



(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)



1-0035189

(51)⁷ A61B 5/11

(13) B

(21) 1-2019-00401

(22) 23/01/2019

(30) 2018-010096 24/01/2018 JP

(45) 25/04/2023 421

(43) 25/07/2019 376A

(73) HITACHI, LTD. (JP)

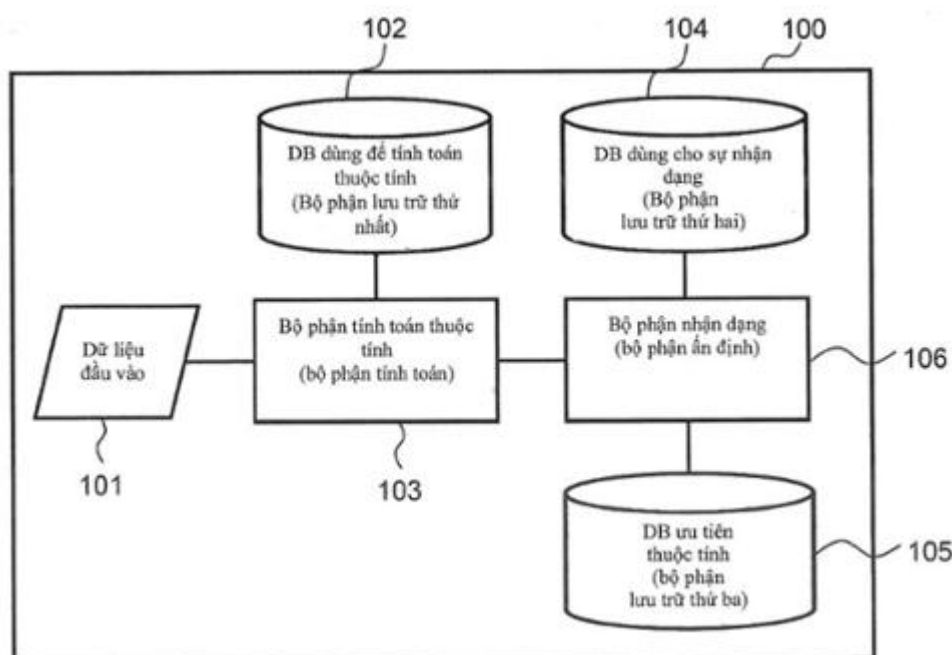
6-6, Marunouchi 1-chome, Chiyoda-ku, Tokyo 1008280, Japan

(72) Takuto SATO (JP); Hiroki OHASHI (JP); Katsuyuki NAKAMURA (JP);
Mohammad Osamh Adel AL-NASER (JO); Sheraz AHMED (PK).

(74) Công ty TNHH một thành viên Sở hữu trí tuệ VCCI (VCCI-IP CO.,LTD)

(54) HỆ THỐNG ĐẶC TẢ VÀ PHƯƠNG PHÁP ĐẶC TẢ

(57) Sáng chế đề cập đến hệ thống và phương pháp đặc tả. Hệ thống đặc tả bao gồm: bộ phận lưu trữ thứ nhất lưu trữ thông số để đưa ra đích xử lý trong không gian biểu diễn trung gian có mỗi sự biểu diễn trung gian chỉ báo mỗi thuộc tính của đích xử lý là mỗi kích cỡ; bộ phận tính toán đưa ra dữ liệu đầu vào thu được từ cảm biến định trước hoặc trị số thu được nhờ thực hiện quy trình tiền xử lý định trước trên dữ liệu đầu vào trong không gian biểu diễn trung gian dựa trên thông số được lưu trữ trong bộ phận lưu trữ thứ nhất; bộ phận lưu trữ thứ hai lưu trữ thông số mà được sử dụng khi đích xử lý được định rõ trong không gian biểu diễn trung gian; bộ phận lưu trữ thứ ba lưu trữ quyền ưu tiên của mỗi sự biểu diễn trung gian chỉ báo mỗi thuộc tính của đích xử lý tương ứng với đích xử lý; và bộ phận ấn định định rõ đích xử lý trong dữ liệu đầu vào dựa trên thông số được lưu trữ trong bộ phận lưu trữ thứ hai, quyền ưu tiên được lưu trữ trong bộ phận lưu trữ thứ ba, và đầu ra của bộ phận tính toán.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến hệ thống đặc tả và phương pháp đặc tả và tốt hơn là được ứng dụng cho, ví dụ, hệ thống đặc tả định rõ đích xử lý mà được bao gồm trong dữ liệu đầu vào thu được từ cảm biến và phương pháp đặc tả.

Tình trạng kỹ thuật của sáng chế

Việc nhận dạng, phân biệt, và phát hiện đích được chỉ báo bởi dữ liệu cảm biến nhờ sử dụng các loại dữ liệu cảm biến khác nhau chẳng hạn như nhận dạng hành động, nhận dạng đối tượng, và nhận dạng giọng nói được yêu cầu.

Chẳng hạn như, đối với hệ thống giám sát, nhật ký cuộc sống, hoặc tương tự, việc phân tích dữ liệu cảm biến thu được từ máy ảnh hoặc thiết bị tương tự và nhận dạng tự động hành vi của con người được yêu cầu.

Nhiều công nghệ nhận dạng đã được phát triển để thỏa mãn các nhu cầu nêu trên. Như một phương pháp điển hình, phương pháp được lấy ví dụ thu thập dữ liệu nhận biết (dữ liệu cảm biến tương ứng với các nhóm tương ứng và dữ liệu nhãn của các tên nhóm của chúng) đối với mỗi nhóm mà là đích nhận dạng và nhận biết sự tương ứng giữa dữ liệu cảm biến và nhãn nhóm nhờ sử dụng phương pháp chẳng hạn như mạng nơron, SVM (Support Vector Machine – Máy vectơ hỗ trợ), và phương pháp tương tự.

Chẳng hạn như, đơn yêu cầu cấp bằng sáng chế Nhật Bản số JP-A-2017-94055 (tài liệu sáng chế 1) bộc lộ phương pháp nhận dạng chuyển động mà nhận dạng chuyển động của cơ thể để điều khiển chức năng của thiết bị. Cụ thể hơn, tài liệu sáng chế 1 bộc lộ phương pháp trong đó sự tạo giả chuyển động bởi người dùng là chuyển động chủ ý được phát hiện từ tín hiệu đầu ra của cảm biến tín hiệu sinh học, và tín hiệu điều khiển để điều khiển chức năng của thiết bị đích tương ứng với mẫu tín hiệu tham chiếu được tạo ra

trong trường hợp trong đó mẫu tín hiệu của sự tạo giả chuyển động được phát hiện tương ứng với mẫu tín hiệu tham chiếu.

Tuy nhiên, theo phương pháp mà thu thập dữ liệu nhận biết của mỗi nhóm của đích nhận dạng trước đó và nhận biết sự tương ứng giữa dữ liệu cảm biến và nhãn nhóm bởi dữ liệu nhận biết, khi nhóm khác (nhóm mới) với nhóm đích xác định trước (nhóm hiện có) được bổ sung một lần nữa vào đích nhận dạng, cần thiết phải thu thập dữ liệu nhận biết của nhóm mới và xây dựng lại mô hình nhận dạng.

Theo đó, “Zero-data Learning of New Tasks” (Larochelle, H, Erhan, D và Bengio, Y; báo cáo của hội nghị AAAI lần thứ hai mươi về trí tuệ nhân tạo, 2008), đơn yêu cầu cấp bằng sáng chế Nhật bản số JP-A-2017-91192 (tài liệu sáng chế 2), và tài liệu tương tự đề xuất phương pháp mà cho phép sự nhận dạng kể cả trong trường hợp trong đó có một lượng nhỏ dữ liệu nhận biết của nhóm mới, hoặc không có dữ liệu nhận biết. Nói chung, phương pháp trong trường hợp trong đó không có dữ liệu nhận biết của đích nhận dạng mới được gọi là nhận biết không có quan sát (zero-shot), phương pháp trong trường hợp trong đó chỉ có một mẫu được gọi là nhận biết với một quan sát (one-shot), và phương pháp trong trường hợp trong đó có một lượng nhỏ dữ liệu nhận biết được gọi là nhận biết với một vài quan sát (few-shot).

Theo các phương pháp nhận dạng chung khác, mà trong đó mỗi nhóm được phân phối trong không gian đầu vào (không gian đặc tính) trước đó được học từ sự phân phối dữ liệu nhận biết nằm trong không gian đầu vào (hoặc không gian đặc tính thu được nhờ thay đổi không gian đầu vào). Khi mẫu mới được đưa ra, sự nhận dạng nhóm được thực hiện dựa trên vị trí của mẫu trong không gian đầu vào (không gian đặc tính). Theo đó, theo nhận biết không có quan sát, tín hiệu đầu vào (hoặc lượng đặc tính mà tín hiệu đầu vào được chuyển đổi) được chuyển đổi thành sự biểu diễn trung gian (mà thường được gọi là “thuộc tính”, và thuật ngữ này được sử dụng trong đơn sáng chế này), và

sự phân loại của mỗi mẫu được thực hiện nằm trong không gian thuộc tính. Thuộc tính mà là sự biểu diễn trung gian của dữ liệu đầu vào và nhãn đầu ra chỉ báo thuộc tính của đích phân loại. Chẳng hạn như, đối với lượng đặc tính bao gồm SIFT (Scale-Invariant Feature Transform - Biến đổi đặc tính bất biến tỷ lệ) lượng đặc tính mà thường được sử dụng thông thường để nhận dạng hình ảnh hoặc dạng tương tự, con người rất khó diễn giải bằng trực giác giá trị. Tuy nhiên, đối với thuộc tính, con người có thể diễn giải thuộc tính trong nhiều trường hợp. Phần mô tả được đưa ra dưới đây về ví dụ cụ thể của thuộc tính. Phép chiếu $f:X \rightarrow A$ từ không gian đầu vào X vào không gian thuộc tính A được học nhờ sử dụng dữ liệu nhận biết hiện có trong khi trong đó mỗi nhóm được phân phối nằm trong không gian thuộc tính được định rõ bởi phương pháp mà không sử dụng dữ liệu nhận biết. Do đó, trong trường hợp trong đó không có dữ liệu nhận biết, đó là, trên đích nhận dạng mới (nhóm mới), sự nhận dạng có thể được thực hiện.

Chẳng hạn như, bộ phận phân biệt có thể được tạo cấu hình để phát hiện thuộc tính “liệu nó có bốn chân” khi dữ liệu của con chó, gia súc, con người, con gà, hoặc tương tự được bao gồm trong dữ liệu nhận biết trong nhiệm vụ nhận dạng đối tượng sử dụng dữ liệu hình ảnh. Tương tự, bộ phận phân biệt có thể được tạo cấu hình để phát hiện thuộc tính “liệu nó có hoa văn sọc” khi dữ liệu của con cá vẹt, áo sơ mi sọc, cá heo, áo sơ mi trơn, hoặc tương tự được bao gồm. Theo cách này, chuyển đổi $f:X \rightarrow A$ từ đầu vào sang thuộc tính có thể được xây dựng từ dữ liệu nhận biết hiện có.

Ở đây, chẳng hạn như, trong trường hợp trong đó “ngựa vằn” được mong muốn để được phát hiện làm đích nhận dạng, khi ngựa vằn được xác định là một vật có “bốn chân” và “hoa văn sọc” (trong trường hợp này, vị trí của sự phân phối (điểm) của nhóm “ngựa vằn” nằm trong không gian thuộc tính được xác định thủ công), khi hình ảnh của ngựa vằn được đưa ra bởi sự chuyển đổi $f:X \rightarrow A$ học từ dữ liệu nhận biết hiện có mặc dù hình ảnh của ngựa

vẫn không được bao gồm trong dữ liệu nhận biết, “liệu nó có bốn chân” và “liệu nó có hoa văn sọc” có thể được đánh giá. Khi xác suất cả hai là “Có” cao, hình ảnh có thể được đánh giá để biểu diễn ngựa vằn.

Một cách ngẫu nhiên, sự chuyển đổi từ vector đầu vào sang vector thuộc tính là khác với sự trích xuất đặc tính mà được sử dụng trong học máy nói chung mà trong đó sự phân phối, điểm đại diện, hoặc tương tự của mỗi nhóm nằm trong không gian thuộc tính có thể được xác định theo phương pháp mà không sử dụng dữ liệu nhận biết.

Trong nhiều trường hợp, phương pháp lân cận gần nhất hoặc phương pháp tương tự được sử dụng làm phương pháp điển hình của việc phân loại nhóm nằm trong không gian thuộc tính.

Danh mục tài liệu trích dẫn

Tài liệu sáng chế

Tài liệu sáng chế 1: Đơn yêu cầu cấp bằng sáng chế Nhật bản số JP-A-2017-94055

Tài liệu sáng chế 2: Đơn yêu cầu cấp bằng sáng chế Nhật bản số JP-A-2017-91192

Tài liệu không phải sáng chế

Tài liệu không phải sáng chế 1: Larochelle, H, Erhan, D và Bengio, Y; báo cáo của hội nghị AAAI lần thứ hai mươi về trí tuệ nhân tạo, 2008

Theo cách này, theo nhận biết không có quan sát/với một quan sát/với một vài quan sát, sự nhận dạng có thể được thực hiện kể cả trong trường hợp trong đó có (không có, chỉ có một, hoặc có lượng nhỏ) dữ liệu nhận biết của đích nhận dạng mới. Mặt khác, khi số lượng thuộc tính được tăng lên (số lượng kích cỡ của không gian thuộc tính được tăng lên) để cho phép sự phân loại của nhiều đích khác nhau, thuộc tính mà không quan trọng đối với sự phân

loại của đích nhận dạng khác có thể được bao gồm mặc dù nó rất hữu ích cho sự phân loại của đích nhận dạng nhất định.

Chẳng hạn như, trong ví dụ trước về “ngựa vằn”, “màu sắc bầu trời trong nền” được coi là hữu ích cho việc phân loại thời tiết của địa điểm của hình ảnh, nhưng không quan trọng để đánh giá liệu ngựa vằn có được thể hiện trong hình ảnh hay không.

Như đã được nêu trên, trong không gian trong đó thuộc tính mà không quan trọng được bao gồm, khi phương pháp lân cận gần nhất hoặc phương pháp tương tự trong đó tất cả thuộc tính (kích cỡ) được coi tương đương được ứng dụng, kết quả nhận dạng ảnh hưởng bởi thuộc tính (kích cỡ) mà không quan trọng, điều này gây ra vấn đề là độ chính xác nhận dạng bị suy giảm.

Bản chất kỹ thuật của sáng chế

Sáng chế được tạo ra khi xem xét các vấn đề nêu trên, và mục đích của sáng chế là đề xuất hệ thống đặc tả và phương pháp đặc tả trong đó đích xử lý có thể được định rõ với độ chính xác cao ngay cả trong trường hợp trong đó không có dữ liệu nhận biết của nhóm mới hoặc chỉ có lượng nhỏ dữ liệu nhận biết có thể thu được.

Theo một khía cạnh để giải quyết vấn đề nêu trên, hệ thống đặc tả bao gồm: bộ phận lưu trữ thứ nhất lưu trữ thông số để đưa ra đích xử lý trong không gian biểu diễn trung gian có mỗi sự biểu diễn trung gian chỉ báo mỗi thuộc tính của đích xử lý là mỗi kích cỡ; bộ phận tính toán đưa ra dữ liệu đầu vào thu được từ cảm biến định trước hoặc trị số thu được nhờ thực hiện quy trình tiền xử lý định trước trên dữ liệu đầu vào trong không gian biểu diễn trung gian dựa trên thông số được lưu trữ trong bộ phận lưu trữ thứ nhất; bộ phận lưu trữ thứ hai lưu trữ thông số mà được sử dụng khi đích xử lý được định rõ trong không gian biểu diễn trung gian; bộ phận lưu trữ thứ ba lưu trữ quyền ưu tiên của mỗi sự biểu diễn trung gian chỉ báo mỗi thuộc tính của đích

xử lý tương ứng với đích xử lý; và bộ phận ấn định định rõ đích xử lý trong dữ liệu đầu vào dựa trên thông số được lưu trữ trong bộ phận lưu trữ thứ hai, quyền ưu tiên được lưu trữ trong bộ phận lưu trữ thứ ba, và đầu ra của bộ phận tính toán.

