



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

(19) Cộng hòa xã hội chủ nghĩa Việt Nam (VN) (11)
CỤC SỞ HỮU TRÍ TUỆ



1-0048696

(51)^{2020.01} G06N 3/00; G06F 16/00; G06N 20/00 (13) B

(21) 1-2021-04164

(22) 07/07/2021

(45) 25/07/2025 448

(43) 27/01/2023 418A

(73) TẬP ĐOÀN VINGROUP - CÔNG TY CP (VN)

Số 7, đường Bằng Lăng 1, khu đô thị sinh thái Vinhomes Riverside, phường Việt Hưng, quận Long Biên, thành phố Hà Nội, Việt Nam

(72) Khuất Quang Ngọc (VN); Nguyễn Minh Hiếu (VN); Lee Quy William (US).

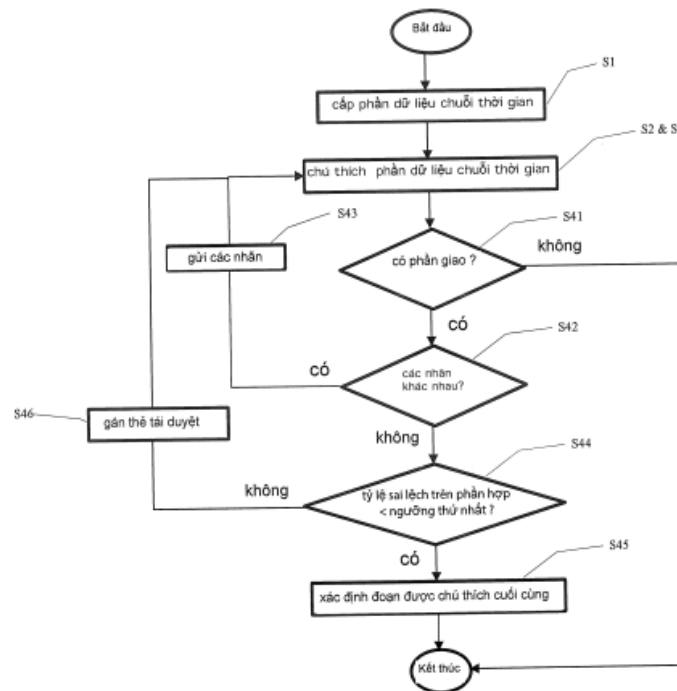
(74) Công ty TNHH Tầm nhìn và Liên danh (VISION & ASSOCIATES CO.LTD.)

(54) PHƯƠNG PHÁP VÀ HỆ THỐNG DÁN NHÃN ĐỐI VỚI DỮ LIỆU CHUỖI THỜI GIAN

(21) 1-2021-04164

(57) Sáng chế đề cập đến phương pháp và hệ thống dán nhãn. Phương pháp dán nhãn bao gồm bước xác định xem liệu có phần giao giữa đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai hay không; nếu có, xác định xem liệu đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai có các nhãn khác nhau hay không; nếu có, gửi nhãn của đoạn được chú thích thứ nhất đến đơn vị tác tử thứ hai và nhãn của đoạn được chú thích thứ hai đến đơn vị tác tử thứ nhất; nếu không, xác định xem liệu tỷ lệ sai lệch trên phần hợp giữa đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai dưới ngưỡng thứ nhất hay không; nếu có, xác định phân giao giữa đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai làm đoạn được chú thích cuối cùng cho phần dữ liệu chuỗi thời gian; nếu không, gán thẻ tái duyệt cho phần hợp của phần dữ liệu chuỗi thời gian mà có khoảng thời gian từ bắt đầu đoạn được chú thích thứ nhất đến kết thúc đoạn được chú thích thứ hai.

FIG. 1B



Lĩnh vực kỹ thuật được đề cập

Các phương án đề cập đến phương pháp và hệ thống dán nhãn đối với dữ liệu chuỗi thời gian.

Tình trạng kỹ thuật của sáng chế

Trong các hệ thống dán nhãn trước đây, các hình ảnh được chú thích thủ công bằng các người chú thích. Có hai nhược điểm đối với giải pháp này.

Thứ nhất là cần mức lao động đáng kể của con người để chú thích thủ công. Ví dụ, lượng dữ liệu là lớn và thường được thu ở 50Hz, do đó thời gian và nỗ lực được sử dụng để chú thích và xác thực thủ công là lớn. Lượng dữ liệu cần thiết để huấn luyện kiểu mô hình AI này là lớn do tính đa dạng trong cách mà con người thực hiện các hoạt động. Ví dụ: Có nhiều cách khác nhau để một người thực hiện hoạt động đi bộ, và thậm chí đa dạng hơn nhiều để một người khác thực hiện cùng hoạt động đó.

Thứ hai là tính không chính xác của chú thích do tính chủ quan trong nhận thức của con người. Định nghĩa về các hoạt động là mơ hồ, do đó có tính không nhất quán cao trong các kết quả chú thích giữa các trình chú thích khác nhau trên cùng phần dữ liệu chuỗi thời gian.

Do đó, cần có phương pháp làm tăng tốc độ chú thích mà vẫn giải quyết vấn đề về tính đa dạng lớn trong dữ liệu.

Bản chất kỹ thuật của sáng chế

Sáng chế đề xuất phương pháp và hệ thống dán nhãn đối với dữ liệu chuỗi thời gian.

Các vấn đề được giải quyết theo các phương án không bị giới hạn ở đó và bao gồm các giải pháp kỹ thuật sau và cũng như các mục đích hoặc hiệu quả có thể hiểu được từ các phương án này.

Theo một khía cạnh, sáng chế đề xuất phương pháp dán nhãn bao gồm các bước S1: cấp phần dữ liệu chuỗi thời gian cho đơn vị tác tử thứ nhất và đơn vị tác tử thứ hai;

S2: chú thích, bởi đơn vị tác tử thứ nhất, phần dữ liệu chuỗi thời gian để tạo ra phần dữ liệu chuỗi thời gian được chú thích thứ nhất có ít nhất một đoạn được chú thích thứ nhất; S3: chú thích, bởi đơn vị tác tử thứ hai, phần dữ liệu chuỗi thời gian để tạo ra phần dữ liệu chuỗi thời gian được chú thích thứ hai có ít nhất một đoạn được chú thích thứ hai; S4: xử lý, bởi thiết bị dán nhãn, phần dữ liệu chuỗi thời gian được chú thích thứ nhất và phần dữ liệu chuỗi thời gian được chú thích thứ hai, trong đó bước xử lý bao gồm:

S41: xác định xem liệu có phần giao giữa đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai hay không, nếu có, đi đến S42, nếu không, kết thúc việc xử lý;

S42: xác định xem liệu đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai có các nhãn khác nhau hay không, nếu có, đi đến S43, nếu không, đi đến S44;

S43: gửi nhãn của đoạn được chú thích thứ nhất đến đơn vị tác tử thứ hai và nhãn của đoạn được chú thích thứ hai đến đơn vị tác tử thứ nhất;

S44: xác định xem liệu tỷ lệ sai lệch trên phần hợp giữa đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai dưới ngưỡng thứ nhất hay không, nếu có, đi đến S45, nếu không, đi đến S46;

S45: xác định phần giao giữa đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai làm đoạn được chú thích cuối cùng đối với phần dữ liệu chuỗi thời gian; và

S46: gán thẻ tái duyệt cho phần hợp của phần dữ liệu chuỗi thời gian mà có khoảng thời gian từ bắt đầu đoạn được chú thích thứ nhất đến kết thúc đoạn được chú thích thứ hai.

Theo một phương án của sáng chế, phương pháp còn bao gồm bước S48 là bước khuyến nghị, sau khi xác định xem liệu có phần giao giữa mỗi đoạn được chú thích thứ nhất và mỗi đoạn được chú thích thứ hai, ít nhất một đoạn riêng lẻ của một đơn vị trong số các đơn vị tác tử thứ nhất và đơn vị tác tử thứ hai, ít nhất một đoạn riêng lẻ này không giao với bất kỳ đoạn nào của đơn vị còn lại trong số các đơn vị tác

tử thứ nhất và đơn vị tác tử thứ hai, cho ít nhất một hoặc nhiều đơn vị trong số các đơn vị tác tử thứ nhất và đơn vị tác tử thứ hai làm các thẻ khuyến nghị.

Theo một phương án của sáng chế, các bước S2 và S3 để chú thích phần dữ liệu chuỗi thời gian được thực hiện bởi các mô hình trí tuệ nhân tạo (artificial intelligence, AI).

Theo một phương án của sáng chế, các bước S2 và S3 để chú thích phần dữ liệu chuỗi thời gian được thực hiện bởi người dùng.

Theo một phương án của sáng chế, đơn vị tác tử thứ nhất và đơn vị tác tử thứ hai bao gồm các mô hình trí tuệ nhân tạo (AI) để giúp người dùng tái duyệt khác biệt giữa các kết quả làm việc của người dùng trên phần dữ liệu chuỗi thời gian.

Theo một phương án của sáng chế, bước S44 bao gồm bước xác định xem liệu tỷ lệ phân giao trên phần hợp giữa đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai vượt quá ngưỡng thứ hai hay không, nếu có, đi đến S45, nếu không, đi đến S46.

Theo một phương án của sáng chế, bước S43 còn bao gồm bước quay về S2 để chú thích phần dữ liệu chuỗi thời gian dựa vào nhãn của đoạn được chú thích thứ hai được gửi đến đơn vị tác tử thứ nhất; và quay về S3 để chú thích phần dữ liệu chuỗi thời gian dựa vào nhãn của đoạn được chú thích thứ nhất được gửi đến đơn vị tác tử thứ hai.

Theo một phương án của sáng chế, bước S46 còn bao gồm bước quay về S2 để chú thích phần dữ liệu chuỗi thời gian dựa vào thẻ tái duyệt được gán; và quay về S3 để chú thích phần dữ liệu chuỗi thời gian dựa vào thẻ tái duyệt được gán.

Theo một phương án của sáng chế, tỷ lệ sai lệch trên phần hợp giữa đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai được tính bằng cách sử dụng phương trình sau:

$$XoLI = \frac{(|start_1 - start_2| + |end_1 - end_2|)}{\max(\max(end_1, end_2) - \min(start_1, start_2), 10^{-6})}$$

trong đó, $start_1$ là bắt đầu của đoạn được chú thích thứ nhất, $start_2$ là bắt đầu của đoạn được chú thích thứ hai, end_1 là kết thúc của đoạn được chú thích thứ nhất, và end_2 là kết thúc của đoạn được chú thích thứ hai.

