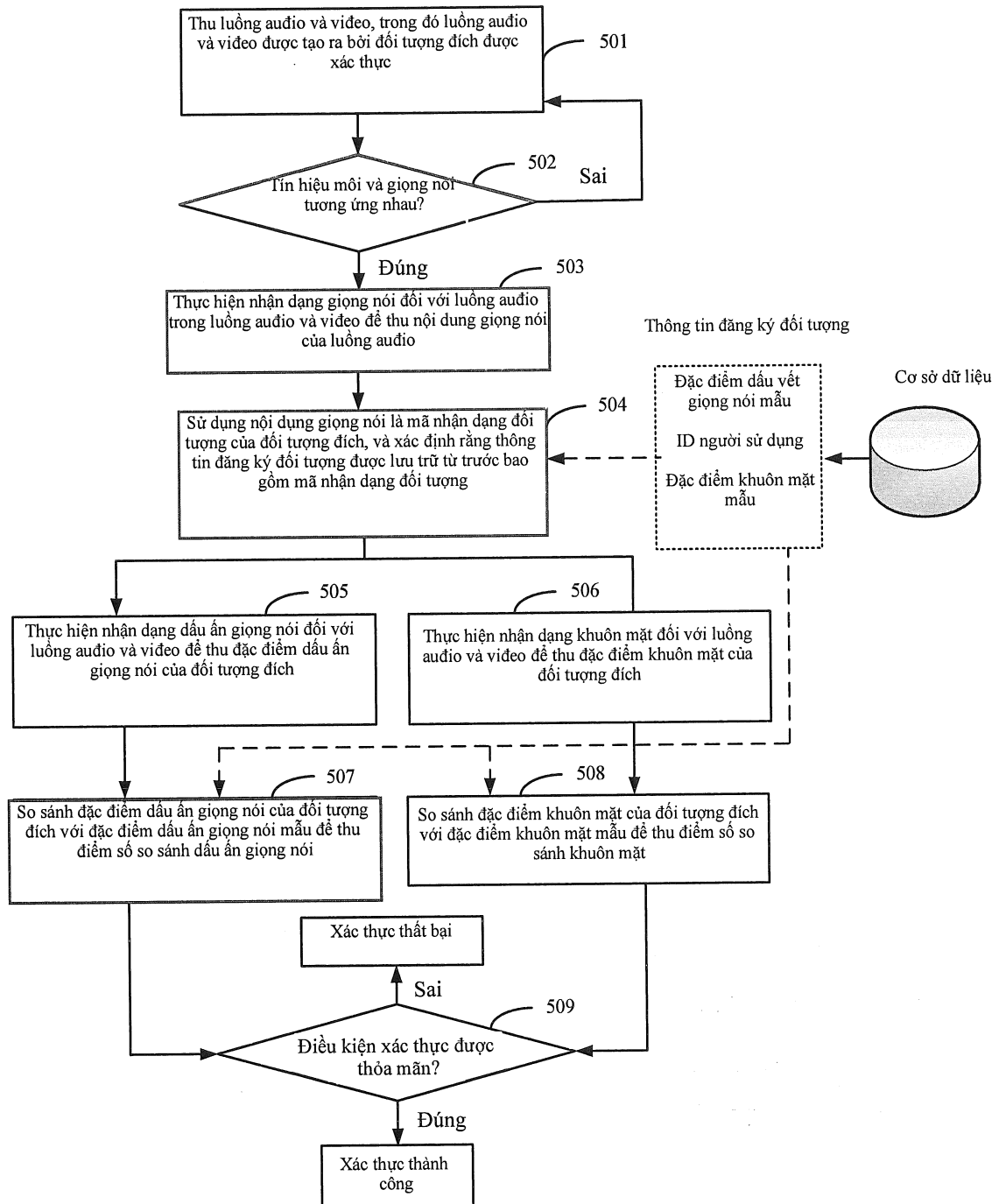


FIG. 5

6/7

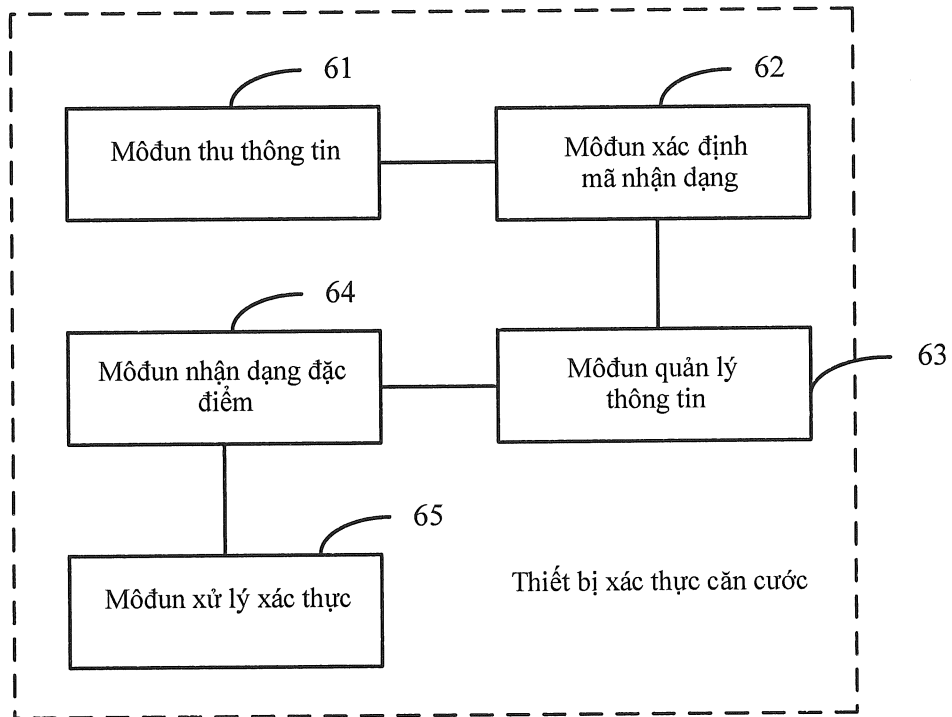