Theo một khía cạnh, phương pháp đặc tả, trong hệ thống đặc tả mà bao gồm bộ phận lưu trữ thứ nhất lưu trữ thông số để đưa ra đích xử lý trong không gian biểu diễn trung gian có mỗi sự biểu diễn trung gian chỉ báo mỗi thuộc tính của đích xử lý là mỗi kích cỡ, bộ phận lưu trữ thứ hai lưu trữ thông số mà được sử dụng khi đích xử lý được định rõ trong không gian biểu diễn trung gian, và bộ phận lưu trữ thứ ba lưu trữ quyền ưu tiên của mỗi sự biểu diễn trung gian chỉ báo mỗi thuộc tính của đích xử lý tương ứng với đích xử lý, bao gồm: bước thứ nhất, bộ phận tính toán đưa ra dữ liệu đầu vào thu được từ cảm biến định trước hoặc trị số thu được nhờ thực hiện quy trình tiền xử lý định trước trên dữ liệu đầu vào trong không gian biểu diễn trung gian dựa trên thông số được lưu trữ trong bộ phận lưu trữ thứ nhất; và bước thứ hai, bộ phận ấn định định rõ đích xử lý trong dữ liệu đầu vào dựa trên thông số được lưu trữ trong bộ phận lưu trữ thứ hai, quyền ưu tiên được lưu trữ trong bộ phận lưu trữ thứ ba, và đầu ra của bộ phận tính toán.

Theo cấu hình này, đích xử lý có thể được định rõ với độ chính xác cao ngay cả trong trường hợp trong đó không có dữ liệu nhận biết của nhóm mới hoặc chỉ có lượng nhỏ dữ liệu nhận biết có thể thu được, chẳng hạn.

Theo sáng chế, đích xử lý có thể được định rõ với độ chính xác cao.

Các vấn đề, cấu hình, và hiệu quả khác với các vấn đề, cấu hình và hiệu quả nêu trên sẽ trở nên rõ ràng từ phần mô tả sau đây của các phương án.

Mô tả vắn tắt các hình vẽ

Fig.1 là hình vẽ minh họa một ví dụ về cấu hình của hệ thống nhận dạng theo phương án thứ nhất;

Fig.2 là hình vẽ minh họa một ví dụ về cấu hình phần cứng của hệ thống nhận dạng theo phương án thứ nhất;

Fig.3 là hình vẽ minh họa một ví dụ về dữ liệu đầu ra của bộ phận tính toán thuộc tính theo phương án thứ nhất;

Fig.4A và Fig.4B là các hình vẽ minh họa một ví dụ về sự phân phối của mẫu dữ liệu nằm trong không gian thuộc tính theo phương án thứ nhất;

Fig.5 là hình vẽ minh họa một ví dụ về trường hợp dùng để giải thích sự nhận dạng sai của phương pháp thông thường;

Fig.6 là hình vẽ minh họa một ví dụ về điểm mẫu giả trong không gian thuộc tính theo phương án thứ nhất;

Fig.7 là hình vẽ minh họa một ví dụ về quyền ưu tiên của mỗi thuộc tính theo phương án thứ nhất;

Fig.8 là hình vẽ minh họa ví dụ khác về cấu hình của hệ thống nhận dạng theo phương án thứ nhất;

Fig.9 là hình vẽ minh họa một ví dụ về GUI để thiết đặt quyền ưu tiên của mỗi thuộc tính đối với mỗi nhóm theo phương án thứ nhất;

Fig.10 là hình vẽ minh họa một ví dụ về cấu hình của hệ thống nhận dạng theo phương án thứ hai;

Fig.11 là hình vẽ minh họa một ví dụ về cấu hình của hệ thống nhận dạng theo phương án thứ ba;

Fig.12 là hình vẽ minh họa một ví dụ về cấu hình của hệ thống nhận dạng theo phương án thứ tư;

Fig.13 là hình vẽ minh họa một ví dụ về cấu hình của hệ thống nhận dạng theo phương án thứ tư;

Fig.14 là hình vẽ minh họa một ví dụ về cấu hình của bộ phận đánh giá nhìn tập trung theo phương án thứ tư; và

Fig.15 là hình vẽ minh họa một ví dụ về cấu hình của bộ phận đánh giá nhìn tập trung theo phương án thứ tư.

Mô tả chi tiết sáng chế

Sau đây, phương án của sáng chế sẽ được mô tả một cách chi tiết dựa vào các hình vẽ kèm theo.

(1) Phương án thứ nhất

Phương pháp cải thiện đối với phương pháp mà được gọi chung là nhận biết không có quan sát, nhận biết với một quan sát, và nhận biết với một vài quan sát được đề xuất theo phương án này. Khi phương pháp học máy khác được sử dụng, dữ liệu mẫu (dữ liệu nhận biết) của các đối tượng của nhóm của đích nhận dạng được thu thập trước đó, và mô hình nhận dạng được xây dựng dựa trên dữ liệu nhận biết. Vì lý do này, nhóm (mà được gọi là nhóm hiện có trong bản mô tả này) trong đó (lượng lớn) dữ liệu nhận biết có thể được thu thập là có thể nhận dạng, nhưng nhóm (mà được gọi là nhóm mới trong bản mô tả này) mà không có dữ liệu nhận biết hoặc lượng dữ liệu nhận biết cực nhỏ là không thể nhận dạng. Ngược lại với điều này, theo nhận biết không có quan sát, nhận biết với một quan sát, và nhận biết với một vài quan sát, kể cả trong trường hợp trong đó dữ liệu mẫu của các đối tượng của nhóm của đích nhận dạng hoàn toàn không được bao gồm trong dữ liệu nhận biết, kể cả trong trường hợp trong đó chỉ một mẫu được bao gồm trong đó, và kể cả trong trường hợp trong đó chỉ lượng nhỏ được bao gồm trong đó, nhóm có thể được nhận dạng nhờ sử dụng thuộc tính (sự biểu diễn trung gian). Theo phương án này, hệ thống mới và phương pháp mới để cải thiện độ chính xác tại thời điểm thực hiện nhận dạng sử dụng thuộc tính được cung cấp. Sau đây, cấu hình cụ thể sẽ được mô tả.

Trên Fig.1, số chỉ dẫn 100 biểu thị hệ thống nhận dạng theo phương án thứ nhất nói chung. Hệ thống nhận dạng 100 bao gồm dữ liệu đầu vào 101, DB

(database - cơ sở dữ liệu) dùng cho sự tính toán thuộc tính 102, bộ phận tính toán thuộc tính 103, DB dùng cho sự nhận dạng 104, DB ưu tiên thuộc tính 105, và bộ phận nhận dạng 106.

Ví dụ về các đối tượng mà được nhận dạng bởi hệ thống nhận dạng 100 bao gồm hành vi của con người (hệ thống nhận dạng hành động), đối tượng (hệ thống nhận dạng đối tượng) mà được chụp thành hình ảnh, video, hoặc tương tự, giọng nói (hệ thống nhận dạng giọng nói), chữ cái (hệ thống nhận dạng chữ cái) và tương tự. Sau đây, quy trình cụ thể (bước) sẽ được mô tả nhờ sử dụng chủ yếu hệ thống nhận dạng hành động làm ví dụ. Tuy nhiên, đích nhận dạng của hệ thống nhận dạng 100 không giới hạn ở hệ thống nhận dạng hành động.

Các cảm biến khác nhau chẳng hạn như máy ảnh, micro, cảm biến khoảng cách, cảm biến gia tốc, cảm biến con quay, cảm biến điện cơ, cảm biến xung nhịp/nhịp tim, cảm biến chiếu sáng, và cảm biến áp có thể được sử dụng riêng lẻ hoặc có thể được sử dụng kết hợp để thu dữ liệu đầu vào 101. Loại hoặc số lượng các cảm biến không giới hạn. Ngoài ra, giá trị đo của cảm biến có thể được lưu trữ một lần trong đĩa cứng hoặc vật tương tự để được nhập vào bộ phận tính toán thuộc tính 103, và sau đó quy trình tiếp theo có thể được thực hiện (quy trình ngoại tuyến). Nếu không, giá trị đo đọc từ cảm biến có thể được nhập trực tiếp vào bộ phận tính toán thuộc tính 103, và sau đó quy trình tiếp theo có thể được thực hiện (quy trình thời gian thực).

Trong lĩnh vực học máy, các thuật ngữ chẳng hạn như “nhận dạng”, “phân loại”, “phát hiện”, và “định vị” có thể được sử dụng. Trong bản mô tả này, ví dụ về hệ thống “nhận dạng” được mô tả. Tuy nhiên, phạm vi bảo hộ của sáng chế không giới hạn ở vấn đề nhận dạng vì hệ thống nhận dạng được ứng dụng dễ dàng vào vấn đề phân loại, vấn đề phát hiện, và vấn đề định vị. Một cách ngẫu nhiên, “đặc tả” được sử dụng một cách thích hợp làm thuật ngữ mà chỉ báo khái niệm nổi trội của các thuật ngữ chẳng hạn như “nhận dạng”, “phân loại”, “phát hiện”, và “định vị”.

Chẳng hạn như, trong bản mô tả này, trong trường hợp trong đó một nhóm mà là đích phân loại nhất thiết phải tương ứng với một dữ liệu đầu vào, “phân loại” chỉ báo nhóm nhận dạng. Trong trường hợp trong đó điều không được biết là liệu nhóm mà là đích nhận dạng có được thể hiện đối với một dữ liệu đầu vào hay không, hoặc điều không được biết là nhóm là một hay nhiều, “nhận dạng” chỉ báo việc nhận dạng trong đó nhóm, khi nhóm là đích nhận dạng, được thể hiện.

Ở đây, trong ví dụ về nhận dạng hành động và phân loại hành động sử dụng cảm biến mang được, trong trường hợp trong đó tất cả các giá trị đo của cảm biến trong khoảng thời gian nhất định tương ứng với một hành động đích bất kỳ, hành động cần được nhận dạng, mà là vấn đề “phân loại”. Trong trường hợp trong đó 0 hoặc nhiều hành động đích tương ứng với dữ liệu cảm biến trong khoảng thời gian nhất định, khi hành động đích bất kỳ được bao gồm, hành động được thực hiện trong khoảng thời gian nhất định cần được nhận dạng, đó là vấn đề “nhận dạng”.

Tương tự, trong ví dụ về nhận dạng đối tượng và phân loại đối tượng sử dụng hình ảnh, trong trường hợp trong đó toàn bộ một hình ảnh thể hiện một đối tượng đích bất kỳ, đối tượng cần được nhận dạng, đó là vấn đề “phân loại”. Trong trường hợp trong đó 0 hoặc nhiều đối tượng đích được thể hiện trong hình ảnh, khi đối tượng đích bất kỳ được bao gồm, mà trong đó đối tượng cần được nhận dạng trong hình ảnh, đó là vấn đề “nhận dạng”.

Đầu tiên, phần tổng quan của hệ thống nhận dạng 100 sẽ được mô tả. Hệ thống nhận dạng 100 đọc giá trị của dữ liệu đầu vào 101. Hơn nữa, bộ phận tính toán thuộc tính 103 thực hiện quy trình trích xuất đặc tính hoặc tương tự khi cần nhờ sử dụng giá trị, và phép chiếu của dữ liệu đầu vào vào không gian thuộc tính được thực hiện nhờ sử dụng thông số được lưu trữ trong DB dùng cho sự tính toán thuộc tính 102. Bộ phận nhận dạng 106 thu đầu ra của chúng. Hơn nữa, bộ phận nhận dạng đọc thông số, mà được lưu trữ trong DB dùng

cho sự nhận dạng 104, để thực hiện quy trình nhận dạng trong không gian thuộc tính, bổ sung quyền ưu tiên (sự quan trọng) của mỗi thuộc tính được lưu trữ trong DB ưu tiên thuộc tính 105, và thực hiện quy trình nhận dạng của dữ liệu đầu vào. Mỗi quy trình, dữ liệu được lưu trữ trong mỗi DB, và tương tự sẽ được mô tả một cách chi tiết dưới đây.

Ngoài ra, DB dùng cho sự tính toán thuộc tính 102 là một ví dụ về bộ phận lưu trữ thứ nhất lưu trữ thông số để đưa ra đích xử lý trong không gian biểu diễn trung gian trong đó mỗi sự biểu diễn trung gian chỉ báo các thuộc tính của đích xử lý được thiết đặt là mỗi kích cỡ. Bộ phận tính toán thuộc tính 103 là một ví dụ về bộ phận tính toán đưa ra dữ liệu đầu vào thu được từ cảm biến định trước hoặc giá trị thu được nhờ thực hiện quy trình tiền xử lý định trước trên dữ liệu đầu vào trong không gian biểu diễn trung gian dựa trên thông số được lưu trữ trong bộ phận lưu trữ thứ nhất. DB dùng cho sự nhận dạng 104 là một ví dụ về bộ phận lưu trữ thứ hai lưu trữ thông số được sử dụng để định rõ đích xử lý trong không gian biểu diễn trung gian. DB ưu tiên thuộc tính 105 là một ví dụ về bộ phận lưu trữ thứ ba lưu trữ quyền ưu tiên của mỗi sự biểu diễn trung gian chỉ báo mỗi thuộc tính của đích xử lý tương ứng với đích xử lý. Bộ phận nhận dạng 106 là một ví dụ về bộ phận ấn định định rõ đích xử lý trong dữ liệu đầu vào dựa trên thông số được lưu trữ trong bộ phận lưu trữ thứ hai, quyền ưu tiên được lưu trữ trong bộ phận lưu trữ thứ ba, và đầu ra của bộ phận tính toán.

Fig.2 là hình vẽ minh họa một ví dụ về cấu hình phần cứng trong hệ thống nhận dạng 100. Chẳng hạn như, hệ thống nhận dạng 100 có thể được thực hiện nhờ sử dụng máy tính 200 mà bao gồm cấu hình phần cứng được minh họa trên Fig.2. Máy tính 200 là máy tính xách tay, thiết bị đầu cuối dạng bảng, hoặc thiết bị tương tự. Máy tính bao gồm thiết bị đầu vào 201, bộ xử lý trung tâm 202, thiết bị lưu trữ chính 203, thiết bị lưu trữ phụ 204, thiết bị đầu ra 205, và buýt 206. Trong thiết bị đầu vào 201, bộ xử lý trung tâm 202, thiết

bị lưu trữ chính 203, thiết bị lưu trữ phụ 204, và thiết bị đầu ra 205 có thể trao đổi dữ liệu thông qua buýt 206.

Thiết bị đầu vào 201 là bàn phím, chuột, hoặc thiết bị tương tự. Thiết bị đầu vào nhập vào các loại thông tin khác nhau. Chẳng hạn như, bộ xử lý trung tâm 202 là CPU (Central Processing Unit – Bộ phận xử lý trung tâm). Bộ xử lý trung tâm thực thi chương trình được lưu trữ trong thiết bị lưu trữ chính 203. Thiết bị lưu trữ chính 203 là thiết bị lưu trữ khả biến và tốc độ cao chẳng hạn như DRAM (Dynamic Random Access Memory – Bộ nhớ truy cập ngẫu nhiên động) và lưu trữ hệ điều hành (OS), chương trình ứng dụng, và tương tự. Khi bộ xử lý trung tâm 202 thực thi OS, chức năng cơ bản của máy tính 200 được thực hiện, và khi chương trình ứng dụng được thực thi, bộ phận tính toán thuộc tính 103 và bộ phận nhận dạng 106 (chức năng của hệ thống nhận dạng 100) được thực hiện. Quy trình của bộ phận tính toán thuộc tính 103 và bộ phận nhận dạng 106 sẽ được mô tả một cách chi tiết dưới đây.

Thiết bị lưu trữ phụ 204 là thiết bị lưu trữ không khả biến và cỡ lớn chẳng hạn như thiết bị lưu trữ từ tính hoặc bộ nhớ tác động nhanh và lưu trữ phần mềm được thực thi bởi bộ xử lý trung tâm 202, dữ liệu được sử dụng để thực thi phần mềm, và tương tự. Chẳng hạn như, phần mềm được thực thi bởi bộ xử lý trung tâm 202 được đọc từ thiết bị lưu trữ phụ 204 để được tải vào thiết bị lưu trữ chính 203 và được thực thi bởi bộ xử lý trung tâm 202. Ngoài ra, DB dùng cho sự tính toán thuộc tính 102, DB dùng cho sự nhận dạng 104, và DB ưu tiên thuộc tính 105 được lưu trữ trong thiết bị lưu trữ phụ 204.

Thiết bị đầu ra 205 là màn hình tinh thể lỏng, máy chiếu, máy in, hoặc thiết bị tương tự và có thể xuất kết quả nhận dạng. Ngoài ra, GPU (Graphics Processing Unit – Bộ xử lý đồ họa) có thể được sử dụng để làm cho thao tác của bộ phận tính toán thuộc tính 103 và bộ phận nhận dạng 106 được thực hiện nhanh chóng.

Dữ liệu đầu vào 101 có thể được truyền đến máy tính 200 theo cách mà

cảm biến và máy tính 200 được kết nối nhờ sử dụng cáp USB (Universal Serial Bus – Buýt nối tiếp vạn năng) hoặc tương tự. Dữ liệu đầu vào 101 có thể được truyền thông qua mạng. Dữ liệu đầu vào 101 có thể được ghi một lần trên phương tiện chẳng hạn như CD (Compact Disc – Đĩa compac) và DVD (Digital Versatile Disc – Đĩa đa năng kỹ thuật số) và có thể được đọc bởi máy tính 200.