Theo một khía cạnh khác, sáng chế đề xuất hệ thống dán nhãn bao gồm đơn vị tác tử thứ nhất được tạo cấu hình để chú thích phần dữ liệu chuỗi thời gian để tạo ra phần dữ liệu chuỗi thời gian được chú thích thứ nhất có ít nhất một đoạn được chú thích thứ nhất; đơn vị tác tử thứ hai được tạo cấu hình để chú thích phần dữ liệu chuỗi thời gian để tạo ra phần dữ liệu chuỗi thời gian được chú thích thứ hai có ít nhất một đoạn được chú thích thứ hai; thiết bị dán nhãn được tạo cấu hình để xử lý phần dữ liệu chuỗi thời gian được chú thích thứ nhất và phần dữ liệu chuỗi thời gian được chú thích thứ hai, trong đó bước xử lý bao gồm:

S41: xác định xem liệu có phần giao giữa đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai hay không, nếu có, đi đến S42, nếu không, kết thúc việc xử lý;

S42: xác định xem liệu đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai có các nhãn khác nhau hay không, nếu có, đi đến S43, nếu không, đi đến S44;

S43: gửi nhãn của đoạn được chú thích thứ nhất đến đơn vị tác tử thứ hai và nhãn của đoạn được chú thích thứ hai đến đơn vị tác tử thứ nhất;

S44: xác định xem liệu tỷ lệ sai lệch trên phần hợp giữa đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai dưới ngưỡng thứ nhất hay không, nếu có, đi đến S45, nếu không, đi đến S46;

S45: xác định phần giao giữa đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai làm đoạn được chú thích cuối cùng đối với phần dữ liệu chuỗi thời gian; và

S46: gán thẻ tái duyệt cho phần hợp của phần dữ liệu chuỗi thời gian mà có khoảng thời gian từ bắt đầu đoạn được chú thích thứ nhất đến kết thúc đoạn được chú thích thứ hai.

Theo một phương án của sáng chế, hệ thống còn được tạo cấu hình để khuyến nghị, sau khi xác định xem liệu có phần giao giữa mỗi đoạn được chú thích thứ nhất và mỗi đoạn được chú thích thứ hai, ít nhất một đoạn riêng lẻ của một đơn vị trong số các đơn vị tác tử thứ nhất và đơn vị tác tử thứ hai, ít nhất một đoạn riêng lẻ này không giao với bất kỳ đoạn nào của đơn vị còn lại trong số các đơn vị tác tử thứ nhất và đơn vị tác

tử thứ hai, cho ít nhất một hoặc nhiều đơn vị trong số các đơn vị tác tử thứ nhất và đơn vị tác tử thứ hai làm các thẻ khuyến nghị.

Theo một phương án của sáng chế, đơn vị tác tử thứ nhất và đơn vị tác tử thứ hai sử dụng các mô hình trí tuệ nhân tạo (AI) để chú thích phần dữ liệu chuỗi thời gian.

Theo một phương án của sáng chế, đơn vị tác tử thứ nhất và đơn vị tác tử thứ hai được sử dụng bởi người dùng để chú thích phần dữ liệu chuỗi thời gian.

Theo một phương án của sáng chế, đơn vị tác tử thứ nhất được tạo cấu hình để chú thích phần dữ liệu chuỗi thời gian dựa vào nhãn của đoạn được chú thích thứ hai được gửi đến đơn vị tác tử thứ nhất; và đơn vị tác tử thứ hai được tạo cấu hình để chú thích phần dữ liệu chuỗi thời gian dựa vào nhãn của đoạn được chú thích thứ nhất được gửi đến đơn vị tác tử thứ hai.

Theo một phương án của sáng chế, đơn vị tác tử thứ nhất được tạo cấu hình để chú thích phần dữ liệu chuỗi thời gian dựa vào thẻ tái duyệt được gán; và đơn vị tác tử thứ hai được tạo cấu hình để chú thích phần dữ liệu chuỗi thời gian dựa vào thẻ tái duyệt được gán.

Theo một khía cạnh khác nữa, sáng chế đề xuất vật ghi lâu dài đọc được bằng máy tính bao gồm các lệnh mà, khi được thực thi bởi ít nhất một bộ xử lý của máy, khiến cho máy này thực hiện các hoạt động bao gồm S1: cấp phần dữ liệu chuỗi thời gian cho đơn vị tác tử thứ nhất và đơn vị tác tử thứ hai; S2: chú thích, bởi đơn vị tác tử thứ nhất, phần dữ liệu chuỗi thời gian để tạo ra phần dữ liệu chuỗi thời gian được chú thích thứ nhất có ít nhất một đoạn được chú thích thứ nhất; S3: chú thích, bởi đơn vị tác tử thứ hai, phần dữ liệu chuỗi thời gian để tạo ra phần dữ liệu chuỗi thời gian được chú thích thứ hai có ít nhất một đoạn được chú thích thứ hai; S4: xử lý, bởi thiết bị dán nhãn, phần dữ liệu chuỗi thời gian được chú thích thứ nhất và phần dữ liệu chuỗi thời gian được chú thích thứ hai, trong đó bước xử lý bao gồm:

S41: xác định xem liệu có phần giao giữa đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai hay không, nếu có, đi đến S42, nếu không, kết thúc việc xử lý;

S42: xác định xem liệu đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai có các nhãn khác nhau hay không, nếu có, đi đến S43, nếu không, đi đến S44;

S43: gửi nhãn của đoạn được chú thích thứ nhất đến đơn vị tác tử thứ hai và nhãn của đoạn được chú thích thứ hai đến đơn vị tác tử thứ nhất;

S44: xác định xem liệu tỷ lệ sai lệch trên phần hợp giữa đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai dưới ngưỡng thứ nhất hay không, nếu có, đi đến S45, nếu không, đi đến S46;

S45: xác định phần giao giữa đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai làm đoạn được chú thích cuối cùng đối với phần dữ liệu chuỗi thời gian; và

S46: gán thẻ tái duyệt cho phần hợp của phần dữ liệu chuỗi thời gian mà có khoảng thời gian từ bắt đầu đoạn được chú thích thứ nhất đến kết thúc đoạn được chú thích thứ hai.

Mô tả vắn tắt các hình vẽ

Các mục đích, dấu hiệu và ưu điểm nêu trên và các mục đích, dấu hiệu và ưu điểm khác của sáng chế sẽ trở nên rõ ràng hơn đối với người có hiểu biết trung bình về lĩnh vực kỹ thuật này bằng cách mô tả chi tiết các phương án làm ví dụ của sáng chế với tham chiếu đến các hình vẽ kèm theo, trong đó:

Fig.1A là sơ đồ khối của hệ thống dán nhãn theo một phương án của sáng chế.

Fig.1B là lưu đồ minh họa quy trình của phương pháp dán nhãn theo một phương án của sáng chế.

Fig.1C là lưu đồ minh họa quy trình của phương pháp dán nhãn theo một phương án khác của sáng chế.

Fig.2 là lưu đồ minh họa quy trình của phương pháp dán nhãn theo một phương án của sáng chế.

Fig.3 là lưu đồ thể hiện quy trình ở giai đoạn thứ nhất theo một phương án của sáng chế.

Fig.4 là lưu đồ thể hiện quy trình ở giai đoạn thứ hai theo một phương án của sáng chế.

Các hình vẽ từ Fig.5A đến Fig.5D là các sơ đồ để giải thích giai đoạn thứ hai.

Fig.6 là lưu đồ thể hiện quy trình ở giai đoạn thứ ba theo một phương án của sáng chế.

Fig.7 là lưu đồ thể hiện quy trình ở giai đoạn thứ tư theo một phương án của sáng chế.

Mô tả chi tiết sáng chế

Trong khi sáng chế có thể có nhiều biến đổi và dạng thay thế, các phương án cụ thể của sáng chế được thể hiện bằng các ví dụ trên hình vẽ và sẽ được mô tả chi tiết trong bản mô tả này. Tuy nhiên, điều này không nhằm làm giới hạn sáng chế ở các dạng cụ thể được bộc lộ. Ngược lại, sáng chế bao gồm tất cả các biến đổi, sửa đổi và tương tự nằm trong mục đích và phạm vi của yêu cầu bảo hộ kèm theo.

Nên hiểu rằng, mặc dù các thuật ngữ “thứ nhất”, “thứ hai”, và tương tự có thể được sử dụng trong bản mô tả này để mô tả các phần tử khác nhau, nhưng các phần tử này không bị giới hạn bởi các thuật ngữ này. Các thuật ngữ này chỉ được sử dụng để phân biệt phần tử này với phần tử khác. Ví dụ, phần tử thứ nhất có thể được gọi là phần tử thứ hai, và tương tự, phần tử thứ hai có thể được gọi là phần tử thứ nhất mà không tách rời khỏi phạm vi của sáng chế. Khi được sử dụng ở đây, thuật ngữ “và/hoặc” bao gồm bất kỳ và tất cả các kết hợp của một hoặc nhiều trong số các mục được liệt kê liên quan.

Nên hiểu rằng khi phần tử được mô tả là “được kết nối” hoặc “được ghép nối” với phần tử khác, thì phần tử này có thể được kết nối hoặc được ghép nối trực tiếp với phần tử khác hoặc các phần tử xen giữa có thể có mặt. Ngược lại, khi phần tử được mô tả là “được kết nối trực tiếp” hoặc “được ghép nối trực tiếp” với phần tử khác, thì không có các phần tử xen giữa.

Thuật ngữ được sử dụng ở đây chỉ nhằm mục đích mô tả các phương án cụ thể và không nhằm giới hạn sáng chế. Như được sử dụng ở đây, các dạng số ít nhằm bao gồm các dạng số nhiều, trừ khi nội dung biểu thị rõ ràng khác. Cũng nên hiểu rằng các thuật ngữ “bao gồm”, “gồm”, “gồm có”, và/hoặc “chứa”, khi được sử dụng trong bản

mô tả này, chỉ rõ sự có mặt các dấu hiệu, tổng thể, bước, hoạt động, phần tử, bộ phận đã nêu hoặc sự kết hợp của chúng, nhưng không loại trừ sự có mặt hoặc bỏ sung của một hoặc nhiều các dấu hiệu, tổng thể, bước, hoạt động, phần tử, bộ phận khác hoặc sự kết hợp của chúng.

Trừ khi được định nghĩa khác, tất cả các thuật ngữ bao gồm thuật ngữ kỹ thuật và khoa học được sử dụng ở đây có nghĩa giống như thường được hiểu bởi người có hiểu biết trung bình trong lĩnh vực kỹ thuật mà sáng chế này thuộc về. Cũng nên hiểu thêm rằng các thuật ngữ, như được định nghĩa trong từ điển hay được sử dụng, cần được giải thích ý nghĩa thống nhất với nghĩa trong ngữ cảnh của lĩnh vực kỹ thuật và không được giải thích theo nghĩa lý tưởng hoặc quá trang trọng trừ khi được chỉ ra rõ ràng ở đây.

Sau đây, các phương án sẽ được mô tả chi tiết với tham chiếu đến các hình vẽ kèm theo, các thành phần giống nhau hoặc các thành phần tương ứng sẽ được ký hiệu bằng các số chỉ dẫn giống nhau bất kể các số chỉ dẫn, và do đó phần mô tả của chúng sẽ không được lặp lại.