FIG. 6

7/7

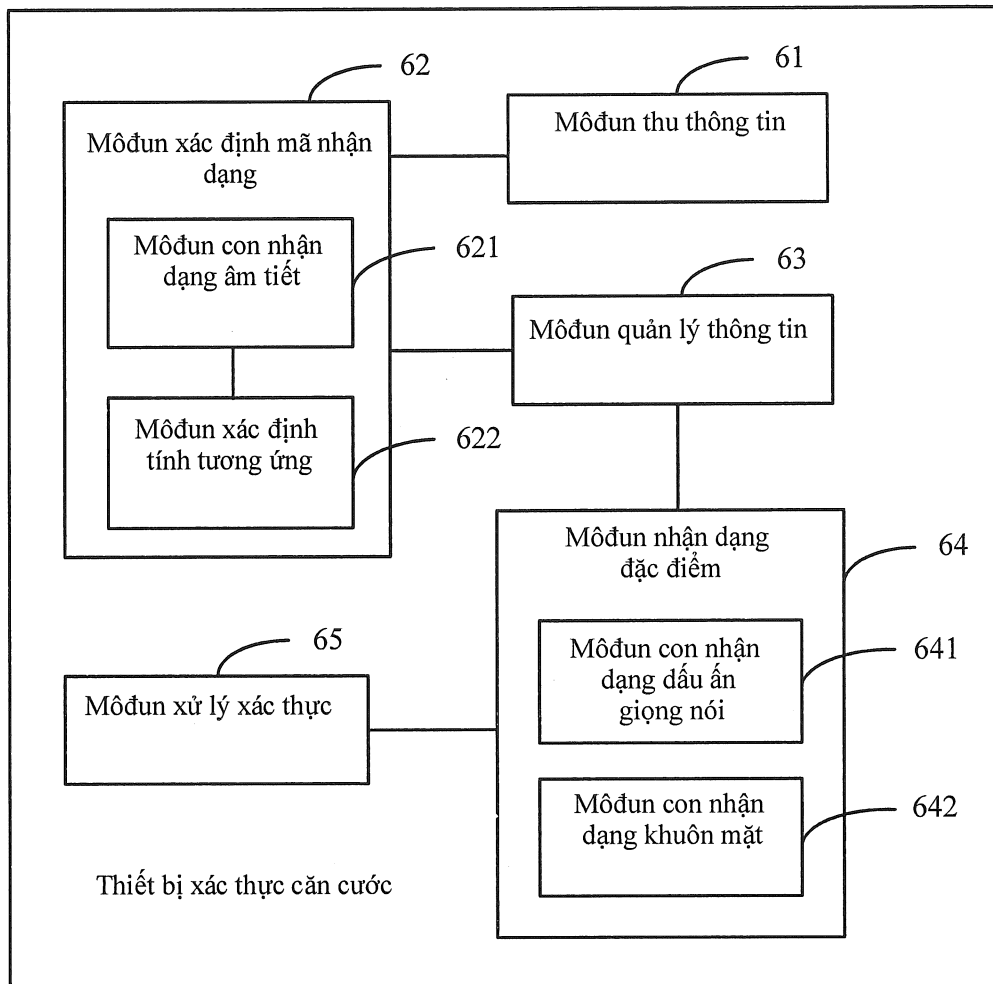


FIG. 7