Sau đây, mỗi quy trình, định dạng dữ liệu, và tương tự sẽ được mô tả một cách chi tiết nhờ sử dụng trường hợp mà trong đó hệ thống nhận dạng 100 là hệ thống nhận dạng hành động làm ví dụ. Để nhận dạng tự động hành động phức tạp với độ chính xác cao, phương pháp đã biết trong đó hành động phức tạp được phân tích thành nhiều yếu tố cơ bản và được xem xét, và sự nhận dạng được thực hiện chính xác để cải thiện độ chính xác nhận dạng. Đó là, đầu tiên, trong bước thứ nhất, sự nhận dạng của mỗi yếu tố được thực hiện từ dữ liệu đầu vào, và trong bước thứ hai, nhận dạng hành động được thực hiện dựa trên kết quả nhận dạng của mỗi yếu tố.

Chuyển động được lấy ví dụ làm yếu tố đơn giản mà định hình hành động phức tạp. Ở đây, chuyển động chỉ báo hành động vật lý tương đối đơn giản chẳng hạn như “đưa cánh tay lên”, “cúi”, và “di động” và chuyển động của đối tượng. Hành động chỉ báo hành vi mà có ý nghĩa về mặt kích cỡ cao nhờ bao hàm các chuyển động chẳng hạn như “di chuyển một vật” và “tìm kiếm một đối tượng”, hoặc một vật. Như một ví dụ, hành động “di chuyển một vật” có thể được biểu diễn là sự kết hợp của ba chuyển động “nắm”, “di động”, và “thả”.

Một cách ngẫu nhiên, thực tế là, rất khó phân biệt rõ ràng “chuyển động” và “hành động”. Tuy nhiên, ý chính trong hệ thống nhận dạng 100 là biểu diễn hiện tượng phức tạp theo sự kết hợp của hiện tượng cơ bản và đơn giản. Theo ý nghĩa đó, không cần thiết phải xác định rõ ràng “chuyển động” và “hành động” nêu trên trong việc thực hiện hệ thống nhận dạng 100.

Chẳng hạn như, khi cần, “di động” được giải thích là một ví dụ về chuyển động trong phần mô tả ở trên có thể được xem xét là hành động, và chuyển động thuần của chân, chuyển động thuần của cánh tay, hoặc tương tự có thể được xem xét là chuyển động. Ngoài ra, “di chuyển một vật” được giải thích là một ví dụ về hành động có thể được xem xét là yếu tố chuyển động cơ bản đối với hành động phức tạp hơn “dọn phòng”.

Như đã nêu trên, khi hành động được biểu diễn chính xác, hành động có thể được chia thành hai hoặc nhiều bước và được xem xét. Chẳng hạn như, hành vi của con người và/hoặc chuyển động của con người được lấy ví dụ làm đích (đích xử lý được định rõ bởi bộ phận ấn định) được nhận dạng bởi bộ phận nhận dạng 106. Ngoài ra, yếu tố cơ bản định hình hành động có thể sử dụng tư thế (pose – dáng) của cơ thể, đối tượng hiện có gần đó, và tương tự khác với chuyển động.

Thuộc tính được tính toán bởi bộ phận tính toán thuộc tính 103 chỉ báo biểu diễn định tính của yếu tố cơ bản định hình hành động. Chuyển động, tư thế, đối tượng gần đó, và tương tự mà được mô tả ở trên là các ví dụ về thuộc tính. Nói cách khác, các sự biểu diễn trung gian chỉ báo các thuộc tính của đích xử lý bao gồm sự xuất hiện hoặc khả năng xuất hiện của chuyển động hoặc tư thế mà là yếu tố cơ bản định hình hành động của con người. Ngoài ra, các sự biểu diễn trung gian chỉ báo các thuộc tính của đích xử lý bao gồm sự có mặt của đối tượng được định trước hoặc khả năng có mặt của đối tượng được định trước xung quanh người dùng mà hành động được mong muốn được định rõ.

Chẳng hạn như, hành động “di chuyển giấy” có thể được biểu diễn theo sự kết hợp của các chuyển động “nắm”, “di động”, và “thả” và đối tượng “giấy”. Về mặt này, không gian thuộc tính được coi là không gian bốn chiều tương ứng với “nắm”, “di động”, “thả”, và “giấy”. Khi “1” được gán giá trị trong trường hợp trong đó chuyển động, đối tượng, hoặc dạng tương tự được nhận dạng trong dữ liệu đầu vào là kết quả của quy trình của bộ phận tính toán

thuộc tính 103 (được mô tả sau), và “0” được gán giá trị trong trường hợp trong đó không có nhận dạng, hành động “di chuyển giấy” có thể được thể hiện là điểm (1,1,1,1) trên không gian thuộc tính.

Thực ra, trong trường hợp trong đó thuộc tính (kích cỡ) khác được sử dụng để nhận dạng hành động khác, khi không gian thuộc tính được xác định là “nắm”, “di động”, “thả”, “nhảy”, “giấy”, và “mèo”, hành động “di chuyển giấy” được biểu diễn là điểm (1,1,1,0,1,0).

Ở đây, ví dụ được mô tả trong đó liệu thuộc tính được nhận dạng được biểu diễn bởi “0” hoặc “1” (sự có mặt, sự xuất hiện, hoặc tương tự). Tuy nhiên, sự biểu diễn có thể được đưa ra là số thực giữa “0” đến “1” nhờ sử dụng khả năng nhận dạng (khả năng có mặt, khả năng xuất hiện, hoặc tương tự) hoặc có thể sử dụng giá trị lớn hơn “1” hoặc giá trị âm.

Ở đây, sự tính toán thuộc tính là khác với quy trình trích xuất đặc tính mà thường được sử dụng trong học máy trong sự nhận dạng của đối tượng đích trong không gian thuộc tính không cần thiết phụ thuộc vào dữ liệu nhận biết (nhận biết không có quan sát) hoặc có thể được thực hiện bởi lượng dữ liệu nhận biết cực nhỏ (nhận biết với một quan sát, nhận biết với một vài quan sát).

Trong quy trình trích xuất đặc tính nói chung, khi quy trình nhận dạng được thực hiện nằm trong không gian đặc tính sau khi trích xuất đặc tính, không có sự biết trước theo cách mà mỗi nhóm của đích nhận dạng được phân phối nằm trong không gian đặc tính. Do đó, cần thiết tạo ra mô hình nhờ sử dụng (lượng lớn) dữ liệu nhận biết.

Tuy nhiên, trong không gian thuộc tính, mà không thu dữ liệu nhận biết của hành động “di chuyển giấy”, nó có thể được đánh giá rằng hành động “di chuyển giấy” được biểu diễn là điểm (1,1,1,0,1,0) nằm trong không gian thuộc tính có sáu kích cỡ “nắm”, “di động”, “thả”, “nhảy”, “giấy”, và “mèo” như trong ví dụ trước. Một cách ngẫu nhiên, trường hợp này tương ứng với trường

hợp mà trong đó định nghĩa của nhóm đích trên không gian thuộc tính được đưa ra một cách thủ công. Sau đó, phương pháp đưa ra định nghĩa theo phương pháp khác sẽ được mô tả.

Tuy nhiên, đối với phép chiếu từ dữ liệu đầu vào vào không gian thuộc tính, cần thiết thực hiện việc học nhờ sử dụng dữ liệu nhận biết. DB dùng cho sự tính toán thuộc tính 102 lưu trữ thông số cho việc học. Ngoài ra, bộ phận tính toán thuộc tính 103 đọc thông số được lưu trữ trong DB dùng cho sự tính toán thuộc tính 102 và thực hiện tính toán phép chiếu của dữ liệu đầu vào vào không gian thuộc tính.

Thông số mà được lưu trữ trong DB dùng cho sự tính toán thuộc tính 102 có thể thu được nhờ sử dụng phương pháp gọi là sự học có giám sát sử dụng dữ liệu nhận biết của nhóm hiện có. Phương pháp hiện có tùy ý có thể được sử dụng.

Chẳng hạn như, trong trường hợp trong đó dữ liệu gia tốc, dữ liệu con quay, và tương tự mà thu được từ cảm biến mang được sử dụng làm dữ liệu đầu vào, và chuyển động được sử dụng làm thuộc tính, RNN (Recurrent Neural Network – Mạng nơron tái phát) chẳng hạn như LSTM (Long-Short Term Memory – Bộ nhớ dài-ngắn hạn) có thể được sử dụng. Trong trường hợp này, chẳng hạn như, chuyển động mà là đích nhận dạng được thực hiện trong trạng thái đang được gắn với cảm biến mang được, và thời gian bắt đầu và thời gian kết thúc của chuyển động và nhãn của tên chuyển động được chuẩn bị làm dữ liệu nhận biết. Hơn nữa, thông số để nhận dạng chuyển động từ dữ liệu cảm biến được học nhờ sử dụng LSTM.

Một cách ngẫu nhiên, phương pháp nhận dạng chuyển động không nhất thiết chỉ giới hạn ở LSTM, và HMM (Hidden Markov Model – Mô hình Markov ẩn), CRF (Conditional Random Field – Trường điều kiện ngẫu nhiên), hoặc tương tự có thể được sử dụng. Ngoài ra, dữ liệu đầu vào hoặc thuộc tính không nhất thiết chỉ giới hạn ở dữ liệu của cảm biến mang được hoặc chuyển

động. Chẳng hạn như, trong trường hợp trong đó hình ảnh được sử dụng làm dữ liệu đầu vào, và đối tượng được sử dụng làm thuộc tính, CNN (Convolutional Neural Network – Mạng nơron xoắn), SVM (Support Vector Machine – Máy vectơ hỗ trợ), hoặc tương tự có thể được sử dụng. Ngoài ra, quy trình tiền xử lý chẳng hạn như sự làm trơn và sự trích xuất đặc tính có thể được thực hiện khi cần trước quy trình học, và dữ liệu đầu vào có thể được sử dụng trực tiếp. Thông số tương ứng với mô hình nhận dạng được lưu trữ trong DB dùng cho sự tính toán thuộc tính 102. Chẳng hạn như, trong mạng nơron, cấu trúc mạng, thông số trọng số giữa các nơron, và tương tự được lưu trữ.

Khi nhận dạng chuyển động được thực hiện từ dữ liệu không xác định (chẳng hạn như hình ảnh hoặc video theo góc nhìn người thứ nhất của người dùng), đầu tiên, bộ phận tính toán thuộc tính 103 đọc thông số được lưu trữ trong DB dùng cho sự tính toán thuộc tính 102. Dữ liệu đã cho được đưa vào quy trình tiền xử lý định trước khi cần, và sau đó giá trị đầu ra của mô hình nhận dạng được tính toán. Đầu ra điển hình là sự có mặt, khả năng, hoặc tương tự của chuyển động của mỗi đích nhận dạng, đối tượng, hoặc tương tự.

Fig.3 là hình vẽ minh họa một ví dụ về dữ liệu đầu ra của bộ phận tính toán thuộc tính 103. Trong trường hợp của ví dụ về Fig.3, không gian thuộc tính có bốn kích cỡ, và do đó, đầu ra là vectơ bốn chiều. Nói chung, đầu vào xi là vectơ đa chiều hoặc là kết quả thu được bằng cách thu thập các vectơ.

Sau đây, phần mô tả sẽ được đưa ra về quy trình của bộ phận nhận dạng 106 và dữ liệu được lưu trữ trong DB dùng cho sự nhận dạng 104 theo nhận biết không có quan sát, nhận biết với một quan sát, và nhận biết với một vài quan sát nói chung. Sau đó phần mô tả sẽ được đưa ra về các quy trình của DB ưu tiên thuộc tính 105 và bộ phận nhận dạng 106 mà sử dụng quyền ưu tiên của mỗi thuộc tính được lưu trữ trong DB ưu tiên thuộc tính 105, mà là các đặc tính của phương án này.

Theo nhận biết không có quan sát nói chung, một trong số các phương

pháp mà được sử dụng thường xuyên nhất trong trường hợp trong đó sự nhận dạng nhóm được thực hiện nằm trong không gian thuộc tính là phương pháp lân cận gần nhất. Khi dữ liệu đầu vào để được nhận dạng được thiết đặt là “x”, và “a” được thiết đặt nhờ đưa ra dữ liệu đầu vào trong không gian thuộc tính bởi bộ phận tính toán thuộc tính 103, theo phương pháp lân cận gần nhất, khoảng cách giữa phép chiếu a và mỗi mẫu trong không gian thuộc tính được tính toán, và nhãn nhóm của mẫu gần nhất được nhận dạng là nhãn nhóm của dữ liệu đầu vào x.

Chẳng hạn như, trên Fig.4A, khi “ Δ ” chỉ báo phép chiếu a của dữ liệu đầu vào x, và “ $\bigcirc$ ” và “X” chỉ báo các mẫu của nhóm hiện có, mẫu gần nhất là “ $\bigcirc$ ” kết nối với “ Δ ” bởi đường nét đứt. Do đó, dữ liệu đầu vào X được phân loại vào nhóm $\bigcirc$. Ngẫu nhiên, khi ngưỡng nhất định được cung cấp, trường hợp mà trong đó khoảng cách từ mẫu gần nhất là bằng hoặc lớn hơn ngưỡng có thể được đánh giá là “không có nhóm tương ứng”.

Tuy nhiên, trong ví dụ này, mẫu không được thể hiện đối với nhóm mới (chẳng hạn như, nhóm $\square$). Do đó, dữ liệu đầu vào không được phân loại vào nhóm mới (mặc dù dữ liệu của nhóm $\square$ là đầu vào, dữ liệu không được nhận dạng chính xác).

Về mặt này, mẫu của nhóm mới nằm trong không gian thuộc tính được định rõ bởi phương pháp mà không phụ thuộc vào dữ liệu nhận biết. Chẳng hạn như, trong hệ thống nhận dạng 100, khi hành động của nhóm mới “mang” là đích nhận dạng, và thuộc tính a1 và thuộc tính a2 trên Fig.4A chỉ báo các chuyển động (trạng thái) “di động” và “nằm”, một cách tương ứng, nhóm “mang” có thể được thể hiện là điểm (1,1) nằm trong không gian thuộc tính như được minh họa trên Fig.4B. Như đã nêu trên, kể cả trong trường hợp của nhóm mới, sự nhận dạng có thể được thực hiện nhờ tạo ra mẫu giả nằm trong không gian thuộc tính.

Trong ví dụ này, trong DB dùng cho sự nhận dạng 104, các giá trị tọa

độ của mẫu của nhóm hiện có và mẫu giả của nhóm mới được lưu trữ. Bộ phận nhận dạng 106 tính toán các khoảng cách giữa phép chiếu a của dữ liệu đầu vào x và các mẫu và thực hiện quy trình đưa ra nhãn nhóm của mẫu gần nhất. Khi phương pháp khác với phương pháp lân cận gần nhất được sử dụng, thông số tương ứng được lưu trữ trong DB dùng cho sự nhận dạng 104, và quy trình tương ứng được thực hiện bởi bộ phận nhận dạng 106.

Trong ví dụ này, dữ liệu nhận biết được sử dụng như vốn có đối với mẫu của nhóm hiện có. Tuy nhiên, tương tự như nhóm mới, mẫu giả có thể cũng được tạo ra đối với nhóm hiện có, và mẫu giả có thể được sử dụng kết hợp hoặc riêng lẻ. Ngoài ra, một hoặc nhiều điểm có thể được tạo ra đối với mẫu giả. Chẳng hạn như, trong ví dụ ở trên, các mẫu có cùng số lượng với các mẫu của các nhóm hiện có có thể được tạo ra ngẫu nhiên theo sự phân phối bình thường tập trung vào điểm (1,1). Tại thời điểm đó, ma trận hiệp phương sai của sự phân phối bình thường là siêu tham số và tốt hơn là được quyết định bằng phương pháp chẳng hạn như xác nhận chéo.

Trong ví dụ này, phương pháp được mô tả mà tạo ra mẫu giả (điểm đại diện của mẫu giả) một cách thủ công. Tuy nhiên, mẫu giả có thể được tạo ra bằng phương pháp khác. Như một ví dụ, phương pháp có thể được sử dụng mà được giới thiệu trong tài liệu được biết đến rộng rãi (chẳng hạn như, Socher, R, Ganjoo, M, Manning, C và Y.Ng, A; zero-shot learning Through Cross-Modal Transfer, Advances in Neural Information Processing Systems (NIPS), 2013). Đó là, dữ liệu nhận biết của nhóm hiện có thứ nhất được sử dụng để thu thập thông số θ mà giảm thiểu hàm mục tiêu được biểu diễn theo (biểu thức 1) sau đây.