Sáng chế đề cập để quy trình để chú thích phần dữ liệu chuỗi thời gian với mục đích là cả tự động hóa các bước xác minh dữ liệu và tăng chất lượng dữ liệu.

Tham chiếu đến Fig.1A, hệ thống dán nhãn 100 theo một phương án của sáng chế có thể bao gồm đơn vị tác tử thứ nhất 110, đơn vị tác tử thứ hai 120, và thiết bị dán nhãn 130. Đơn vị tác tử thứ nhất 110 được tạo cấu hình để chú thích phần dữ liệu chuỗi thời gian để tạo ra phần dữ liệu chuỗi thời gian được chú thích thứ nhất có ít nhất một đoạn được chú thích thứ nhất.

Đơn vị tác tử thứ hai 120 được tạo cấu hình để chú thích phần dữ liệu chuỗi thời gian để tạo ra phần dữ liệu chuỗi thời gian được chú thích thứ hai có ít nhất một đoạn được chú thích thứ hai.

Thiết bị dán nhãn 130 được tạo cấu hình để xử lý phần dữ liệu chuỗi thời gian được chú thích thứ nhất và phần dữ liệu chuỗi thời gian được chú thích thứ hai. Quy trình xử lý phần dữ liệu chuỗi thời gian được chú thích thứ nhất và phần dữ liệu chuỗi thời gian được chú thích thứ hai trong thiết bị dán nhãn 130 có thể được mô tả sau.

Theo một phương án của sáng chế, đơn vị tác tử thứ nhất 110 và đơn vị tác tử

thứ hai 120 được sử dụng bởi người dùng để chú thích phần dữ liệu chuỗi thời gian. Ít nhất một trong số đơn vị tác tử thứ nhất 110 và đơn vị tác tử thứ hai 120 có thể là thiết bị mà thông tin được nhập vào bởi người dùng. Ví dụ, đơn vị tác tử thứ nhất 110 có thể là thiết bị tính toán mà nhận phần dữ liệu chuỗi thời gian để tạo ra phần dữ liệu chuỗi thời gian được chú thích thứ nhất có ít nhất một đoạn được chú thích thứ nhất. Thao tác của người dùng có thể là sử dụng đơn vị tác tử thứ nhất 110 để chú thích thủ công phần dữ liệu chuỗi thời gian với danh sách nhãn cho trước. Và đối với mỗi nhãn được chú thích, người dùng có thể phải quyết định khi nào nhãn này bắt đầu và kết thúc trên phần dữ liệu chuỗi thời gian. Người dùng có thể được cung cấp công cụ dán nhãn nội bộ để phù hợp với phần dữ liệu chuỗi thời gian. Người dùng có thể dán nhãn vào các mẫu video bằng cách thực hiện các thao tác kéo và thả. Sau quy trình dán nhãn, dữ liệu video được đồng bộ với phần dữ liệu chuỗi thời gian tương ứng để nhận kết quả dán nhãn cuối cùng.

Theo một phương án của sáng chế, đơn vị tác tử thứ nhất 110 và đơn vị tác tử thứ hai 120 sử dụng các mô hình trí tuệ nhân tạo (AI) để chú thích phần dữ liệu chuỗi thời gian.

Theo một phương án của sáng chế, một trong số đơn vị tác tử thứ nhất 110 và đơn vị tác tử thứ hai 120 có thể được sử dụng bởi người dùng để chú thích phần dữ liệu chuỗi thời gian và đơn vị còn lại trong số đơn vị tác tử thứ nhất 110, đơn vị tác tử thứ hai 120 có thể bao gồm mô hình AI để chú thích phần dữ liệu chuỗi thời gian.

Theo một phương án của sáng chế, đơn vị tác tử thứ nhất 110 và đơn vị tác tử thứ hai 120 có thể bao gồm các mô hình trí tuệ nhân tạo (AI) để giúp người dùng tái duyệt khác biệt giữa các kết quả làm việc của người dùng trên phần dữ liệu chuỗi thời gian.

Theo một phương án của sáng chế, đơn vị tác tử thứ nhất 110 được tạo cấu hình để chú thích phần dữ liệu chuỗi thời gian dựa vào nhãn của đoạn được chú thích thứ hai được gửi đến đơn vị tác tử thứ nhất 110; và đơn vị tác tử thứ hai 120 được tạo cấu hình để chú thích phần dữ liệu chuỗi thời gian dựa vào nhãn của đoạn được chú thích thứ nhất được gửi đến đơn vị tác tử thứ hai 120. Ở S42, sau khi thiết bị dán nhãn 130 xác định rằng đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai có các nhãn khác nhau, nhãn của đoạn được chú thích thứ nhất được gửi trở về đơn vị tác tử

thứ hai và nhãn của đoạn được chú thích thứ hai được gửi trở về đơn vị tác tử thứ nhất để khuyến nghị. Dựa vào các nhãn được khuyến nghị, đơn vị tác tử thứ nhất 110 và đơn vị tác tử thứ hai 120 có thể xem xét lại các lựa chọn nhãn khi có sự khác nhau về nhãn của đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai.

Theo một phương án của sáng chế, đơn vị tác tử thứ nhất 110 được tạo cấu hình để chú thích phần dữ liệu chuỗi thời gian dựa vào thẻ tái duyệt được gán; và đơn vị tác tử thứ hai 120 được tạo cấu hình để chú thích phần dữ liệu chuỗi thời gian dựa vào thẻ tái duyệt được gán. Ở bước S46, sau khi thiết bị dán nhãn 130 gán thẻ tái duyệt cho đoạn của phần dữ liệu chuỗi thời gian mà có khoảng thời gian từ bắt đầu đoạn được chú thích thứ nhất đến kết thúc đoạn được chú thích thứ hai, đoạn này được đánh dấu là “cần được tái duyệt” bởi thẻ tái duyệt được gửi trở về đơn vị tác tử thứ nhất 110 và đơn vị tác tử thứ hai 120 để khuyến nghị. Dựa vào thẻ tái duyệt được gán, đơn vị tác tử thứ nhất 110 và đơn vị tác tử thứ hai 120 có thể kiểm tra và điều chỉnh khoảng thời gian của đoạn được đánh dấu khi có sự khác nhau về khoảng thời gian của đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai.

Trên Fig.1A, chỉ đơn vị tác tử thứ nhất 110 và đơn vị tác tử thứ hai 120 được thể hiện để thuận tiện cho việc mô tả, nhưng hệ thống dán nhãn theo một phương án của sáng chế có thể bao gồm hai hoặc nhiều đơn vị tác tử.

Theo một phương án của sáng chế, hệ thống dán nhãn 100 có thể là hệ thống phân tán trong đó đơn vị tác tử thứ nhất 110, đơn vị tác tử thứ hai 120, và thiết bị dán nhãn 130 có thể được truyền thông với nhau thông qua mạng 140. Mạng 140 có thể là mạng nối dây hoặc mạng không dây. Mạng 140 có thể là internet, mạng nội bộ, mạng mở rộng, LAN hoặc tương tự.

Tham chiếu đến Fig.1B, phương pháp dán nhãn theo một phương án của sáng chế được mô tả. Phương pháp này có thể bao gồm các bước:

S1: cấp phần dữ liệu chuỗi thời gian cho đơn vị tác tử thứ nhất 110 và đơn vị tác tử thứ hai 120;

S2: chú thích, bởi đơn vị tác tử thứ nhất 110, phần dữ liệu chuỗi thời gian để tạo ra phần dữ liệu chuỗi thời gian được chú thích thứ nhất có ít nhất một đoạn được chú thích thứ nhất;

S3: chú thích, bởi đơn vị tác tử thứ hai 120, phần dữ liệu chuỗi thời gian để tạo ra phần dữ liệu chuỗi thời gian được chú thích thứ hai có ít nhất một đoạn được chú thích thứ hai;

S4: xử lý, bởi thiết bị dán nhãn 130, phần dữ liệu chuỗi thời gian được chú thích thứ nhất và phần dữ liệu chuỗi thời gian được chú thích thứ hai, trong đó bước xử lý bao gồm các bước:

S41: xác định xem liệu có phần giao giữa đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai hay không, nếu có, đi đến S42, nếu không, kết thúc việc xử lý;

S42: xác định xem liệu đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai có các nhãn khác nhau hay không, nếu có, đi đến S43, nếu không, đi đến S44;

S43: gửi nhãn của đoạn được chú thích thứ nhất đến đơn vị tác tử thứ hai 120 và nhãn của đoạn được chú thích thứ hai đến đơn vị tác tử thứ nhất 110;

S44: xác định xem liệu tỷ lệ sai lệch trên phần hợp giữa đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai dưới ngưỡng thứ nhất hay không, nếu có, đi đến S45, nếu không, đi đến S46. Tốt hơn là, ngưỡng thứ nhất được xác định thông qua thực nghiệm và nằm trong khoảng 0,15-0,5. Cụ thể hơn, ngưỡng thứ nhất có thể là 0,2; 0,25; 0,3; 0,35; 0,4 và 0,45. Theo một phương án của sáng chế, tỷ lệ phần giao trên phần hợp có thể được sử dụng thay cho tỷ lệ sai lệch trên phần hợp giữa đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai. Trong trường hợp này, S44 có thể bao gồm bước xác định xem liệu tỷ lệ phần giao trên phần hợp giữa đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai vượt quá ngưỡng thứ hai hay không, nếu có, đi đến S45, nếu không, đi đến S46;

S45: xác định phần giao giữa đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai làm đoạn được chú thích cuối cùng đối với phần dữ liệu chuỗi thời gian; và

S46: gán thẻ tái duyệt cho phần hợp của phần dữ liệu chuỗi thời gian mà có khoảng thời gian từ bắt đầu đoạn được chú thích thứ nhất đến kết thúc đoạn được chú thích thứ hai.

Dựa vào FIG. 1C minh họa quy trình của phương pháp dán nhãn theo một phương án khác của sáng chế, phương pháp còn bao gồm bước S48 là bước khuyến nghị, sau khi xác định xem liệu có phần giao giữa mỗi đoạn được chú thích thứ nhất và mỗi đoạn được chú thích thứ hai, ít nhất một đoạn riêng lẻ của một đơn vị trong số các đơn vị tác tử thứ nhất và đơn vị tác tử thứ hai, ít nhất một đoạn riêng lẻ này không giao với bất kỳ đoạn nào của đơn vị còn lại trong số các đơn vị tác tử thứ nhất và đơn vị tác tử thứ hai, cho ít nhất một hoặc nhiều đơn vị trong số các đơn vị tác tử thứ nhất và đơn vị tác tử thứ hai làm các thẻ khuyến nghị. Bước S48 có thể được thực hiện sau khi thiết bị dán nhãn 130 xác định rằng tất cả các đoạn được dán nhãn thứ nhất và thứ hai được so sánh với nhau (S47).

Theo một phương án của sáng chế, các bước S2 và S3 để chú thích phần dữ liệu chuỗi thời gian được thực hiện bởi các mô hình trí tuệ nhân tạo (AI).