Phép toán 1

$$J(\Theta) = \sum_{y \in Y_s} \sum_{x^{(i)} \in X_y} \|w_y - \theta^{(2)} f(\theta^{(1)} x^{(i)})\|^2$$

$$Z_1 = \sqrt{R^2 + \left(\omega L - \frac{1}{\omega C}\right)^2} \quad \dots(\text{Biểu thức 1})$$

Ở đây, Y_s là tập hợp của các nhóm hiện có. X_y là tập hợp của dữ liệu đầu vào của nhóm y hoặc kết quả thu được nhờ thực hiện quy trình trích xuất đặc tính định trước trên đó. W_y thu được khi tên nhãn của nhóm y được vector hóa thành cơ sở của không gian W nhờ sử dụng phương pháp hiện có chẳng hạn như word2vec. Sau đây, sự biểu diễn của từ chẳng hạn như tên nhãn của nhóm trong vector được gọi là “nhúng từ”, và vector mà được tạo ra là kết quả và biểu diễn từ được gọi là “vector được nhúng (của từ)”. Ngoài ra, không gian, chẳng hạn như không gian W , mà có vector được nhúng làm cơ sở được gọi là “không gian được nhúng (của từ)”. $\theta^{(1)}$ và $\theta^{(2)}$ là các thông số. f là hàm năng lượng phi tuyến tính chẳng hạn như tanh.

Một cách định tính, trong (biểu thức 1), thông số tại thời điểm đưa ra dữ liệu đầu vào (vector đặc tính của dữ liệu đầu vào) trên không gian W được học nhờ sử dụng dữ liệu nhận biết của nhóm hiện có, và không gian W tương ứng với không gian thuộc tính mà được mô tả trong bản mô tả này. Đầu tiên, mẫu trên không gian thuộc tính của nhóm mới, đó là, không gian W có thể thu được nhờ sử dụng phương pháp hiện có chẳng hạn như word2vec mà được sử dụng khi w_y được chiếu vào không gian W . Cụ thể hơn, tài liệu được biết đến rộng rãi nêu trên được đề cập.

Theo cách này, khi phương pháp nhận biết với một quan sát được sử dụng, quy trình nhận dạng nằm trong không gian thuộc tính có thể cũng được thực hiện trên dữ liệu của nhóm mới nhờ sử dụng phương pháp lân cận gần nhất, chẳng hạn. Tuy nhiên, trong phương pháp thông thường, tất cả thuộc tính được coi tương đương. Do đó, trong trường hợp trong đó quyền ưu tiên của mỗi thuộc tính là khác nhau phụ thuộc vào nhóm của đích nhận dạng, độ chính xác nhận dạng có thể bị làm giảm.

Trong trường hợp được minh họa trên Fig.5, trong dữ liệu để được nhận dạng (sau đây, được gọi là dữ liệu thử nghiệm), chẳng hạn như, thuộc

tính a2 quan trọng hơn thuộc tính a1, và sự phân loại thành nhóm X được coi là phù hợp. Tuy nhiên, khi phương pháp lân cận gần nhất được ứng dụng, dữ liệu thử nghiệm được phân loại vào nhóm O vì $d_1 < d_2$. Lý do của việc phân loại (sự nhận dạng sai) đó là các thuộc tính được coi tương đương để ứng dụng phương pháp lân cận gần nhất mặc dù thuộc tính a2 quan trọng hơn thuộc tính a1 trong trường hợp này.

Chẳng hạn như, trong hệ thống nhận dạng 100, trong trường hợp trong đó mong muốn để nhận dạng hành động “đánh máy” và hành động “uống cà phê”, các thuộc tính liên quan đến đối tượng, chuyển động, và tư thế được sử dụng như được minh họa trong hàng thứ nhất của bảng trên Fig.6, và mẫu giả trong không gian thuộc tính của mỗi nhóm được xác định thủ công như được minh họa trong hàng thứ hai và hàng thứ ba của bảng trên Fig.6.

Vị trí trong đó “0” hoặc “1” lần lượt được ghi tương ứng với trường hợp mà trong đó thuộc tính không được thể hiện hoặc được thể hiện. Trong vị trí trong đó “?” được ghi, thuộc tính có thể được thể hiện hoặc không được thể hiện đối với hành động, và có thể không được coi là quan trọng. Chẳng hạn như, khi uống cà phê, một người uống trong khi đang ngồi hoặc uống trong khi đang đứng. Do đó, thuộc tính liên quan đến tư thế “(có hoặc không) ngồi” không quan trọng trong việc nhận dạng hành động “uống cà phê”.

Kể cả đối với thuộc tính mà không quan trọng, chẳng hạn như, khi giá trị trung gian chẳng hạn như “0”, “1”, và “0,5” được xác định một cách bắt buộc, và khoảng cách nằm trong không gian thuộc tính được tính toán tại thời điểm nhận dạng dữ liệu thử nghiệm, mặc dù thuộc tính quan trọng khác chẳng hạn như “cốc cà phê” phù hợp với định nghĩa của hành động chỉ báo “uống cà phê”, thuộc tính chỉ báo “ngồi” là khác với định nghĩa (mặc dù giá trị của thuộc tính có thể là giá trị bất kỳ trong thực tế), mà kết quả là nhận dạng không chính xác.

Về mặt này, theo phương án này, DB ưu tiên thuộc tính 105 giữ lại

quyền ưu tiên của mỗi thuộc tính đối với mỗi nhóm, và bộ phận nhận dạng 106 thực hiện quy trình nhận dạng theo sự xem xét quyền ưu tiên. Cụ thể hơn, DB ưu tiên thuộc tính 105 giữ lại bảng được thể hiện trên Fig.7 trước đó. Trong ví dụ trên Fig.7, chẳng hạn như, w13, w15, w21, w22, w24, và w26 có thể thiết đặt là “0”, và những cái khác có thể thiết đặt là “1”.

Tại thời điểm nhận dạng dữ liệu thử nghiệm, bộ phận nhận dạng 106 tính toán khoảng cách giữa dữ liệu thử nghiệm và mỗi mẫu và mẫu giả theo (biểu thức 2) sau đây chẳng hạn, và đưa ra nhãn nhóm của mẫu có khoảng cách ngắn nhất là nhãn nhóm của dữ liệu thử nghiệm.

Phép toán 2

$$d^{(k)} = \sqrt{\sum_{i=1}^N w_{y_k i} (a_i^{(k)} - x_i)^2} \quad \dots (\text{Biểu thức 2})$$

Ở đây, $d^{(k)}$ chỉ báo khoảng cách giữa mẫu thử và mẫu thứ k. $w_{y_k i}$ chỉ báo trọng số của thuộc tính i đối với nhóm y_k như được thể hiện trong bảng trên Fig.7. $a_i^{(k)}$ chỉ báo giá trị của thuộc tính i của mẫu thứ k. x_i chỉ báo giá trị của thuộc tính i của dữ liệu thử nghiệm. Khi phép tính được thực hiện như trên, việc gán trọng số tại thời điểm tính toán khoảng cách có thể được thực hiện đối với mỗi nhóm theo quyền ưu tiên của mỗi thuộc tính, để cải thiện độ chính xác nhận dạng.

Đối với phương pháp sử dụng quyền ưu tiên của thuộc tính, chẳng hạn như, phương pháp trong đó chỉ thuộc tính được sử dụng mà có ưu tiên cao hơn ngưỡng định trước, phương pháp trong đó chỉ các thuộc tính k định trước được sử dụng theo thứ tự ưu tiên cao hơn, phương pháp trong đó k được thay đổi từ “1” đến số định trước để thực hiện quy trình nhận dạng một cách tương ứng và để tập hợp các kết quả, và phương pháp trong đó thứ tự đưa ra bộ phận phân biệt dữ liệu được thay đổi theo quyền ưu tiên được cân nhắc cũng như ví dụ trong đó sự gán trọng số được thực hiện trên giá trị của mỗi thuộc tính như được mô tả ở đây. Các chi tiết của phương pháp sử dụng không bị giới hạn.

Như được minh họa trên Fig.7, khi quyền ưu tiên của mỗi thuộc tính được thiết đặt đối với mỗi nhóm, chẳng hạn như, GUI (Graphical User Interface – Giao diện người dùng đồ họa) có thể được chuẩn bị sao cho quyền ưu tiên của mỗi thuộc tính được thiết đặt bởi người dùng. Trong trường hợp này, hệ thống nhận dạng 100 (cấu hình được minh họa trên Fig.1) được thiết đặt là một cấu hình (hệ thống nhận dạng 800) trong đó GUI thiết đặt ưu tiên thuộc tính 801 được bổ sung như được minh họa trên Fig.8. Khi bộ xử lý trung tâm 202 thực thi chương trình ứng dụng, GUI thiết đặt ưu tiên thuộc tính 801 được thực hiện. Một cách ngẫu nhiên, GUI thiết đặt ưu tiên thuộc tính 801 là một ví dụ về bộ phận giao diện người dùng mà được sử dụng khi người dùng thiết đặt quyền ưu tiên được lưu trữ trong bộ phận lưu trữ thứ ba.

Fig.9 là hình vẽ minh họa một ví dụ về GUI để thiết đặt quyền ưu tiên của mỗi thuộc tính đối với mỗi nhóm. Trong hệ thống nhận dạng 800 có GUI này, việc liệu mỗi khớp của con người có uốn cong hay không được thiết đặt làm thuộc tính, và hành động được nhận dạng.

Đầu tiên, người dùng nhập tên nhóm hành động để thiết đặt quyền ưu tiên của thuộc tính vào hộp văn bản 901. Tên nhóm hành động có thể được lựa chọn từ danh sách thả xuống hoặc tương tự thay vì được nhập vào trực tiếp hộp văn bản.

Tiếp theo, người dùng nhấp chuột lựa chọn vị trí khớp 902 mà được coi là quan trọng trong việc nhận dạng hành động của tên nhóm hành động nhập vào trong hộp văn bản 901 trong số các vị trí khớp 902 được chỉ báo bởi dấu ○ trong hình vẽ. Sau khi nhấp chuột, chế độ hiển thị của dấu ○ được thay đổi (chẳng hạn như, một đường xiên được hình thành) để chỉ báo rằng vị trí khớp 902 đã được lựa chọn. Khi vị trí khớp được chọn 902 được nhấp một lần nữa, sự lựa chọn có thể bị hủy.

Khi lựa chọn kết thúc, người dùng nhấp chuột nút thiết đặt 903. Bằng cách này, chẳng hạn như, GUI thiết đặt ưu tiên thuộc tính 801 thiết đặt quyền

ưu tiên của thuộc tính tương ứng với vị trí khớp được chọn 902 là “1” và thiết đặt các quyền ưu tiên khác là “0” và lưu trữ thiết đặt trong DB ưu tiên thuộc tính 105. Một cách ngẫu nhiên, trong trường hợp trong đó giá trị trung gian hoặc tương tự được mong muốn thiết đặt cũng như giá trị “0” hoặc “1”, chẳng hạn như, hộp văn bản có thể được hiển thị gần mỗi khớp, và người dùng có thể nhập vào quyền ưu tiên trong hộp văn bản nhờ sử dụng giá trị số tùy ý. Nếu không, người dùng có thể nhập vào quyền ưu tiên nhờ sử dụng thanh trượt hoặc tương tự.

Theo cách này, kể cả trong trường hợp trong đó không có dữ liệu nhận biết của nhóm mới, hoặc chỉ có một lượng dữ liệu nhận biết cực nhỏ, nhận dạng tự động có thể được thực hiện với độ chính xác cao hơn nhờ xem xét quyền ưu tiên của mỗi thuộc tính đối với mỗi nhóm. Theo cấu hình mô tả ở trên, chẳng hạn như, trong trường hợp trong đó thiết bị chẳng hạn như các thiết bị điện gia dụng và các máy móc công nghiệp được vận hành bằng nhận dạng cử chỉ, người dùng có thể tự do định nghĩa cử chỉ mà là đích của lệnh hoạt động mới tại thời điểm tùy ý, và cử chỉ định nghĩa một lần nữa có thể được nhận dạng với độ chính xác cao. Do đó, sự tiện lợi của người dùng được cải thiện đáng kể.

(2) Phương án thứ hai

Theo hệ thống được mô tả theo phương án thứ nhất, người dùng có thể nhận dạng sự nhận dạng của nhóm mới trong đó không có dữ liệu nhận biết, hoặc chỉ có lượng dữ liệu nhận biết cực nhỏ với độ chính xác cao nhờ xem xét quyền ưu tiên của mỗi thuộc tính đối với mỗi nhóm. Tuy nhiên, điều được xem xét là có các trường hợp mà trong đó người dùng rất khó thiết đặt thủ công quyền ưu tiên của mỗi thuộc tính đối với mỗi nhóm và trường hợp mà trong đó nó cần thời gian kể cả nếu có thể.

Về mặt này, theo phương án này, hệ thống nhận dạng với chức năng thiết đặt tự động ưu tiên thuộc tính được mô tả trong đó hệ thống có thể ước

tính quyền ưu tiên của mỗi thuộc tính đối với mỗi nhóm và có thể thiết đặt quyền ưu tiên một cách tự động. Một cách ngẫu nhiên, trong cùng cấu hình với cấu hình theo phương án thứ nhất, các số chỉ dẫn giống nhau được sử dụng, và phần giải thích của chúng được bỏ qua khi thích hợp.

Theo phương án này, sự nhận dạng có thể được thực hiện với độ chính xác cao mà không khiến người dùng thiết đặt thủ công quyền ưu tiên của mỗi thuộc tính đối với mỗi nhóm.

Fig.10 là hình vẽ minh họa một ví dụ về cấu hình của hệ thống nhận dạng với chức năng thiết đặt tự động ưu tiên thuộc tính (hệ thống nhận dạng 1000) theo phương án này. Hệ thống nhận dạng 1000 bao gồm bộ phận thiết đặt tự động ưu tiên thuộc tính 1001 ngoài cấu hình (hệ thống nhận dạng 100) được minh họa trên Fig.1. Hơn nữa, bộ phận thiết đặt tự động ưu tiên thuộc tính 1001 là một ví dụ về bộ phận thiết đặt đánh giá và thiết đặt quyền ưu tiên được lưu trữ trong bộ phận lưu trữ thứ ba. Một cách ngẫu nhiên, bộ phận thiết đặt tự động ưu tiên thuộc tính 1001 được thực hiện khi bộ xử lý trung tâm 202 thực thi chương trình ứng dụng.

Sau đây, phần mô tả chi tiết sẽ được đưa ra cho quy trình của bộ phận thiết đặt tự động ưu tiên thuộc tính 1001. Phương pháp nhúng từ chẳng hạn như word2vec có thể được sử dụng như một ví dụ. Trong trường hợp này, như được mô tả theo phương án thứ nhất, không gian được nhúng của từ được sử dụng là không gian thuộc tính. Chẳng hạn như, trường hợp được xem xét trong đó nhóm c của đích nhận dạng được nhận dạng nhờ sử dụng thuộc tính $A = \{a_1, a_2, \dots, a_n\}$. Đầu tiên, tên nhãn của nhóm c được nhúng trong không gian được nhúng nhờ sử dụng phương pháp chẳng hạn như word2vec, và vector được nhúng thu được là kết quả được thiết đặt là x_c . Hơn nữa, tương tự, các tên thuộc tính chỉ báo $a_1, a_2, \dots, a_n$ được nhúng trong không gian được nhúng, và vector được nhúng thu được làm kết quả được thiết đặt là $X = \{x_1, x_2, \dots, x_n\}$.

Ở đây, điều được xem xét là x_c được biểu diễn nhờ sử dụng sự kết hợp

tuyến tính của X. Nó thu được khi hàm mục tiêu được biểu diễn sau đây (biểu thức 3) được giảm thiểu đối với $W = \{w_1, w_2, \dots, w_n\}$.

Phép toán 3

$$\left(x_c - \sum_{i=1}^n w_i x_i\right)^2 + \lambda \|W\|_p \dots (\text{Biểu thức 3})$$

Tuy nhiên, $\|W\|_p$ của số hạng thứ hai của (biểu thức 3) là định mức chính quy được biểu diễn sau đây (biểu thức 4).

Phép toán 4

$$\|W\|_p = \left(|w_1|^p + |w_2|^p + \dots + |w_n|^p\right)^{1/p} \dots (\text{Biểu thức 4})$$

Chẳng hạn như, “2” được sử dụng làm giá trị của p. Cụ thể là, trong trường hợp trong đó nghiệm rải rác được mong muốn thu được, đó là, trong trường hợp trong đó nghiệm là “0” khác với số lượng nhỏ các yếu tố trong W được mong muốn thu được, $p=1$ hoặc tương tự có thể được sử dụng. λ là siêu tham số và được quyết định bằng phương pháp chẳng hạn như xác nhận chéo.