Theo một phương án của sáng chế, các bước S2 và S3 để chú thích phần dữ liệu chuỗi thời gian được thực hiện bởi người dùng.

Theo một phương án của sáng chế, đơn vị tác tử thứ nhất 110 và đơn vị tác tử thứ hai 120 bao gồm các mô hình trí tuệ nhân tạo (AI) để giúp người dùng tái duyệt khác biệt giữa các kết quả làm việc của người dùng trên phần dữ liệu chuỗi thời gian.

Theo một phương án của sáng chế, bước S43 còn bao gồm bước quay về S2 để chú thích phần dữ liệu chuỗi thời gian dựa vào nhãn của đoạn được chú thích thứ hai được gửi đến đơn vị tác tử thứ nhất 110; và quay về S3 để chú thích phần dữ liệu chuỗi thời gian dựa vào nhãn của đoạn được chú thích thứ nhất được gửi đến đơn vị tác tử thứ hai 120. Các bước từ S2 đến S43 có thể được lặp lại cho đến khi đoạn được chú thích cuối cùng được quyết định.

Theo một phương án của sáng chế, bước S46 còn bao gồm bước quay về S2 để chú thích phần dữ liệu chuỗi thời gian dựa vào thẻ tái duyệt được gán; và quay về S3 để chú thích phần dữ liệu chuỗi thời gian dựa vào thẻ tái duyệt được gán. Các bước từ S2 đến S46 có thể được lặp lại cho đến khi đoạn được chú thích cuối cùng được quyết định.

Theo một phương án của sáng chế, tỷ lệ sai lệch trên phần hợp (discrepancy-over-union, XoU) giữa đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai được tính bằng cách sử dụng phương trình sau:

$$XoU = \frac{(|start_1 - start_2| + |end_1 - end_2|)}{\max(\max(end_1, end_2) - \min(start_1, start_2), 10^{-6})}$$

trong đó, $start_1$ là bắt đầu của đoạn được chú thích thứ nhất, $start_2$ là bắt đầu của đoạn được chú thích thứ hai, end_1 là kết thúc của đoạn được chú thích thứ nhất, và end_2 là kết thúc của đoạn được chú thích thứ hai.

Fig.2 là lưu đồ minh họa quy trình của hệ thống dán nhãn theo một phương án của sáng chế.

Tham chiếu đến Fig.2, phương pháp dán nhãn theo một phương án của sáng chế có thể được thực hiện dưới dạng quy trình bao gồm bốn giai đoạn.

Giai đoạn thứ nhất có thể được thực hiện bởi nhiều đơn vị tác tử. Trong giai đoạn thứ nhất, mỗi đơn vị tác tử có thể nhận phần dữ liệu chuỗi thời gian. Và, đối với phần dữ liệu chuỗi thời gian này, mỗi đơn vị tác tử có thể chú thích phần dữ liệu chuỗi thời gian để tạo ra ít nhất một đoạn được chú thích. Mỗi trong số ít nhất một đoạn được chú thích xác định điểm bắt đầu và điểm kết thúc của các sự kiện (các nhãn) xảy ra bên trong phần dữ liệu chuỗi thời gian cho trước.

Phần dữ liệu chuỗi thời gian là chuỗi các phép đo được ghi từ cùng một nguồn trong các nhãn thời gian liên tiếp. Các phép đo này có thể là dữ liệu cảm biến, giá cổ phần, các khung ảnh, v.v. và nhãn trong phần dữ liệu chuỗi thời gian có thể được xem là điểm sự kiện đã xảy ra trong dữ liệu cho trước với ý nghĩa ngữ nghĩa được xác định, với ràng buộc là nhãn thời gian bắt đầu và nhãn thời gian kết thúc.

Giai đoạn thứ hai có thể được thực hiện trong thiết bị dán nhãn 130. Trong giai đoạn thứ hai, thiết bị dán nhãn 130 so sánh nhiều đoạn được chú thích và tạo ra khuyến nghị cho mỗi tác tử. Thiết bị dán nhãn 130 có thể tạo ra khuyến nghị nhờ điều kiện liên quan đến phần giao. Khi khuyến nghị được tạo ra, thì thiết bị dán nhãn 130 có thể gửi khuyến nghị đến nhiều đơn vị tác tử. Trong trường hợp này, giai đoạn thứ ba có thể tiến hành. Mặt khác, khi khuyến nghị không được tạo ra, thì thiết bị dán nhãn 130 có thể tiếp tục đến giai đoạn thứ tư.

Giai đoạn thứ ba có thể được thực hiện bởi nhiều đơn vị tác tử. Mỗi trong số các đơn vị tác tử có thể nhận khuyến nghị và cố định/thay đổi đoạn được chú thích dựa vào các đánh giá của chúng. Quy trình sẽ giống với giai đoạn thứ nhất, ngoại trừ là các tác tử đã được cho các gợi ý từ các khuyến nghị và các thẻ tái duyệt để xem xét lại các quyết định của chúng về kết quả hiện thời. Giai đoạn thứ ba có thể không được thực hiện nếu khuyến nghị không được tạo ra ở giai đoạn thứ hai.

Giai đoạn thứ tư có thể được thực hiện bởi thiết bị dán nhãn 130. Theo một phương án, thiết bị dán nhãn 130 có thể tạo ra đoạn được chú thích cuối cùng sử dụng nhiều đoạn được chú thích được tạo ra ở giai đoạn thứ nhất. Theo cách khác, thiết bị dán nhãn 130 có thể tạo ra đoạn được chú thích cuối cùng sử dụng nhiều đoạn được chú thích được điều chỉnh theo khuyến nghị.

Phần mô tả chi tiết của mỗi giai đoạn có thể được mô tả chi tiết với tham chiếu đến các hình vẽ dưới đây.

Fig.3 là lưu đồ thể hiện quy trình ở giai đoạn thứ nhất theo một phương án của sáng chế.

Đơn vị tác tử thứ nhất 110 có thể nhận phần dữ liệu chuỗi thời gian (S105).

Đơn vị tác tử thứ nhất 110 có thể xác định nhãn thứ nhất từ phần dữ liệu chuỗi thời gian. Đơn vị tác tử thứ nhất 110 có thể xác định nhãn thứ nhất bằng cách chọn điểm bắt đầu thứ nhất và điểm kết thúc thứ nhất. Nghĩa là, đơn vị tác tử thứ nhất 110 có thể chú thích đoạn giữa điểm bắt đầu thứ nhất và điểm kết thúc thứ nhất với nhãn thứ nhất.

Đơn vị tác tử thứ nhất 110 có thể tạo ra phần dữ liệu chuỗi thời gian được chú thích thứ nhất có ít nhất một đoạn được chú thích thứ nhất (S110).

Đơn vị tác tử thứ nhất 110 có thể gửi phần dữ liệu chuỗi thời gian được chú thích thứ nhất đến thiết bị dán nhãn 130 (S125).

Đơn vị tác tử thứ hai 120 có thể nhận phần dữ liệu chuỗi thời gian (S130).

Đơn vị tác tử thứ hai 120 có thể xác định nhãn thứ hai từ phần dữ liệu chuỗi thời gian. Đơn vị tác tử thứ hai 120 có thể xác định nhãn thứ hai bằng cách chọn điểm bắt đầu thứ hai và điểm kết thúc thứ hai. Nghĩa là, đơn vị tác tử thứ hai 120 có thể chú thích đoạn giữa điểm bắt đầu thứ hai và điểm kết thúc thứ hai với nhãn thứ hai.

Đơn vị tác tử thứ hai 120 có thể tạo ra phần dữ liệu chuỗi thời gian được chú thích thứ hai có ít nhất một đoạn được chú thích thứ hai (S135).

Đơn vị tác tử thứ hai 120 có thể gửi phần dữ liệu chuỗi thời gian được chú thích thứ hai đến thiết bị dán nhãn 130 (S150).

Fig.4 là lưu đồ thể hiện quy trình ở giai đoạn thứ hai theo một phương án của sáng chế. Các hình vẽ từ Fig.5A đến Fig.5D là các sơ đồ để giải thích giai đoạn thứ hai.

Thiết bị dán nhãn 130 có thể so sánh đoạn được chú thích thứ nhất với đoạn được chú thích thứ hai (S205). Cụ thể, thiết bị dán nhãn 130 có thể so sánh xem liệu đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai giao nhau hay không. Phần dữ liệu chuỗi thời gian được chú thích thứ nhất có thể bao gồm các đoạn được chú thích thứ nhất. Phần dữ liệu chuỗi thời gian được chú thích thứ hai có thể bao gồm các đoạn được chú thích thứ hai. Mỗi trong số các đoạn được chú thích thứ nhất có thể được so sánh với mỗi trong số các đoạn được chú thích thứ hai để xác định xem liệu có phần giao giữa đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai hay không.

Theo một phương án của sáng chế, khi đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai không giao nhau hoặc không còn cặp nào của đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai để so sánh, thì thiết bị dán nhãn 130 có thể kết thúc việc xử lý. Theo một phương án khác của sáng chế, khi không có cặp nào của đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai để so sánh, các đoạn riêng lẻ của một đơn vị trong số các đơn vị tác tử thứ nhất và đơn vị tác tử thứ hai, các đoạn riêng lẻ này không giao với bất kỳ đoạn nào của đơn vị còn lại trong số các đơn vị tác tử thứ nhất và đơn vị tác tử thứ hai, có thể được khuyến nghị cho ít nhất một hoặc nhiều đơn vị trong số các đơn vị tác tử thứ nhất và đơn vị tác tử thứ hai làm các thể khuyến nghị. Ví dụ, như được minh họa trên Fig.5A, đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai không giao nhau khi điểm kết thúc của đoạn được chú thích thứ nhất nhỏ hơn điểm bắt đầu của đoạn được chú thích thứ hai.

Khi đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai giao nhau, thì thiết bị dán nhãn 130 có thể xác định xem liệu đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai có các nhãn khác nhau hay không. Đáp lại việc xác định rằng đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai có các nhãn khác nhau,

thiết bị dán nhãn 130 có thể tạo ra khuyến nghị cho đơn vị tác tử thứ nhất 110 và khuyến nghị cho đơn vị tác tử thứ hai 120. Theo một phương án, như được thể hiện trên Fig.5B, thiết bị dán nhãn 130 có thể tạo ra nhãn thứ hai làm khuyến nghị thứ nhất cho đơn vị tác tử thứ nhất 110 và nhãn thứ nhất làm khuyến nghị thứ hai cho đơn vị tác tử thứ hai 120.

Thiết bị dán nhãn 130 có thể gửi các khuyến nghị đã tạo đến đơn vị tác tử thứ nhất 110 và đơn vị tác tử thứ hai 120 (S215). Thiết bị dán nhãn 130 có thể gửi khuyến nghị thứ nhất đến đơn vị tác tử thứ nhất 110 và khuyến nghị thứ hai đến đơn vị tác tử thứ hai 120.

Khi đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai giao nhau, thì thiết bị dán nhãn 130 có thể xác định xem liệu đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai có các nhãn khác nhau hay không. Đáp lại việc xác định rằng đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai có cùng các nhãn, thiết bị dán nhãn 130 có thể tính toán tỷ lệ sai lệch trên phần hợp giữa đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai (S220). Ví dụ, như được thể hiện trên Fig.5C, khi có vùng giao giữa đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai trên dòng thời gian, thì thiết bị dán nhãn 130 có thể tính toán tỷ lệ sai lệch trên phần hợp giữa đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai.

Thiết bị dán nhãn 130 có thể tính toán tỷ lệ sai lệch trên phần hợp (discrepancy-over-union, XoU) sử dụng phương trình 1 dưới đây.

[Phương trình 1]

$$XoU = \frac{(|start_1 - start_2| + |end_1 - end_2|)}{\max(\max(end_1, end_2) - \min(start_1, start_2), 10^{-6})}$$

Ở đây, $start_1$ là bắt đầu của đoạn được chú thích thứ nhất, $start_2$ là bắt đầu của đoạn được chú thích thứ hai, end_1 là kết thúc của đoạn được chú thích thứ nhất, và end_2 là kết thúc của đoạn được chú thích thứ hai. Đoạn giao có khoảng thời gian từ bắt đầu đoạn được chú thích thứ hai đến kết thúc đoạn được chú thích thứ nhất. Phần hợp có khoảng thời gian từ bắt đầu đoạn được chú thích thứ nhất đến kết thúc đoạn được chú thích thứ hai.

Thiết bị dán nhãn 130 có thể so sánh tỷ lệ sai lệch trên phân hợp với ngưỡng thứ nhất định trước (S225).

Kết quả của việc so sánh, nếu tỷ lệ sai lệch trên phân hợp đã tính lớn hơn hoặc bằng ngưỡng này, thiết bị dán nhãn 130 có thể tạo ra khuyến nghị (S230). Cụ thể, thiết bị dán nhãn 130 có thể tạo ra khuyến nghị thứ ba kết hợp đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai làm khuyến nghị. Ví dụ, như được thể hiện trên Fig.5D, thiết bị dán nhãn 130 có thể tạo ra khuyến nghị thứ ba cho phân hợp có khoảng thời gian từ bắt đầu đoạn được chú thích thứ nhất đến kết thúc đoạn được chú thích thứ hai. Khuyến nghị thứ ba có thể là thẻ tái duyệt mà đánh dấu phân hợp có trạng thái cần được tái duyệt.

Thiết bị dán nhãn 130 có thể gửi khuyến nghị đã tạo đến đơn vị tác tử thứ nhất 110 và đơn vị tác tử thứ hai 120 (S215). Thiết bị dán nhãn 130 có thể truyền khuyến nghị thứ ba đến đơn vị tác tử thứ nhất 110 và đơn vị tác tử thứ hai 120.

Fig.6 là lưu đồ thể hiện quy trình ở giai đoạn thứ ba theo một phương án của sáng chế.

Đơn vị tác tử thứ nhất 110 có thể điều chỉnh dữ liệu chuỗi thời gian được chú thích thứ nhất dựa vào khuyến nghị đã nhận (S305).

Đơn vị tác tử thứ nhất 110 có thể chú thích phần dữ liệu chuỗi thời gian dựa vào khuyến nghị thứ nhất. Ví dụ, đơn vị tác tử thứ nhất 110 có thể điều chỉnh nhãn của đoạn được chú thích thứ nhất dựa vào nhãn của đoạn được chú thích thứ hai được gửi đến đơn vị tác tử thứ nhất 110.

Đơn vị tác tử thứ nhất 110 có thể chú thích phần dữ liệu chuỗi thời gian dựa vào khuyến nghị thứ ba. Ví dụ, đơn vị tác tử thứ nhất 110 có thể điều chỉnh khoảng thời gian của đoạn được chú thích thứ nhất dựa vào thẻ tái duyệt được gán của phân hợp được gửi đến đơn vị tác tử thứ nhất 110.

Đơn vị tác tử thứ nhất 110 có thể gửi phần dữ liệu chuỗi thời gian được điều chỉnh thứ nhất đến thiết bị dán nhãn 130 (S320).

Đơn vị tác tử thứ hai 120 có thể điều chỉnh dữ liệu chuỗi thời gian được chú thích thứ hai dựa vào khuyến nghị đã nhận (S325).

Đơn vị tác tử thứ hai 120 có thể chú thích phần dữ liệu chuỗi thời gian dựa vào khuyến nghị thứ nhất. Ví dụ, đơn vị tác tử thứ hai 120 có thể điều chỉnh nhãn của đoạn được chú thích thứ hai dựa vào nhãn của đoạn được chú thích thứ nhất được gửi đến đơn vị tác tử thứ hai 120.

Đơn vị tác tử thứ hai 120 có thể chú thích phần dữ liệu chuỗi thời gian dựa vào khuyến nghị thứ ba. Ví dụ, đơn vị tác tử thứ hai 120 có thể điều chỉnh khoảng thời gian của đoạn được chú thích thứ hai dựa vào thẻ tái duyệt được gán của phần hợp được gửi đến đơn vị tác tử thứ hai 120.

Đơn vị tác tử thứ hai 120 có thể gửi phần dữ liệu chuỗi thời gian được điều chỉnh thứ hai đến thiết bị dán nhãn 130 (S340).

Sau đó, thiết bị dán nhãn 130 có thể thực hiện quy trình ở giai đoạn thứ hai dựa vào phần dữ liệu chuỗi thời gian được điều chỉnh thứ nhất và phần dữ liệu chuỗi thời gian được điều chỉnh thứ hai.

Fig.7 là lưu đồ thể hiện quy trình ở giai đoạn thứ tư theo một phương án của sáng chế.

Đầu tiên, thiết bị dán nhãn 130 có thể thu đoạn được chú thích cuối cùng dựa vào phần dữ liệu chuỗi thời gian được chú thích thứ nhất và phần dữ liệu chuỗi thời gian được chú thích thứ hai được nhận từ đơn vị tác tử thứ nhất 110 và đơn vị tác tử thứ hai 120 (S405). Phần dữ liệu chuỗi thời gian được chú thích thứ nhất và phần dữ liệu chuỗi thời gian được chú thích thứ hai có thể là phần dữ liệu chuỗi thời gian được chú thích bởi đơn vị tác tử thứ nhất 110 và đơn vị tác tử thứ hai 120, một cách tương ứng, có hoặc không có các khuyến nghị.

Theo một phương án, thiết bị dán nhãn 130 có thể thu miền giao giữa đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai làm đoạn được chú thích cuối cùng. Nghĩa là, khi đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai không được điều chỉnh do khuyến nghị cho đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai không được tạo ra, thì thiết bị dán nhãn 130 có thể thu vùng giao giữa đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai làm đoạn được chú thích cuối cùng.

Theo một phương án, thiết bị dán nhãn 130 có thể thu miền giao giữa đoạn được chú thích thứ nhất đã được điều chỉnh và đoạn được chú thích thứ hai đã được điều chỉnh làm đoạn được chú thích cuối cùng. Nghĩa là, khi khuyến nghị cho đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai được tạo ra để điều chỉnh đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai, thì thiết bị dán nhãn 130 có thể thu miền giao giữa đoạn được chú thích thứ nhất đã được điều chỉnh và đoạn được chú thích thứ hai đã được điều chỉnh làm đoạn được chú thích cuối cùng.

Các hiệu quả và ưu điểm có lợi khác nhau của sáng chế không bị giới hạn ở các phần mô tả trên và sẽ dễ dàng được hiểu hơn bằng cách mô tả các phương án cụ thể của sáng chế.

Các thuật ngữ được mô tả trong bản mô tả này như “bộ phận” đề cập đến thành phần mềm hoặc phần cứng như mảng cổng có thể lập trình theo trường (field-programmable gate array, FPGA) hoặc mạch tích hợp chuyên dụng (Application-Specific Integrated Circuit, ASIC), và bộ phận thực hiện các chức năng nhất định. Tuy nhiên, “bộ phận” không bị giới hạn ở phần mềm hoặc phần cứng. “Bộ phận” có thể được tạo cấu hình dưới dạng vật ghi có thể được xử lý hoặc có thể được tạo cấu hình để được thực thi bởi ít nhất một bộ xử lý. Vì vậy, các ví dụ về “bộ phận” bao gồm các thành phần như các thành phần mềm, các thành phần mềm định hướng đối tượng, các thành phần lớp và các thành phần tác vụ, và quy trình, chức năng, thuộc tính, thủ tục, thủ tục con, đoạn mã chương trình, trình điều khiển, phân sụn, vi mã, mạch, dữ liệu, cơ sở dữ liệu, cấu trúc dữ liệu, bảng, và các biến. Các thành phần và chức năng được bố trí nằm trong “các bộ phận” có thể được kết hợp thành số lượng nhỏ hơn các thành phần và “các bộ phận” hoặc có thể còn được tách khỏi các thành phần và “các bộ phận” bổ sung. Ngoài ra, các thành phần và “các bộ phận” có thể được thực hiện để đọc lại một hoặc nhiều bộ xử lý trung tâm (central processing unit, CPU) trong thiết bị hoặc thẻ đa phương tiện an toàn.

Trong khi sáng chế đã được mô tả với tham chiếu đến các phương án, các phương án này chỉ là các phương án làm ví dụ của sáng chế và không làm giới hạn sáng chế, người có hiểu biết trung bình về lĩnh vực kỹ thuật này sẽ hiểu rằng các sửa đổi và ứng dụng khác nhau, mà không được minh họa trong phần mô tả trên, sẽ được thực hiện mà không tách rời khỏi phạm vi các đặc tính thiết yếu của các phương án làm ví dụ này. Ví dụ, mỗi thành phần được mô tả chi tiết theo các phương án này có

thể được thay đổi. Ngoài ra, nên hiểu rằng sự khác nhau liên quan đến các sửa đổi và ứng dụng này nằm trong phạm vi của sáng chế được xác định trong bộ yêu cầu bảo hộ kèm theo.