W thu được theo cách này có thể được hiểu là quyền ưu tiên của mỗi thuộc tính để biểu diễn nhãn nhóm. Do đó, bộ phận thiết đặt tự động ưu tiên thuộc tính 1001 lưu trữ (các hồ sơ) giá trị trong DB ưu tiên thuộc tính 105. Hơn nữa, bộ phận thiết đặt có thể sử dụng không gian được nhúng của từ làm không gian biểu diễn trung gian, và có thể ước tính, là quyền ưu tiên của mỗi sự biểu diễn trung gian, hệ số của mỗi vector tại thời điểm biểu diễn vector được nhúng biểu diễn nhóm của đích xử lý là sự kết hợp tuyến tính của vector được nhúng của sự biểu diễn trung gian trong không gian được nhúng.

Khi cần, sự chuẩn hóa có thể được thực hiện sao cho mỗi yếu tố của W trở thành giá trị giữa “0” và “1”, hoặc điều kiện hạn chế được biểu diễn sau đây (biểu thức 5) có thể được bổ sung khi vấn đề giảm thiểu của (biểu thức 3) được giải quyết.

Phép toán 5

$$\sum_{i=1}^n w_i = 1 \quad \dots(\text{Biểu thức 5})$$

$$\forall w_i \in W, w_i \geq 0$$

Phương pháp mà sử dụng nhóm hiện có nằm trong không gian thuộc tính, mẫu của nhóm mới, và mẫu giả và đánh giá quyền ưu tiên của mỗi thuộc tính đối với mỗi nhóm được coi là phương pháp thực hiện khác của bộ phận thiết đặt tự động ưu tiên thuộc tính 1001. Như một ví dụ, khi thuộc tính đơn được sử dụng, khả năng phân tách nhóm của khoảng cách mỗi nhóm có thể được tách ra từ nhóm khác được tính toán được sử dụng làm quyền ưu tiên của thuộc tính. Hơn nữa, bộ phận thiết đặt có thể ước tính hiệu quả phân tách nhóm của mỗi sự biểu diễn trung gian đối với nhóm của đích xử lý làm quyền ưu tiên của mỗi sự biểu diễn trung gian.

Chẳng hạn như, khả năng phân tách nhóm của thuộc tính j đối với nhóm k quyết định các thông số $a^{(j)}$ và $b^{(j)}$ khi mẫu và mẫu giả $X = \{x_1, x_2, \dots, x_n\}$ và nhãn của chúng $Y = \{y_1, y_2, \dots, y_n\}$ được sử dụng để giải quyết vấn đề tối ưu hóa của (biểu thức 6) sau đây đầu tiên.

Phép toán 6

$$\min_{a^{(j)}, b^{(j)}, \xi^{(j)}} \frac{1}{2} a^{(j)2} + C \sum_{i=1}^n \xi_i^{(j)}$$

$$s.t. \quad y_i^{(k)} (a^{(j)} x_i^{(j)} + b^{(j)}) \geq 1 - \xi_i^{(j)}, \quad \dots(\text{Biểu thức 6})$$

$$\xi_i^{(j)} \geq 0$$

$$(i = 1, 2, \dots, n)$$

Sau đó, khả năng phân tách nhóm có thể được tính toán bởi (biểu thức 7) sau đây. Ở đây, $x_i^{(j)}$ là giá trị của thuộc tính thứ j của mẫu (giả) thứ i . $y_i^{(k)}$ là biến nhị phân mà là “1” trong trường hợp trong đó mẫu thứ i giữ nhãn của nhóm k và là “-1” trong các trường hợp khác. C là siêu tham số.

Phép toán 7

$$\frac{1}{n} \sum_{i=1}^n f(y_i^{(k)} (ax_i + b)) \quad \dots(\text{Biểu thức 7})$$

$$f(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases}$$

Theo quy trình mô tả ở trên, sự nhận dạng có thể được thực hiện với độ chính xác cao mà không khiến người dùng thiết đặt thủ công quyền ưu tiên của mỗi thuộc tính đối với mỗi nhóm.

(3) Phương án thứ ba

Theo hệ thống được mô tả theo phương án thứ nhất và phương án thứ hai, sự nhận dạng của nhóm mới trong đó không có dữ liệu nhận biết, hoặc chỉ có lượng dữ liệu nhận biết cực nhỏ có thể được thực hiện với độ chính xác cao nhờ xem xét quyền ưu tiên của mỗi thuộc tính đối với mỗi nhóm. Tuy nhiên, độ chính xác nhận dạng, mà người dùng mong muốn, đối với nhóm mới không thể thu được chỉ bằng thuộc tính hiện có. Chẳng hạn như, trong hệ thống nhận dạng hành động, chỉ hành động chỉ liên quan bởi thân phía trên được xử lý, và thuộc tính liên quan đến chuyển động của thân phía trên được sử dụng làm thuộc tính. Tuy nhiên, trong trường hợp trong đó hành động đi kèm với chuyển động của thân phía dưới được coi là nhóm mới, nhóm mới rất khó để nhận dạng với độ chính xác đủ chỉ trong thuộc tính liên quan đến chuyển động của thân phía trên hiện có. Tại điểm này, trong hệ thống thông thường, chỉ số xác định không được cung cấp về mức độ hiệu quả của thuộc tính hiện có trong việc nhận dạng của mỗi nhóm. Do đó, rất khó để đánh giá khách quan tình huống như vậy.

Về mặt này, theo phương án này, hệ thống nhận dạng với chức năng bổ sung thuộc tính được mô tả trong đó hệ thống đánh giá tình huống mô tả ở trên nhờ sử dụng quyền ưu tiên của mỗi thuộc tính, và thuộc tính mới được thiết đặt khi cần để cải thiện hơn nữa độ chính xác nhận dạng. Một cách ngẫu nhiên, trong cùng cấu hình với cấu hình theo phương án thứ nhất, các số chỉ dẫn giống nhau được sử dụng, và phân giải thích của chúng được bỏ qua khi thích hợp.

Fig.11 là hình vẽ minh họa một ví dụ về cấu hình của hệ thống nhận dạng với chức năng bổ sung thuộc tính (hệ thống nhận dạng 1100) theo

phương án này. Hệ thống nhận dạng 1100 bao gồm bộ phận bổ sung thuộc tính 1101 ngoài cấu hình (hệ thống nhận dạng 100) được minh họa trên Fig.1. Một cách ngẫu nhiên, bộ phận bổ sung thuộc tính 1101 được thực hiện khi bộ xử lý trung tâm 202 thực thi chương trình ứng dụng.

Sau đây, phần mô tả sẽ được đưa ra về quy trình của bộ phận bổ sung thuộc tính 1101. Bộ phận bổ sung thuộc tính 1101 tính toán quyền ưu tiên của mỗi thuộc tính hiện có đối với nhóm mới nhờ sử dụng phương pháp được mô tả theo phương án thứ nhất và phương án thứ hai. Hơn nữa, chẳng hạn như, trong trường hợp trong đó sự ưu tiên được tính toán bất kỳ thấp hơn ngưỡng định trước, bộ phận bổ sung thuộc tính 1101 đánh giá rằng nhóm mới không thể được nhận dạng với độ chính xác cao trong các thuộc tính hiện có và bổ sung thuộc tính mới. Hơn nữa, bộ phận bổ sung thuộc tính 1101 là một ví dụ về bộ phận bổ sung bổ sung sự biểu diễn trung gian dựa trên quyền ưu tiên của mỗi sự biểu diễn trung gian đối với nhóm mới.

Khi thuộc tính mới được bổ sung, giao diện để người dùng thiết đặt thủ công thuộc tính mới có thể được chuẩn bị, hoặc thuộc tính mới có thể được thiết đặt một cách tự động.

Chẳng hạn như, phương pháp sau đây được coi là phương pháp thiết đặt một cách tự động. Đó là, bộ phận bổ sung thuộc tính 1101 sử dụng không gian được nhúng từ làm không gian thuộc tính, thực hiện PCA (Principal Component Analysis – Phân tích thành phần chính) nhờ sử dụng vector được nhúng của mỗi thuộc tính, và trích xuất vector đặc trưng đối với giá trị đặc trưng nhỏ nhất. Vector đặc trưng là hướng trong đó sự phân tán vector được nhúng của thuộc tính hiện có là nhỏ nhất, và do đó có thể được coi là hướng bao gồm thông tin mà không được bao gồm trong thuộc tính hiện có.

Về mặt này, bộ phận bổ sung thuộc tính 1101 sử dụng, làm thuộc tính mới, từ gần nhất với vector đặc trưng trong số các từ mà là các ứng viên của thuộc tính được cung cấp trước đó. Chẳng hạn như, nó có thể được đánh giá

dựa trên tính tương đồng cosin giữa vector được nhúng của từ mà là ứng viên của thuộc tính và vector đặc trưng. Một cách ngẫu nhiên, thuộc tính mới được bổ sung theo cách này có thể được sử dụng cho sự nhận dạng của nhóm hiện có, hoặc trọng số của thuộc tính được bổ sung một lần nữa có thể thiết đặt là “0” để không ảnh hưởng độ chính xác nhận dạng đối với nhóm hiện có.

Như đã nêu trên, nhờ bổ sung thuộc tính theo hướng không được bổ sung vào không gian thuộc tính hiện có, thuộc tính được bổ sung mới có thể được sử dụng kể cả trong trường hợp trong đó nhóm mới không thể được nhận dạng với độ chính xác cao trong thuộc tính hiện có, nhờ đó cải thiện độ chính xác.

(4) Phương án thứ tư

Theo phương án này, cụ thể là, phần mô tả chi tiết sẽ được đưa ra cho ví dụ về phương pháp thiết đặt quyền ưu tiên đối với đối tượng trong trường hợp trong đó hành vi của con người được coi là đích nhận dạng, và khả năng có mặt của đối tượng được định trước xung quanh người dùng mà hoạt động được mong muốn để được nhận dạng được sử dụng như một trong số các thuộc tính. Khả năng có mặt của đối tượng được tính toán nhờ sử dụng hình ảnh hoặc video theo góc nhìn người thứ nhất của người dùng mà có thể thu được từ thiết bị camera dạng kính hoặc thiết bị tương tự. Ngoài ra, theo phương án này, giả sử rằng mẫu trên không gian thuộc tính của nhóm mới là không phải mẫu giả và thu được từ lượng nhỏ của dữ liệu thực tế (đó là, sự thiết đặt của nhận biết với một quan sát hoặc nhận biết với một vài quan sát).

Theo các nghiên cứu trước đó, được nhận thấy rằng đối tượng hiện có gần đó có thông tin quan trọng để nhận dạng hành động. Chẳng hạn như, để nhận dạng hành động “đọc sách”, thông tin về việc liệu đối tượng “sách” có được thể hiện hay không là quan trọng cũng như hình dáng (tư thế) của bàn tay cầm sách hoặc chuyển động thân tương đối thấp, mà là rõ ràng. Tương tự, để nhận dạng hành động “gửi SMS (Short Message Service – Dịch vụ tin nhắn

ngắn)), sự có mặt của đối tượng “điện thoại di động” là quan trọng cũng như thông tin về chuyển động của ngón tay. Về mặt này, chẳng hạn như, trong trường hợp trong đó các hành động đó được định nghĩa là nhóm mới, hiển nhiên là sử dụng sự có mặt hoặc khả năng có mặt của đối tượng “sách” và “điện thoại di động” làm thuộc tính.

Tuy nhiên, khi mẫu (dữ liệu nhận biết) của hành động của các nhóm mới thu được, chẳng hạn như, khi dữ liệu được thu thập trong hoàn cảnh chẳng hạn như văn phòng, và dữ liệu mẫu của nhóm hành động “đọc sách” được thu thập, “điện thoại di động” cũng có thể nằm trong khung hình ảnh ngoài đối tượng “sách”. Trong trường hợp này, hành động “đọc sách” được xác định là sự có mặt (khả năng có mặt cao) của cả hai đối tượng “sách” và “điện thoại di động”. Nói cách khác, mặc dù tất cả các thuộc tính khác khớp hành động “đọc sách”, hành động “đọc sách” có thể không được nhận dạng khi “điện thoại di động” không được thể hiện.

Vấn đề phát sinh khi, mặc dù thực tế là quyền ưu tiên của thuộc tính “điện thoại di động” là thấp đối với hành động “đọc sách”, quy trình nhận dạng được thực hiện mà không xem xét thực tế.

Nói chung, điều được xem xét là sự quan sát được hướng tới đối tượng có sự liên quan cao với hành động khi người dùng thực hiện hành động nhất định. Chẳng hạn như, trong trường hợp “đọc sách”, sự quan sát được hướng tới “sách”, và trong trường hợp “gửi SMS”, sự quan sát được hướng tới “điện thoại di động”. Vì lý do này, điều được xem xét là sự quan sát dễ dàng bị thu hút tới đối tượng quan trọng cho hành động.

Về mặt này, theo phương án này, phần mô tả được đưa ra về hệ thống nhận dạng với chức năng thiết đặt ưu tiên thuộc tính sử dụng thông tin nhìn tập trung mà có thể thiết đặt sự có mặt hoặc khả năng có mặt của đối tượng, đó là, quyền ưu tiên của thuộc tính nhờ sử dụng thông tin quan sát (chẳng hạn như sự có mặt của sự quan sát và khả năng nhìn tập trung) của người dùng. Một cách

ngẫu nhiên, trong cùng cấu hình với cấu hình theo phương án thứ nhất, các số chỉ dẫn giống nhau được sử dụng, và phần giải thích của chúng được bỏ qua khi thích hợp.

Fig.12 là hình vẽ minh họa một ví dụ về cấu hình của hệ thống nhận dạng với chức năng thiết đặt ưu tiên thuộc tính sử dụng thông tin nhìn tập trung (hệ thống nhận dạng 1200) theo phương án này. Hệ thống nhận dạng 1200 bao gồm bộ phận thiết đặt ưu tiên thuộc tính sử dụng thông tin nhìn tập trung 1201 ngoài cấu hình (hệ thống nhận dạng 100) được minh họa trên Fig.1. Hơn nữa, bộ phận thiết đặt ưu tiên thuộc tính sử dụng thông tin nhìn tập trung 1201 là một ví dụ về bộ phận thiết đặt mà thiết đặt quyền ưu tiên theo đối tượng dựa trên thông tin quan sát của người dùng đối với đối tượng. Một cách ngẫu nhiên, bộ phận thiết đặt ưu tiên thuộc tính sử dụng thông tin nhìn tập trung 1201 được thực hiện khi bộ xử lý trung tâm 202 thực thi chương trình ứng dụng.

Sau đây, quy trình của bộ phận thiết đặt ưu tiên thuộc tính sử dụng thông tin nhìn tập trung 1201 sẽ được mô tả. Giả sử rằng dữ liệu video cho tổng khung F thu được trong khi dữ liệu nhận biết của nhóm mới được thu thập. Trong trường hợp này, nó có thể được coi là khung F tương ứng với một phần của dữ liệu hành động (nhận biết với một quan sát), hoặc nó có thể được coi là dữ liệu video của tổng khung F được thu thập trong một số phần dữ liệu hành động (nhận biết với một vài quan sát). Bộ phận thiết đặt ưu tiên thuộc tính sử dụng thông tin nhìn tập trung 1201 tính toán quyền ưu tiên I_j của thuộc tính j bởi (biểu thức 8) sau đây nhờ sử dụng dữ liệu của khung F.

Phép toán 8

$$I_j = \frac{1}{F} \sum_{i=1}^F f(o_i^{(j)}) \dots (\text{Biểu thức 8})$$

Ở đây, $f(o_i^{(j)})$ là hàm sao cho trường hợp mà trong đó điểm nhìn tập trung của người dùng là trên đối tượng j được thiết đặt là “1”, và các trường

hợp khác được thiết đặt là “0” trong khung i. Vị trí của điểm nhìn tập trung có thể thu được nhờ sử dụng sản phẩm làm sẵn chẳng hạn như kính theo dõi mắt. Ngoài ra, khi cần, I_j có thể được chuẩn hóa sao cho giá trị tối đa là “1”, và giá trị tối thiểu là “0”.

Phần tiếp theo mô tả phương pháp trích xuất thông tin điểm nhìn tập trung chỉ từ hình ảnh hoặc video theo góc nhìn người thứ nhất nhờ sử dụng sản phẩm mà có thể thu được hình ảnh hoặc video theo điểm quan sát người thứ nhất với chi phí thấp hơn mà không sử dụng sản phẩm đặc biệt để thu thông tin điểm nhìn tập trung chẳng hạn như kính theo dõi mắt.

Fig.13 là hình vẽ minh họa một ví dụ về cấu hình của hệ thống nhận dạng với chức năng thiết đặt ưu tiên thuộc tính sử dụng thông tin nhìn tập trung (hệ thống nhận dạng 1300) theo phương án này. Như được minh họa trên Fig.13, hệ thống nhận dạng 1300 bao gồm bộ phận đánh giá nhìn tập trung 1301. Một cách ngẫu nhiên, bộ phận đánh giá nhìn tập trung 1301 được thực hiện khi bộ xử lý trung tâm 202 thực thi chương trình ứng dụng.