YÊU CẦU BẢO HỘ

1. Phương pháp dán nhãn bao gồm các bước:

S1: cấp phần dữ liệu chuỗi thời gian cho đơn vị tác tử thứ nhất và đơn vị tác tử thứ hai;

S2: chú thích, bởi đơn vị tác tử thứ nhất, phần dữ liệu chuỗi thời gian để tạo ra phần dữ liệu chuỗi thời gian được chú thích thứ nhất có ít nhất một đoạn được chú thích thứ nhất;

S3: chú thích, bởi đơn vị tác tử thứ hai, phần dữ liệu chuỗi thời gian để tạo ra phần dữ liệu chuỗi thời gian được chú thích thứ hai có ít nhất một đoạn được chú thích thứ hai;

S4: xử lý, bởi thiết bị dán nhãn, phần dữ liệu chuỗi thời gian được chú thích thứ nhất và phần dữ liệu chuỗi thời gian được chú thích thứ hai, trong đó bước xử lý bao gồm các bước:

S41: xác định xem liệu có phần giao giữa đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai hay không, nếu có, đi đến S42, nếu không, kết thúc việc xử lý;

S42: xác định xem liệu đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai có các nhãn khác nhau hay không, nếu có, đi đến S43, nếu không, đi đến S44;

S43: gửi nhãn của đoạn được chú thích thứ nhất đến đơn vị tác tử thứ hai và nhãn của đoạn được chú thích thứ hai đến đơn vị tác tử thứ nhất;

S44: xác định xem liệu tỷ lệ sai lệch trên phần hợp giữa đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai dưới ngưỡng thứ nhất hay không, nếu có, đi đến S45, nếu không, đi đến S46;

S45: xác định phần giao giữa đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai làm đoạn được chú thích cuối cùng đối với phần dữ liệu chuỗi thời gian; và

S46: gán thẻ tái duyệt cho phần hợp của phần dữ liệu chuỗi thời gian mà có khoảng thời gian từ bắt đầu đoạn được chú thích thứ nhất đến kết thúc đoạn được chú thích thứ hai.

2. Phương pháp dán nhãn theo điểm 1, trong đó phương pháp này còn bao gồm bước S48 là bước khuyến nghị, sau khi xác định xem liệu có phân giao giữa mỗi đoạn được chú thích thứ nhất và mỗi đoạn được chú thích thứ hai, ít nhất một đoạn riêng lẻ của một đơn vị trong số các đơn vị tác tử thứ nhất và đơn vị tác tử thứ hai, ít nhất một đoạn riêng lẻ này không giao với bất kỳ đoạn nào của đơn vị còn lại trong số các đơn vị tác tử thứ nhất và đơn vị tác tử thứ hai, cho ít nhất một hoặc nhiều đơn vị trong số các đơn vị tác tử thứ nhất và đơn vị tác tử thứ hai làm các thẻ khuyến nghị.

3. Phương pháp dán nhãn theo điểm 1, trong đó các bước S2 và S3 để chú thích phần dữ liệu chuỗi thời gian được thực hiện bởi các mô hình trí tuệ nhân tạo (artificial intelligence, AI).

4. Phương pháp dán nhãn theo điểm 1, trong đó các bước S2 và S3 để chú thích phần dữ liệu chuỗi thời gian được thực hiện bởi người dùng.

5. Phương pháp dán nhãn theo điểm 1, trong đó đơn vị tác tử thứ nhất và đơn vị tác tử thứ hai bao gồm các mô hình trí tuệ nhân tạo (AI) để giúp người dùng tái duyệt khác biệt giữa các kết quả làm việc của người dùng trên phần dữ liệu chuỗi thời gian.

6. Phương pháp dán nhãn theo điểm 1, trong đó bước S44 bao gồm bước xác định xem liệu tỷ lệ phân giao trên phần hợp giữa đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai vượt quá ngưỡng thứ hai hay không, nếu có, đi đến S45, nếu không, đi đến S46.

7. Phương pháp dán nhãn theo điểm 1, trong đó bước S43 còn bao gồm các bước:

quay về S2 để chú thích phần dữ liệu chuỗi thời gian dựa vào nhãn của đoạn được chú thích thứ hai được gửi đến đơn vị tác tử thứ nhất; và

quay về S3 để chú thích phần dữ liệu chuỗi thời gian dựa vào nhãn của đoạn được chú thích thứ nhất được gửi đến đơn vị tác tử thứ hai.

8. Phương pháp dán nhãn theo điểm 1, trong đó bước S46 còn bao gồm các bước:

quay về S2 để chú thích phần dữ liệu chuỗi thời gian dựa vào thẻ tái duyệt được gán; và

quay về S3 để chú thích phần dữ liệu chuỗi thời gian dựa vào thẻ tái duyệt được gán.

9. Phương pháp dán nhãn theo điểm 1, trong đó tỷ lệ sai lệch trên phần hợp giữa đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai được tính bằng cách sử dụng phương trình sau:

$$XoU = \frac{(|start_1 - start_2| + |end_1 - end_2|)}{\max(\max(end_1, end_2) - \min(start_1, start_2), 10^{-6})}$$

trong đó, $start_1$ là bắt đầu của đoạn được chú thích thứ nhất, $start_2$ là bắt đầu của đoạn được chú thích thứ hai, end_1 là kết thúc của đoạn được chú thích thứ nhất, và end_2 là kết thúc của đoạn được chú thích thứ hai.

10. Hệ thống dán nhãn bao gồm:

đơn vị tác tử thứ nhất được tạo cấu hình để chú thích phần dữ liệu chuỗi thời gian để tạo ra phần dữ liệu chuỗi thời gian được chú thích thứ nhất có ít nhất một đoạn được chú thích thứ nhất;

đơn vị tác tử thứ hai được tạo cấu hình để chú thích phần dữ liệu chuỗi thời gian để tạo ra phần dữ liệu chuỗi thời gian được chú thích thứ hai có ít nhất một đoạn được chú thích thứ hai;

thiết bị dán nhãn được tạo cấu hình để xử lý phần dữ liệu chuỗi thời gian được chú thích thứ nhất và phần dữ liệu chuỗi thời gian được chú thích thứ hai, trong đó bước xử lý bao gồm các bước:

S41: xác định xem liệu có phần giao giữa đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai hay không, nếu có, đi đến S42, nếu không, kết thúc việc xử lý;

S42: xác định xem liệu đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai có các nhãn khác nhau hay không, nếu có, đi đến S43, nếu không, đi đến S44;

S43: gửi nhãn của đoạn được chú thích thứ nhất đến đơn vị tác tử thứ hai và nhãn của đoạn được chú thích thứ hai đến đơn vị tác tử thứ nhất;

S44: xác định xem liệu tỷ lệ sai lệch trên phần hợp (discrepancy-over-union, XoU) giữa đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai dưới ngưỡng thứ nhất hay không, nếu có, đi đến S45, nếu không, đi đến S46;

S45: xác định phần giao giữa đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai làm đoạn được chú thích cuối cùng đối với phần dữ liệu chuỗi thời gian; và

S46: gán thẻ tái duyệt cho phần hợp của phần dữ liệu chuỗi thời gian mà có khoảng thời gian từ bắt đầu đoạn được chú thích thứ nhất đến kết thúc đoạn được chú thích thứ hai.

11. Hệ thống dán nhãn theo điểm 10, trong đó hệ thống này còn được tạo cấu hình đề khuyến nghị, sau khi xác định xem liệu có phần giao giữa mỗi đoạn được chú thích thứ nhất và mỗi đoạn được chú thích thứ hai, ít nhất một đoạn riêng lẻ của một đơn vị trong số các đơn vị tác tử thứ nhất và đơn vị tác tử thứ hai, ít nhất một đoạn riêng lẻ này không giao với bất kỳ đoạn nào của đơn vị còn lại trong số các đơn vị tác tử thứ nhất và đơn vị tác tử thứ hai, cho ít nhất một hoặc nhiều đơn vị trong số các đơn vị tác tử thứ nhất và đơn vị tác tử thứ hai làm các thẻ khuyến nghị.

12. Hệ thống dán nhãn theo điểm 10, trong đó đơn vị tác tử thứ nhất và đơn vị tác tử thứ hai sử dụng các mô hình trí tuệ nhân tạo (AI) để chú thích phần dữ liệu chuỗi thời gian.

13. Hệ thống dán nhãn theo điểm 10, trong đó đơn vị tác tử thứ nhất và đơn vị tác tử thứ hai được sử dụng bởi người dùng để chú thích phần dữ liệu chuỗi thời gian.

14. Hệ thống dán nhãn theo điểm 10, trong đó đơn vị tác tử thứ nhất và đơn vị tác tử thứ hai bao gồm các mô hình trí tuệ nhân tạo (AI) để giúp người dùng tái duyệt khác biệt giữa các kết quả làm việc của người dùng trên phần dữ liệu chuỗi thời gian.

15. Hệ thống dán nhãn theo điểm 10, trong đó bước S44 bao gồm bước xác định xem liệu tỷ lệ phần giao trên phần hợp giữa đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai vượt quá ngưỡng thứ hai hay không, nếu có, đi đến S45, nếu không, đi đến S46.

16. Hệ thống dán nhãn theo điểm 10, trong đó

đơn vị tác tử thứ nhất được tạo cấu hình để chú thích phần dữ liệu chuỗi thời gian dựa vào nhãn của đoạn được chú thích thứ hai được gửi đến đơn vị tác tử thứ nhất; và

đơn vị tác tử thứ hai được tạo cấu hình để chú thích phần dữ liệu chuỗi thời gian dựa vào nhãn của đoạn được chú thích thứ nhất được gửi đến đơn vị tác tử thứ hai.

17. Hệ thống dán nhãn theo điểm 10, trong đó

đơn vị tác tử thứ nhất được tạo cấu hình để chú thích phần dữ liệu chuỗi thời gian dựa vào thẻ tái duyệt được gán; và

đơn vị tác tử thứ hai được tạo cấu hình để chú thích phần dữ liệu chuỗi thời gian dựa vào thẻ tái duyệt được gán.

18. Hệ thống dán nhãn theo điểm 10, trong đó tỷ lệ sai lệch trên phần hợp giữa đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai được tính bằng cách sử dụng phương trình sau:

$$XoU = \frac{(|start_1 - start_2| + |end_1 - end_2|)}{\max(\max(end_1, end_2) - \min(start_1, start_2), 10^{-6})}$$

trong đó, $start_1$ là bắt đầu của đoạn được chú thích thứ nhất, $start_2$ là bắt đầu của đoạn được chú thích thứ hai, end_1 là kết thúc của đoạn được chú thích thứ nhất, và end_2 là kết thúc của đoạn được chú thích thứ hai.

19. Vật ghi lâu dài đọc được bằng máy tính bao gồm các lệnh mà, khi được thực thi bởi ít nhất một bộ xử lý của máy, khiến cho máy này thực hiện các hoạt động bao gồm:

S1: cấp phần dữ liệu chuỗi thời gian cho đơn vị tác tử thứ nhất và đơn vị tác tử thứ hai;

S2: chú thích, bởi đơn vị tác tử thứ nhất, phần dữ liệu chuỗi thời gian để tạo ra phần dữ liệu chuỗi thời gian được chú thích thứ nhất có ít nhất một đoạn được chú thích thứ nhất;

S3: chú thích, bởi đơn vị tác tử thứ hai, phần dữ liệu chuỗi thời gian để tạo ra phần dữ liệu chuỗi thời gian được chú thích thứ hai có ít nhất một đoạn được chú thích thứ hai;

S4: xử lý, bởi thiết bị dán nhãn, phần dữ liệu chuỗi thời gian được chú thích thứ nhất và phần dữ liệu chuỗi thời gian được chú thích thứ hai, trong đó bước xử lý bao gồm các bước:

S41: xác định xem liệu có phần giao giữa đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai hay không, nếu có, đi đến S42, nếu không, kết thúc việc xử lý;

S42: xác định xem liệu đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai có các nhãn khác nhau hay không, nếu có, đi đến S43, nếu không, đi đến S44;

S43: gửi nhãn của đoạn được chú thích thứ nhất đến đơn vị tác tử thứ hai và nhãn của đoạn được chú thích thứ hai đến đơn vị tác tử thứ nhất;

S44: xác định xem liệu tỷ lệ sai lệch trên phần hợp giữa đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai dưới ngưỡng thứ nhất hay không, nếu có, đi đến S45, nếu không, đi đến S46;

S45: xác định phần giao giữa đoạn được chú thích thứ nhất và đoạn được chú thích thứ hai làm đoạn được chú thích cuối cùng đối với phần dữ liệu chuỗi thời gian; và

S46: gán thẻ tái duyệt cho phần hợp của phần dữ liệu chuỗi thời gian mà có khoảng thời gian từ bắt đầu đoạn được chú thích thứ nhất đến kết thúc đoạn được chú thích thứ hai.

FIG. 1A

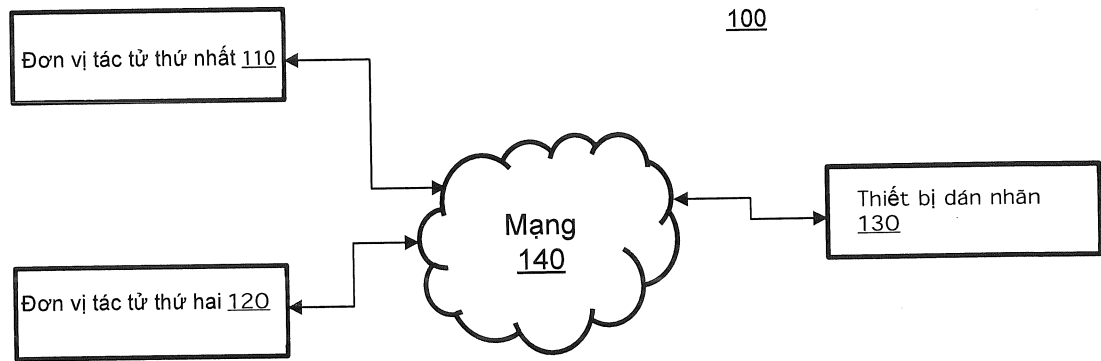


FIG. 1B

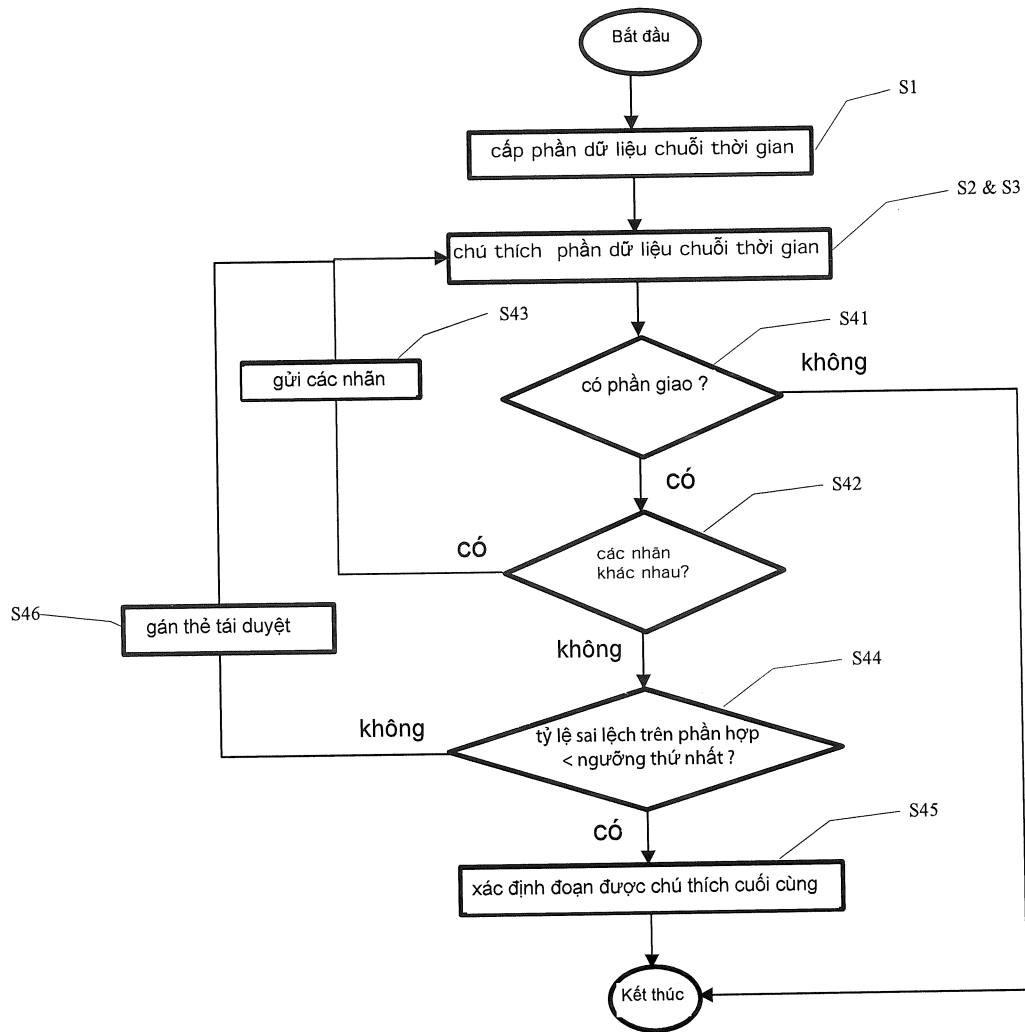


FIG. 1C

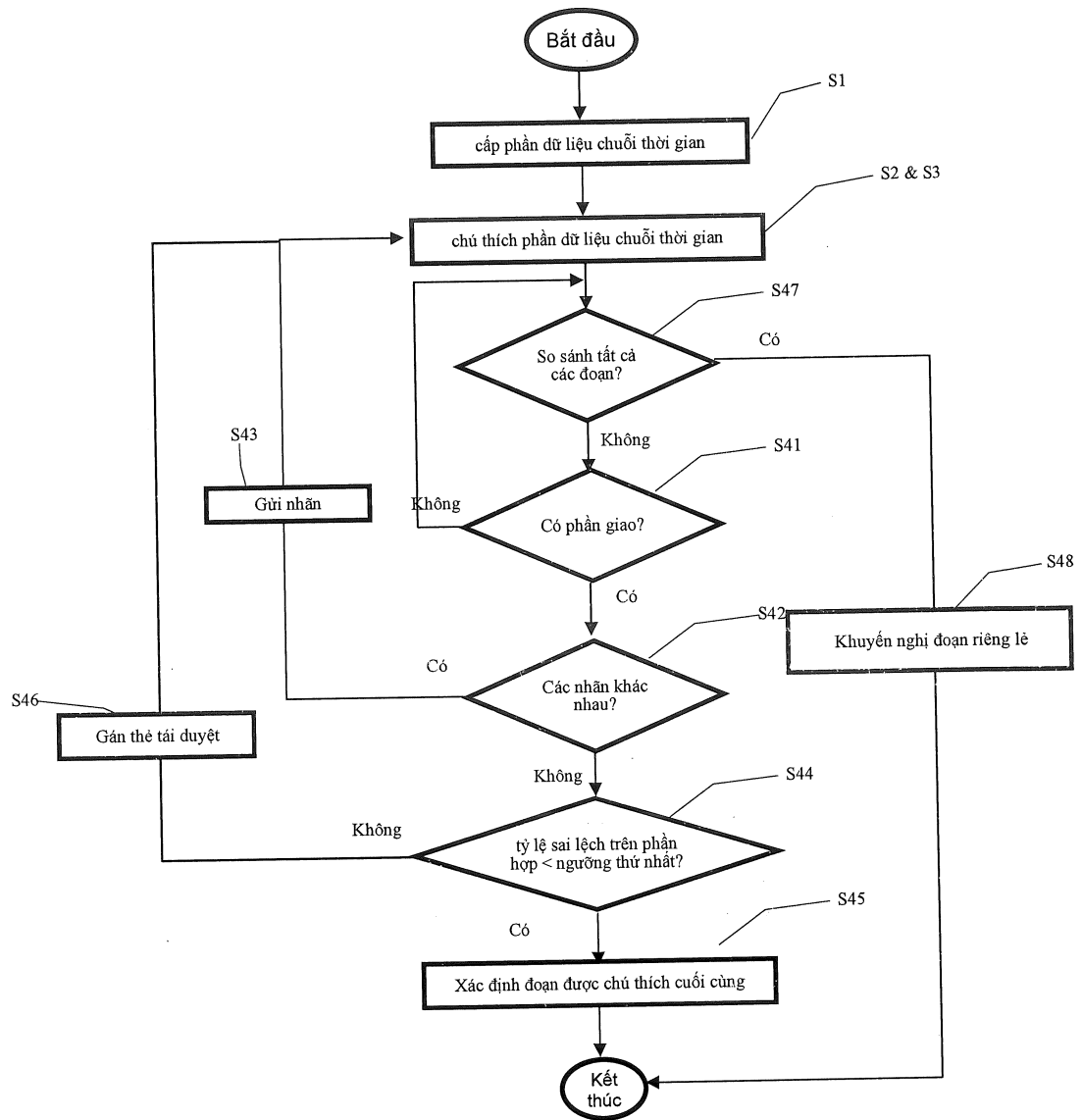


FIG. 2

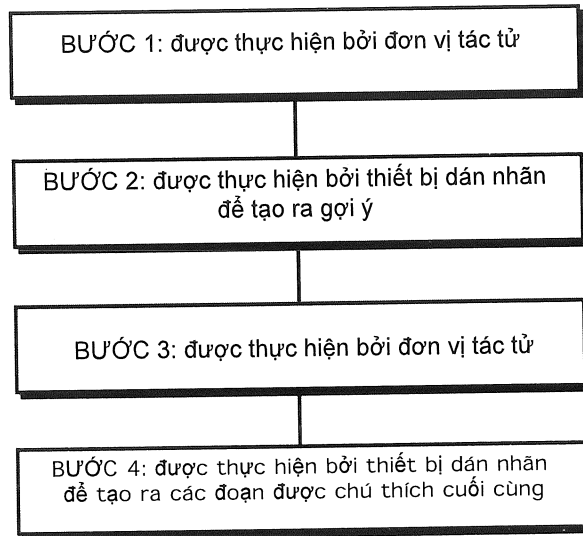


FIG. 3

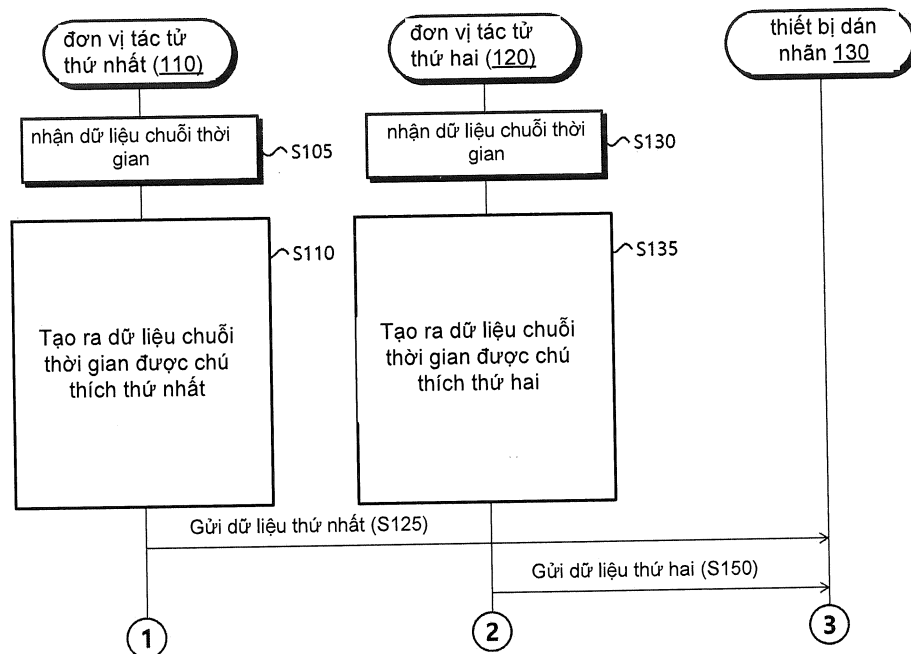


FIG. 4

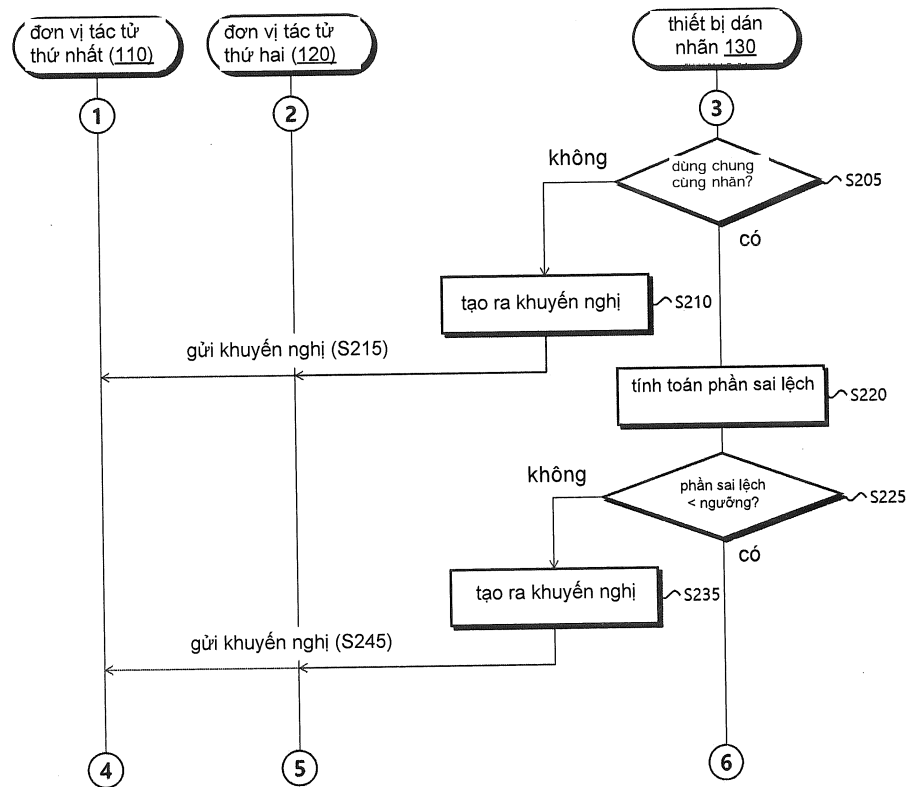