Fig.14 là hình vẽ minh họa một ví dụ về cấu hình của bộ phận đánh giá nhìn tập trung 1301. Bộ phận đánh giá nhìn tập trung 1301 thực hiện đánh giá nhìn tập trung là việc nhận biết đa nhiệm với nhận dạng đối tượng nhờ sử dụng CNN, ví dụ. Hơn nữa, bộ phận đánh giá nhìn tập trung 1301 là một ví dụ về bộ phận đánh giá thực hiện việc nhận biết đa nhiệm trên đặc tả của đối tượng và sự đánh giá nhìn tập trung dựa trên hình ảnh hoặc video theo điểm quan sát người thứ nhất của người dùng và đánh giá thông tin quan sát của người dùng.

Bộ phận đánh giá nhìn tập trung 1301 xử lý hàng của khung hình ảnh mà là dữ liệu đầu vào trong mạng gốc mà là môđun chung giữa nhận dạng đối tượng và đánh giá nhìn tập trung. Một cách định tính, mạng gốc có thể được hiểu là thực hiện sự trích xuất đặc tính mức thấp. Sau đó bộ phận đánh giá nhìn tập trung 1301 đưa ra bản đồ đặc tính mà là đầu ra của mạng gốc đến mạng nhận dạng đối tượng và mạng đánh giá nhìn tập trung làm đầu vào.

Trong mạng nhận dạng đối tượng, bộ phận đánh giá nhìn tập trung 1301 đưa ra khả năng có mặt của mỗi đối tượng và vị trí có mặt của mỗi đối tượng trong mỗi khung làm kết quả ước tính. Ngoài ra, trong mạng đánh giá nhìn tập trung, bộ phận đánh giá nhìn tập trung 1301 nhập vào khả năng có mặt của mỗi đối tượng và vị trí có mặt của mỗi đối tượng trong khung mà là đầu ra (kết quả) của mạng nhận dạng đối tượng ngoài bản đồ đặc tính mà là đầu ra của mạng gốc. Một cách định tính, ngoài đặc tính hình ảnh, thông tin của đối tượng mà dễ thu hút sự quan sát có thể được học để kỳ vọng độ chính xác của đánh giá nhìn tập trung được cải thiện.

Phương pháp điều chỉnh các thông số cùng với mạng gốc, mạng nhận dạng đối tượng, và mạng đánh giá nhìn tập trung sẽ được mô tả như sự chuẩn bị sơ bộ.

Trong mạng nhận dạng đối tượng, bộ phận đánh giá nhìn tập trung 1301 có thể thu được gradien tổn thất (giá trị chỉ báo sự khác biệt từ giá trị chính xác) được biểu diễn bởi $L_{obj} = L_{class} + \alpha L_{loc}$ bởi thuật toán xác định trước (chẳng hạn như, phương pháp lan truyền lỗi nghịch đảo) và có thể cập nhật trọng số (thông số trong mạng nhận dạng đối tượng) của mạng. Trong mạng đánh giá nhìn tập trung, bộ phận đánh giá nhìn tập trung 1301 có thể cập nhật trọng số (thông số trong mạng đánh giá nhìn tập trung) của mạng nhờ thu gradien tổn thất được biểu diễn bởi βL_{gaze} bởi thuật toán xác định trước (chẳng hạn như, phương pháp lan truyền lỗi nghịch đảo).

Ở đây, L_{class} là tổn thất liên quan đến phân loại nhóm của đối tượng, và chẳng hạn như, hàm trung bình mũ có thể được sử dụng. L_{loc} là tổn thất liên quan đến định vị của đối tượng. Chẳng hạn như, khi vị trí của đối tượng được biểu diễn bởi tọa độ trung tâm của hộp giới hạn và chiều rộng và chiều cao của chúng, sai số bình phương từ giá trị chính xác (thực địa) của chúng có thể được sử dụng làm L_{loc} . Chẳng hạn như, khoảng cách Öclit giữa vị trí nhìn tập trung ước tính và vị trí nhìn tập trung của giá trị chính xác thu được nhờ sử

dụng kính theo dõi mắt, và tương tự có thể được sử dụng làm L_{gaze} . α và β là các thông số để thay đổi trọng số của mỗi tổn thất, và giá trị tối ưu có thể được quyết định nhờ xác nhận chéo hoặc tương tự.

Trong mạng gốc, bộ phận đánh giá nhìn tập trung 1301 thu gradient của tổng $L = L_{class} + \alpha L_{loc} + \beta L_{gaze}$ của các tổn thất bởi thuật toán xác định trước (chẳng hạn như, phương pháp lan truyền lỗi nghịch đảo) và cập nhật trọng số (thông số trong mạng gốc) của mạng.

Một cách ngẫu nhiên, bản đồ khả năng mà chỉ báo khả năng của sự có mặt của điểm nhìn tập trung tại mỗi vị trí trong khung có thể được sử dụng làm đầu ra (thông tin quan sát) của mạng đánh giá nhìn tập trung.

Khi đối tượng nhất định được nhận biết, trong trường hợp trong đó bàn tay được thể hiện trong khung, mạng đánh giá vị trí của bàn tay có thể được bổ sung theo sự xem xét sự dễ dàng của việc thu hút sự quan sát đến đối tượng. Trong trường hợp này, như được minh họa trên Fig.15, mạng ước tính bàn tay được bổ sung vào mạng của Fig.14, và đầu ra được nhập vào mạng đánh giá nhìn tập trung. Hơn nữa, nhờ còn sử dụng khả năng có mặt của bàn tay hoặc sự có mặt của bàn tay trong khung hình ảnh và thông tin của vị trí của bàn tay trong khung trong hình ảnh hoặc video theo điểm quan sát người thứ nhất của người dùng, bộ phận đánh giá có thể thực hiện việc nhận biết đa nhiệm trên đặc tả của đối tượng và sự đánh giá của sự quan sát và có thể ước tính thông tin quan sát của người dùng.

Đối với việc điều chỉnh thông số, trong mạng ước tính bàn tay, bộ phận đánh giá nhìn tập trung 1301 thu gradient tổn thất γL_{hand} bởi thuật toán xác định trước (chẳng hạn như, phương pháp lan truyền lỗi nghịch đảo) và cập nhật trọng số (thông số trong mạng ước tính bàn tay) của mạng. Trong trường hợp này, trong mạng gốc, bộ phận đánh giá nhìn tập trung 1301 thu chênh lệch của $L = L_{class} + \alpha L_{loc} + \beta L_{gaze} + \gamma L_{hand}$ bởi thuật toán xác định trước (chẳng hạn như, phương pháp lan truyền lỗi nghịch đảo) và cập nhật trọng số (thông số trong

mạng gốc) của mạng.

Chẳng hạn như, sai số bình phương giữa tọa độ trung tâm của hộp giới hạn chỉ báo vị trí ước tính của bàn tay và chiều rộng và chiều cao của chúng và tọa độ trung tâm của hộp giới hạn của giá trị chính xác và chiều rộng và chiều cao của chúng có thể được sử dụng làm L_{hand} . γ là thông số để điều chỉnh trọng số L_{hand} .

(5) Phương án khác

Một cách ngẫu nhiên, theo phương án mô tả ở trên, trường hợp mà trong đó sáng chế được ứng dụng vào hệ thống nhận dạng (100, 800, 1000, 1100, 1200, và 1300) đã được mô tả. Tuy nhiên, sáng chế không giới hạn ở đó, và có thể được ứng dụng rộng rãi vào các hệ thống đặc tả khác nhau.

Theo phương án mô tả ở trên, phần mô tả đã được đưa ra về trường hợp mà trong đó hệ thống nhận dạng 100 được thực hiện bởi một máy tính 200. Tuy nhiên, sáng chế không giới hạn, và hệ thống nhận dạng 100 có thể được thực hiện bởi các máy tính. Trong trường hợp này, chẳng hạn như, DB dùng cho sự tính toán thuộc tính 102, bộ phận tính toán thuộc tính 103, DB dùng cho sự nhận dạng 104, DB ưu tiên thuộc tính 105, và bộ phận nhận dạng 106 được bố trí thích hợp trong máy tính.

Trước đây, sáng chế được mô tả một cách chi tiết dựa vào các hình vẽ đi kèm. Tuy nhiên, sáng chế không giới hạn ở cấu hình cụ thể này và bao gồm nhiều thay đổi rõ ràng và các cấu hình tương đương mà nằm trong nội dung của các điểm yêu cầu bảo hộ kèm theo. Chẳng hạn như, các phương án được mô tả ở trên được mô tả một cách chi tiết để giải thích sáng chế một cách rõ ràng, và sáng chế không nhất thiết phải giới hạn ở cấu trúc được đề xuất với tất cả các cấu hình được giải thích. Ngoài ra, một phần của cấu hình của phương án có thể được thay thế với cấu hình của phương án khác, và cấu hình của phương án khác có thể được bổ sung vào cấu hình của phương án này. Ngoài

ra, sự bổ sung, sự loại bỏ, và thay thế của cấu hình khác có thể được thực hiện trên một phần của cấu hình theo mỗi phương án.

Một phần hoặc tất cả các cấu hình, chức năng, bộ phận xử lý, phương thức xử lý, hoặc tương tự được mô tả trên đây có thể được thực hiện bởi phần cứng nhờ thiết kế với mạch tích hợp, chẳng hạn. Tương tự, các cấu hình, chức năng, hoặc tương tự có thể được thực hiện bởi phần mềm khi bộ xử lý thực thi chương trình thực hiện mỗi trong số các chức năng. Thông tin chẳng hạn như chương trình, bảng, hoặc tệp để thực hiện các chức năng có thể được lưu trữ trong thiết bị lưu trữ chẳng hạn như thiết bị lưu trữ (bộ nhớ), đĩa cứng, hoặc ổ cứng trạng thái rắn (solid state drive - SSD) hoặc vật ghi chẳng hạn như thẻ IC, thẻ SD, hoặc DVD.

Chỉ các tuyến điều khiển hoặc các tuyến thông tin mà được xem xét là cần thiết đối với mục đích của bản mô tả được minh họa, và hầu hết các tuyến điều khiển và các tuyến thông tin được bao gồm trong các sản phẩm thực tế không cần thiết được minh họa. Trong thực tế, hầu hết tất cả các cấu hình có thể được coi là kết nối với nhau.

Cấu hình mô tả ở trên có thể được thay đổi, bố trí lại, tập hợp, và bỏ qua khi thích hợp trong phạm vi mà không trệtch khỏi bản chất của sáng chế.

YÊU CẦU BẢO HỘ

1. Hệ thống đặc tả, khác biệt ở chỗ bao gồm:

bộ phận lưu trữ thứ nhất lưu trữ thông số để đưa ra đích xử lý trong không gian biểu diễn trung gian có mỗi sự biểu diễn trung gian chỉ báo mỗi thuộc tính của đích xử lý là mỗi kích cỡ;

bộ phận tính toán đưa ra dữ liệu đầu vào thu được từ cảm biến định trước hoặc trị số thu được nhờ thực hiện quy trình tiền xử lý định trước trên dữ liệu đầu vào trong không gian biểu diễn trung gian dựa trên thông số được lưu trữ trong bộ phận lưu trữ thứ nhất;

bộ phận lưu trữ thứ hai lưu trữ thông số mà được sử dụng khi đích xử lý được định rõ trong không gian biểu diễn trung gian;

bộ phận lưu trữ thứ ba lưu trữ quyền ưu tiên của mỗi sự biểu diễn trung gian chỉ báo mỗi thuộc tính của đích xử lý tương ứng với đích xử lý; và

bộ phận ấn định định rõ đích xử lý trong dữ liệu đầu vào dựa trên thông số được lưu trữ trong bộ phận lưu trữ thứ hai, quyền ưu tiên được lưu trữ trong bộ phận lưu trữ thứ ba, và đầu ra của bộ phận tính toán,

hệ thống đặc tả còn bao gồm bộ phận bổ sung bổ sung sự biểu diễn trung gian dựa trên quyền ưu tiên của mỗi sự biểu diễn trung gian đối với nhóm mới.

2. Hệ thống đặc tả theo điểm 1, khác biệt ở chỗ:

đích xử lý mà được định rõ bởi bộ phận ấn định là hành động của con người và/hoặc chuyển động của con người.

3. Hệ thống đặc tả theo điểm 2, khác biệt ở chỗ:

sự có mặt của đối tượng được định trước hoặc khả năng có mặt của đối tượng xung quanh người dùng mà hành động được mong muốn được định rõ của nó được bao gồm như một trong số mỗi sự biểu diễn trung gian của mỗi

thuộc tính của đích xử lý.

4. Hệ thống đặc tả theo điểm 2, khác biệt ở chỗ:

sự xuất hiện hoặc khả năng xuất hiện của chuyển động hoặc tư thế của yếu tố cơ bản định hình hành động của con người được bao gồm như một trong số mỗi sự biểu diễn trung gian chỉ báo mỗi thuộc tính của đích xử lý.

5. Hệ thống đặc tả theo điểm 3, khác biệt ở chỗ bao gồm:

bộ phận thiết đặt mà tính toán sự có mặt của đối tượng hoặc khả năng có mặt của đối tượng dựa trên hình ảnh hoặc video theo góc nhìn người thứ nhất của người dùng.

6. Hệ thống đặc tả theo điểm 1, khác biệt ở chỗ bao gồm:

bộ phận giao diện người dùng để cho phép người dùng thiết đặt quyền ưu tiên được lưu trữ trong bộ phận lưu trữ thứ ba.

7. Hệ thống đặc tả theo điểm 1, khác biệt ở chỗ bao gồm:

bộ phận thiết đặt đánh giá và thiết đặt quyền ưu tiên được lưu trữ trong bộ phận lưu trữ thứ ba.

8. Hệ thống đặc tả theo điểm 7, khác biệt ở chỗ:

bộ phận thiết đặt sử dụng không gian được nhúng của từ làm không gian biểu diễn trung gian, và đánh giá, làm quyền ưu tiên của mỗi sự biểu diễn trung gian, hệ số của mỗi vectơ ở thời điểm biểu diễn vectơ được nhúng biểu diễn nhóm của đích xử lý trong không gian được nhúng làm sự kết hợp tuyến tính của vectơ được nhúng của sự biểu diễn trung gian.

9. Hệ thống đặc tả theo điểm 7, khác biệt ở chỗ:

bộ phận thiết đặt đánh giá hiệu quả phân tách nhóm của mỗi sự biểu diễn trung gian đối với nhóm của đích xử lý làm quyền ưu tiên của sự biểu diễn trung gian.

10. Hệ thống đặc tả theo điểm 5, khác biệt ở chỗ:

bộ phận thiết đặt thiết đặt quyền ưu tiên liên quan đến đối tượng dựa trên thông tin quan sát của người dùng đối với đối tượng.

11. Hệ thống đặc tả theo điểm 10, khác biệt ở chỗ bao gồm:

bộ phận đánh giá thực hiện việc nhận biết đa nhiệm về đặc tả của đối tượng và sự đánh giá của sự quan sát dựa trên hình ảnh hoặc video theo điểm quan sát người thứ nhất của người dùng và đánh giá thông tin quan sát của người dùng.

12. Hệ thống đặc tả theo điểm 11, khác biệt ở chỗ:

bộ phận đánh giá thực hiện việc nhận biết đa nhiệm trên đặc tả của đối tượng và sự đánh giá của sự quan sát nhờ còn sử dụng thông tin về khả năng có mặt của bàn tay hoặc sự có mặt của bàn tay trong khung hình ảnh và vị trí của bàn tay trong khung trong hình ảnh hoặc video theo điểm quan sát người thứ nhất của người dùng và đánh giá thông tin quan sát của người dùng.

13. Phương pháp đặc tả trong hệ thống đặc tả, khác biệt ở chỗ:

hệ thống đặc tả bao gồm bộ phận lưu trữ thứ nhất lưu trữ thông số để đưa ra đích xử lý trong không gian biểu diễn trung gian có mỗi sự biểu diễn trung gian chỉ báo mỗi thuộc tính của đích xử lý là mỗi kích cỡ, bộ phận lưu trữ thứ hai lưu trữ thông số mà được sử dụng khi đích xử lý được định rõ trong không gian biểu diễn trung gian, và bộ phận lưu trữ thứ ba lưu trữ quyền ưu tiên của mỗi sự biểu diễn trung gian chỉ báo mỗi thuộc tính của đích xử lý tương ứng với đích xử lý,

phương pháp đặc tả trong hệ thống đặc tả bao gồm:

bước thứ nhất, trong đó bộ phận tính toán đưa ra dữ liệu đầu vào thu được từ cảm biến định trước hoặc trị số thu được nhờ thực hiện quy trình tiền xử lý định trước trên dữ liệu đầu vào trong không gian biểu diễn trung gian dựa trên thông số được lưu trữ trong bộ phận lưu trữ thứ nhất; và

bước thứ hai, trong đó bộ phận ấn định định rõ đích xử lý trong dữ liệu đầu vào dựa trên thông số được lưu trữ trong bộ phận lưu trữ thứ hai, quyền ưu tiên được lưu trữ trong bộ phận lưu trữ thứ ba, và đầu ra của bộ phận tính toán,

phương pháp đặc tả trong hệ thống đặc tả còn bao gồm bước bổ sung bổ sung sự biểu diễn trung gian dựa trên quyền ưu tiên của mỗi sự biểu diễn trung gian đối với nhóm mới.