FIG. 5A

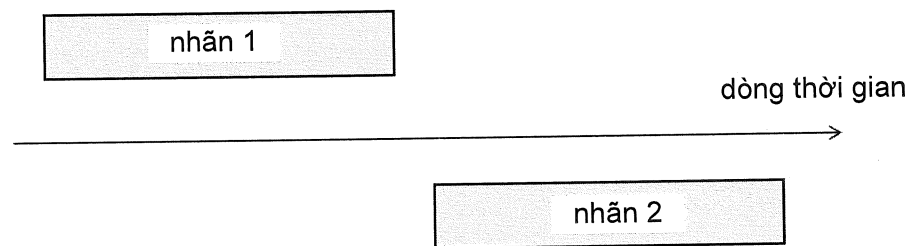


FIG. 5B

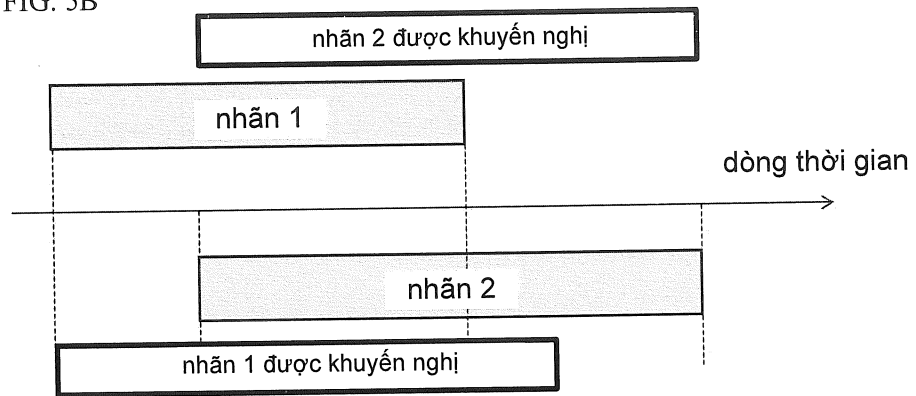


FIG. 5C

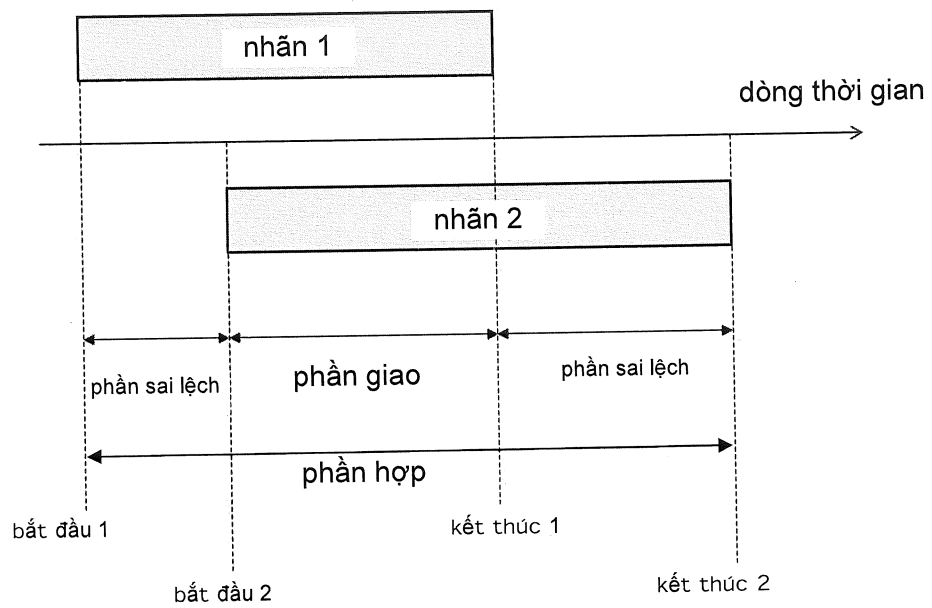


FIG. 5D

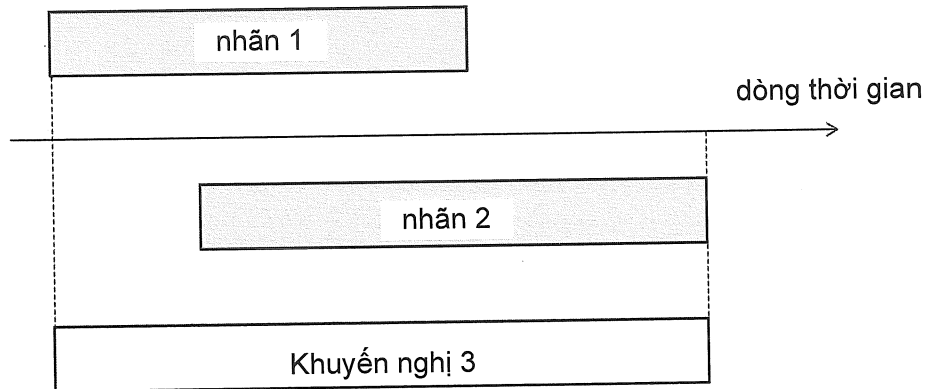


FIG. 6

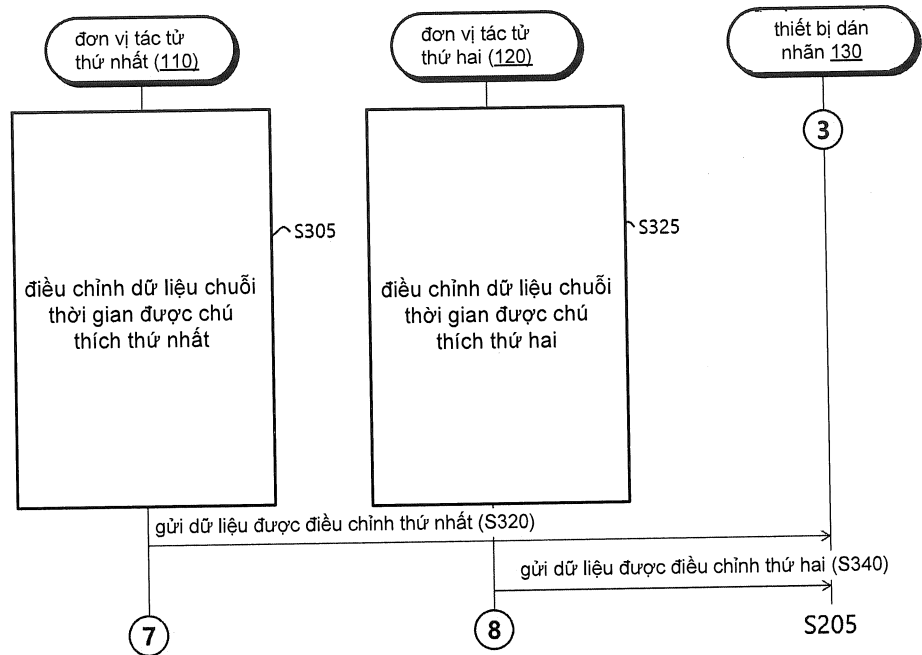


FIG. 7