FIG. 1

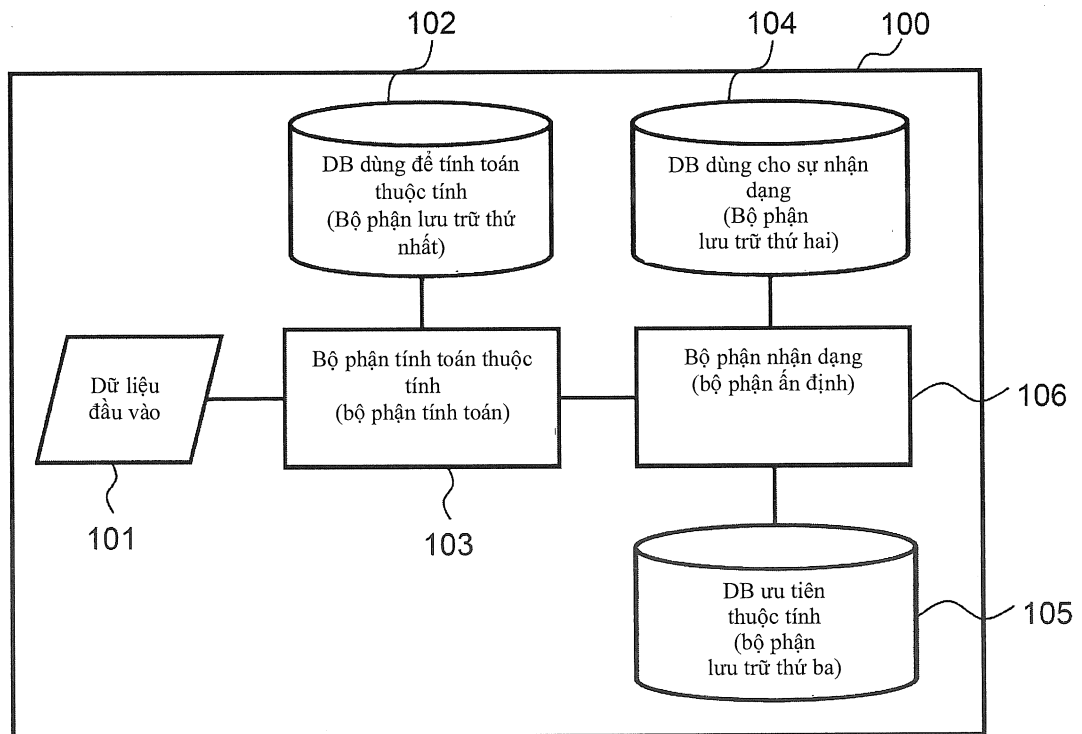


FIG. 2

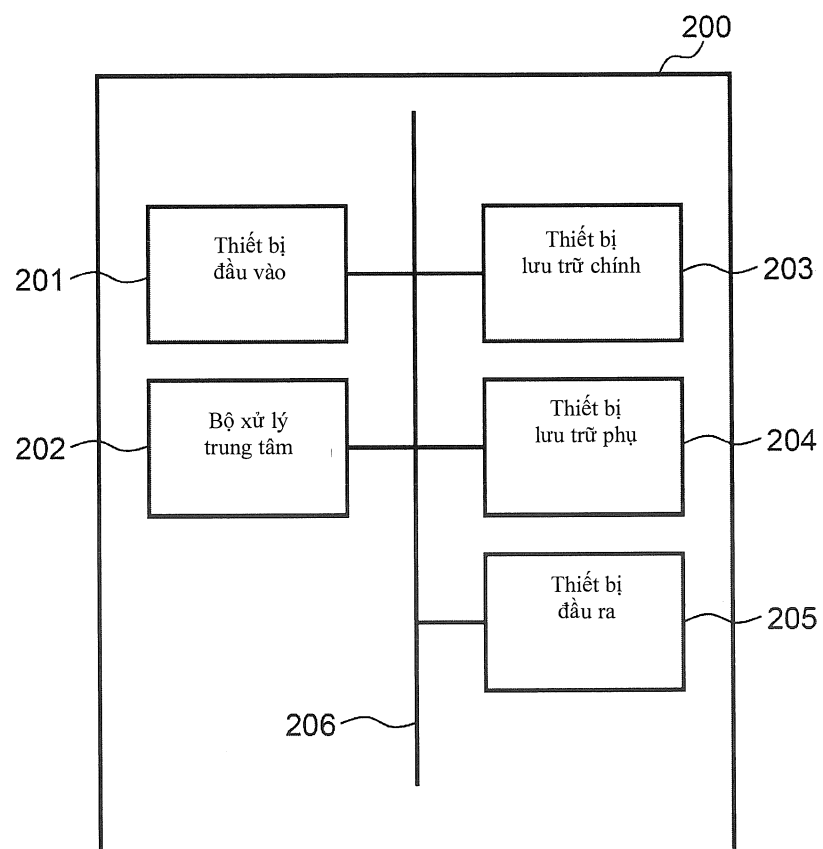


FIG. 3

Đầu vào x	Đầu ra a
x1	(0,91 ; 0,11 ; 0,02 ; 0,87)
x2	(0,10 ; 0,94 ; 0,08 ; 0,91)
...	...

FIG. 4A

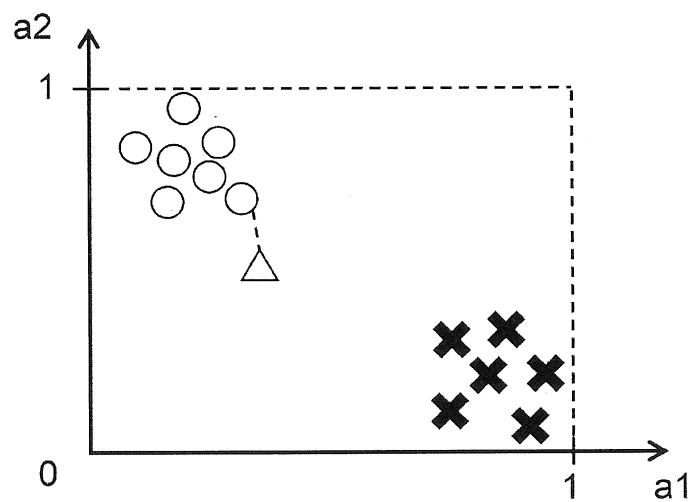


FIG. 4B

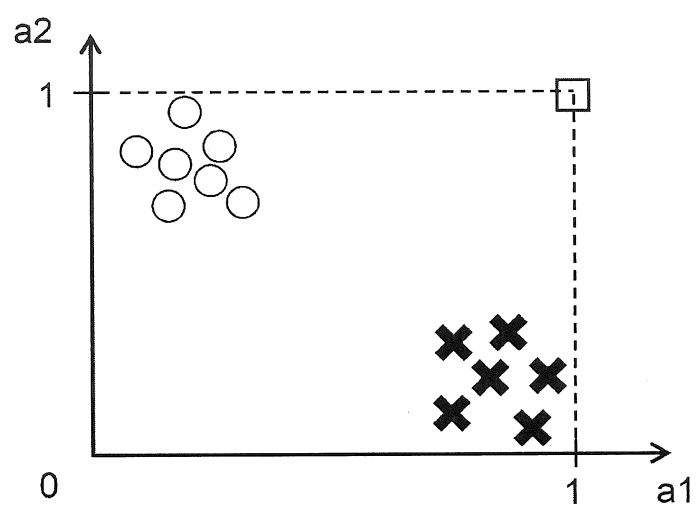
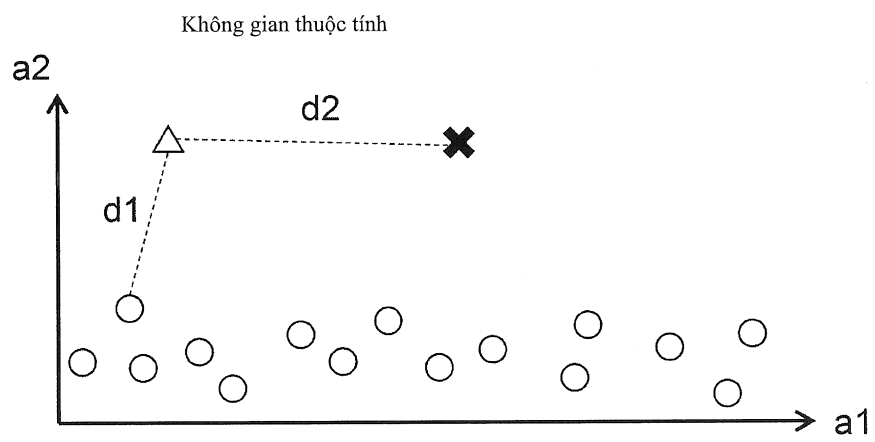


FIG. 5



○ : Mẫu nhóm hiện có

✕ : Mẫu nhóm mới hoặc nhóm giả

△ : Dữ liệu dùng cho quy trình nhận dạng

FIG. 6

Thuộc tính	Đánh máy	Uống cà phê
Màn hình	1	?
Bàn phím	1	?
Cốc cà phê	?	1
Cường độ chuyển động đầu ngón tay	1	0
Chuyển động thẳng đứng của cánh tay	0	1
Ngồi	1	?

FIG. 7

Thuộc tính	Đánh máy	Uống cà phê
Màn hình	W_{11}	W_{21}
Bàn phím	W_{12}	W_{22}
Cốc cà phê	W_{13}	W_{23}
Cường độ chuyển động đầu ngón tay	W_{14}	W_{24}
Chuyển động thẳng đứng của cánh tay	W_{15}	W_{25}
Ngồi	W_{16}	W_{26}

FIG. 8

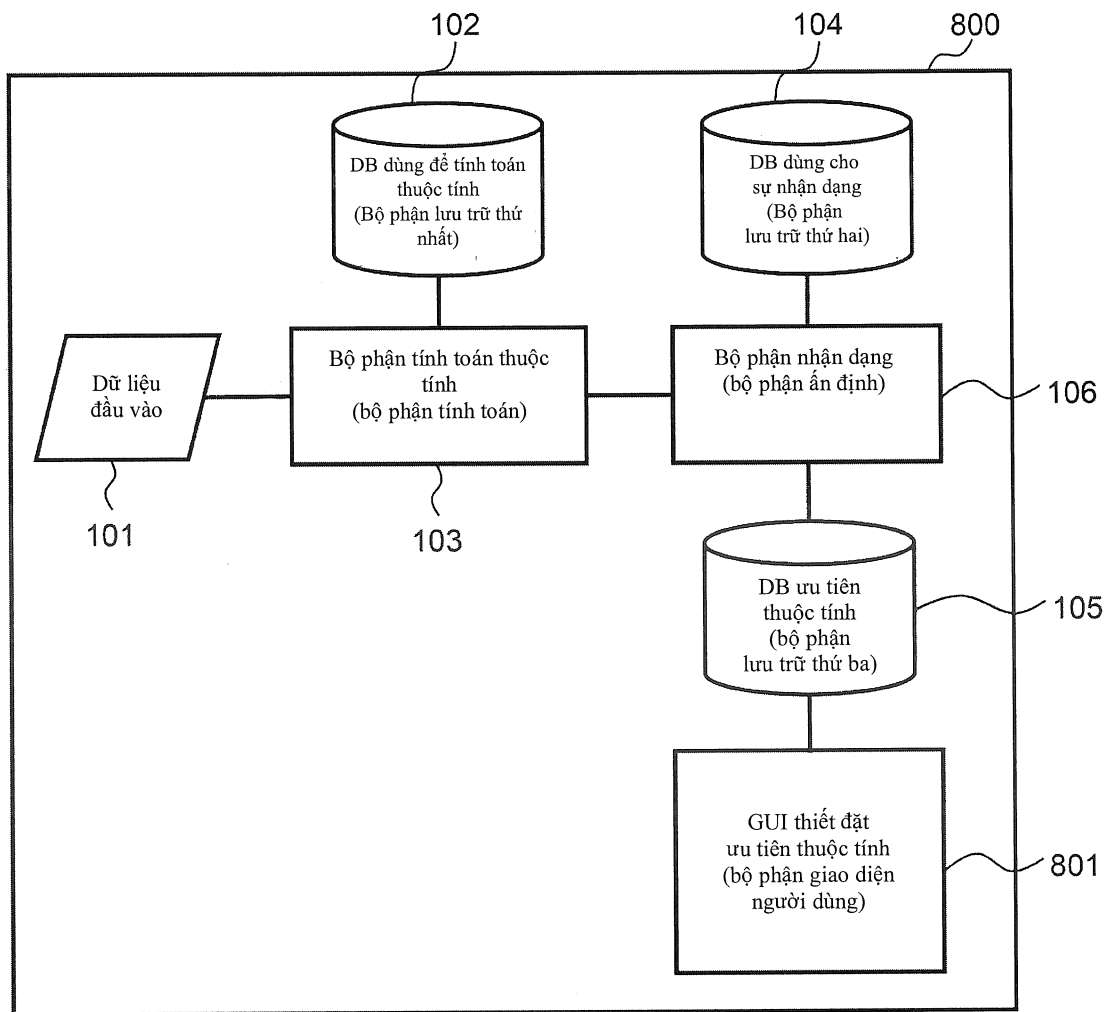


FIG. 9

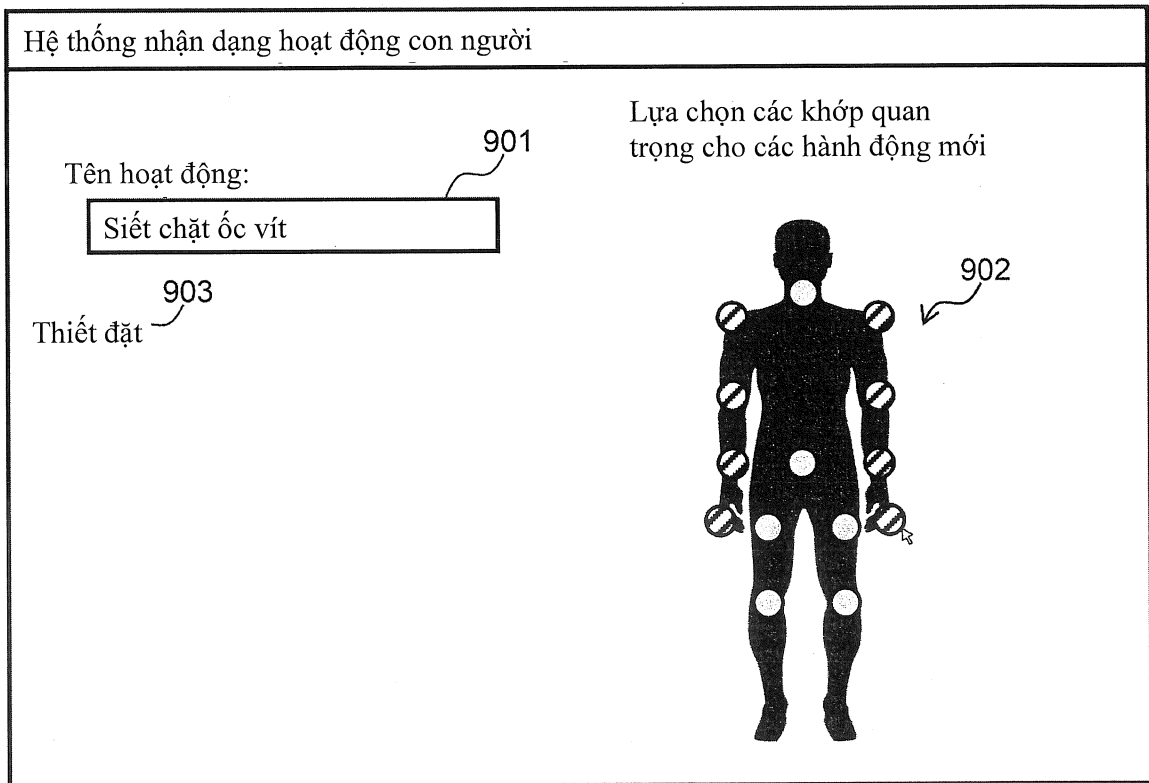


FIG. 10

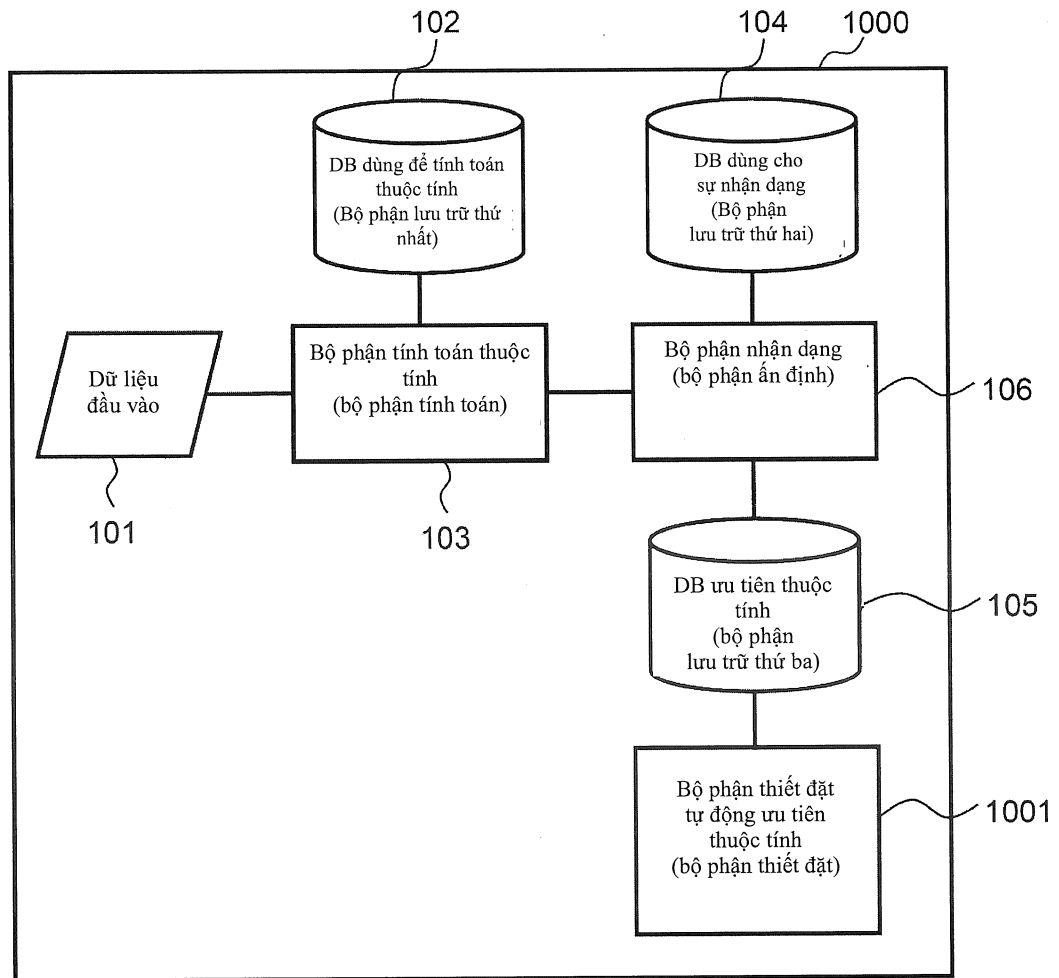


FIG. 11

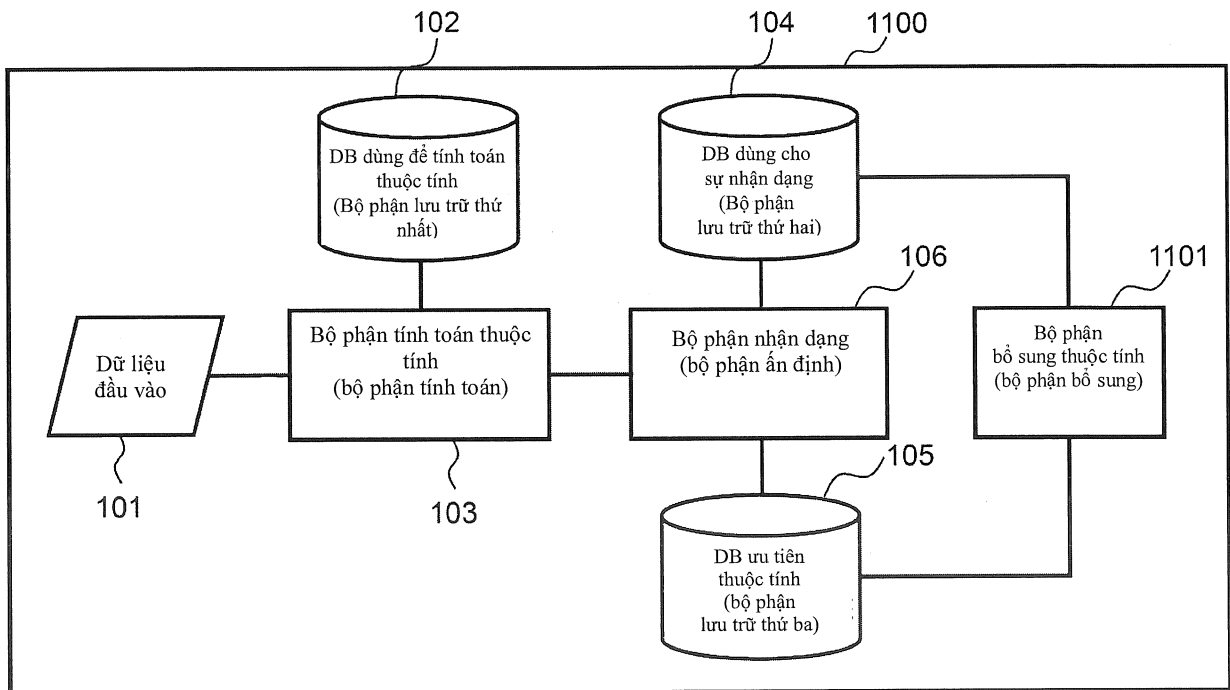


FIG. 12

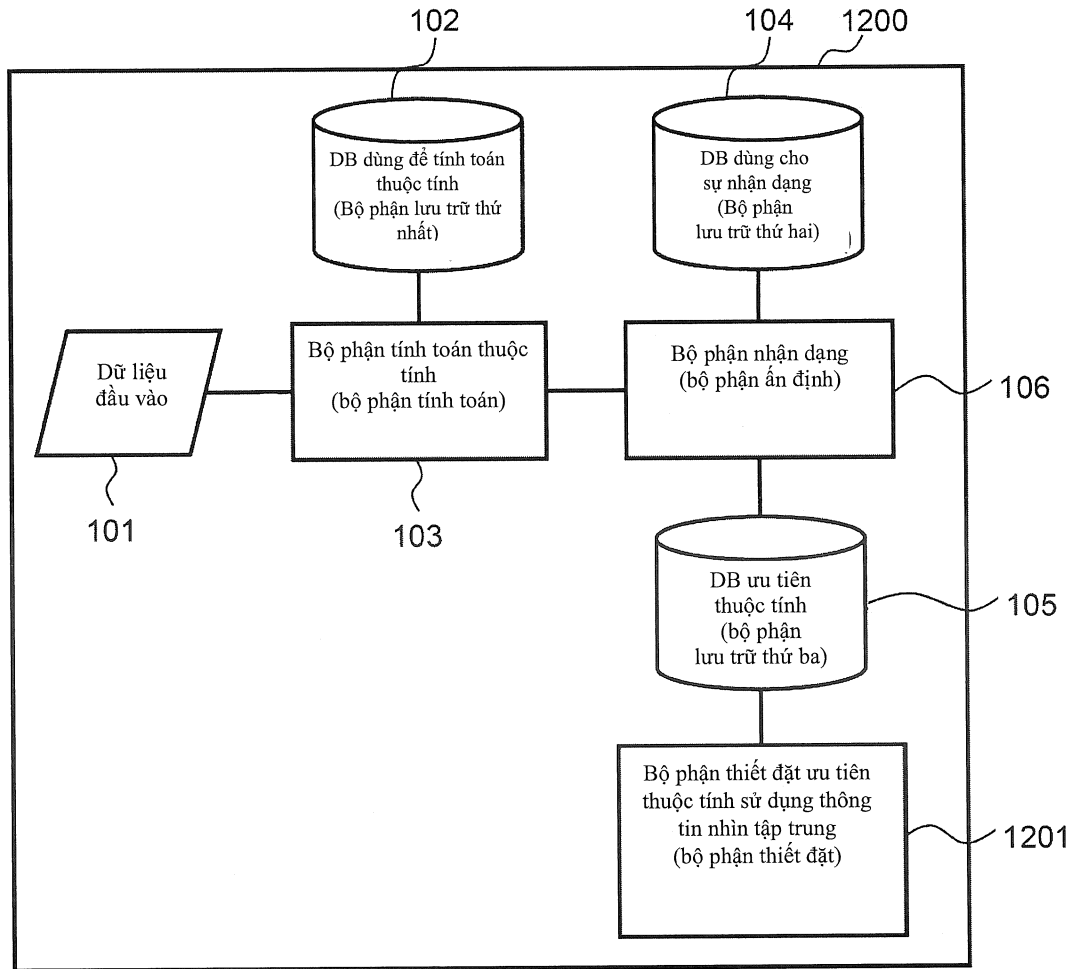


FIG. 13

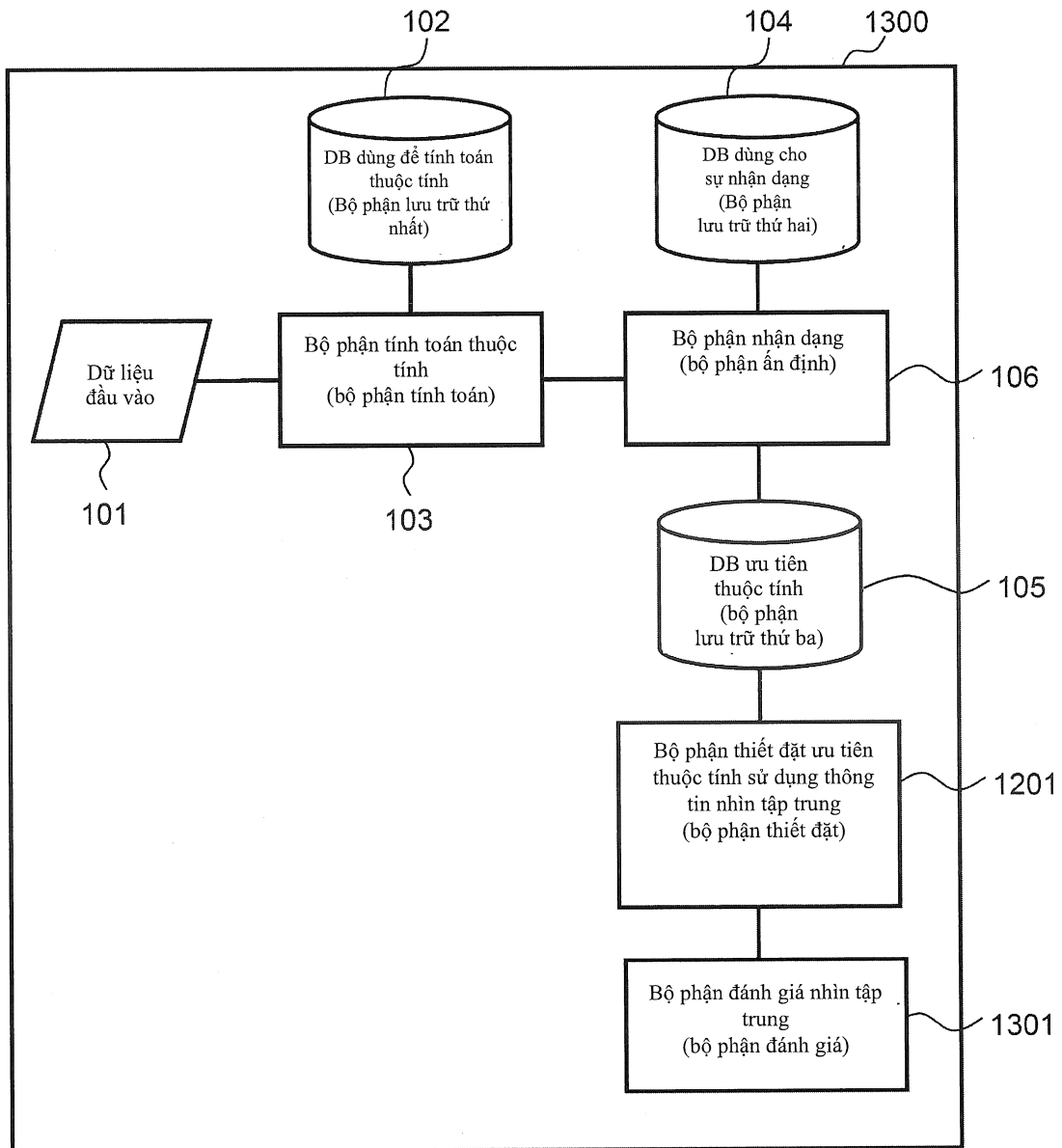


FIG. 14

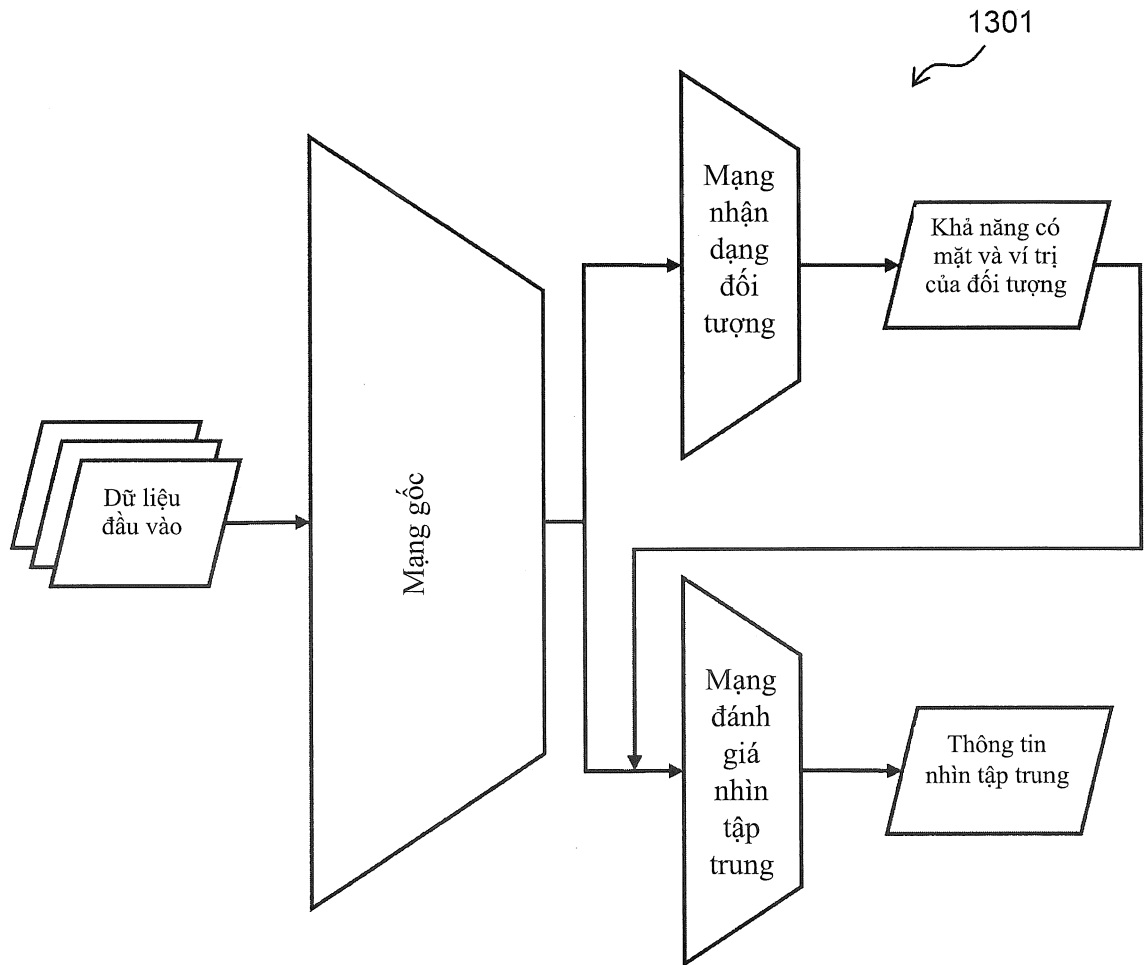


FIG. 15

