



(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)



1-0035225

(51)⁷ G06F 17/30; G06T 7/12; G06T 7/13;
G06K 9/46

(13) B

(21) 1-2018-05755

(22) 07/09/2017

(86) PCT/SG2017/050450 07/09/2017

(87) WO/2018/048356 15/03/2018

(30) 62/384,872 08/09/2016 US; 62/384,860 08/09/2016 US

(45) 25/04/2023 421

(43) 27/05/2019 374A

(73) GOH SOO SIAH (SG)

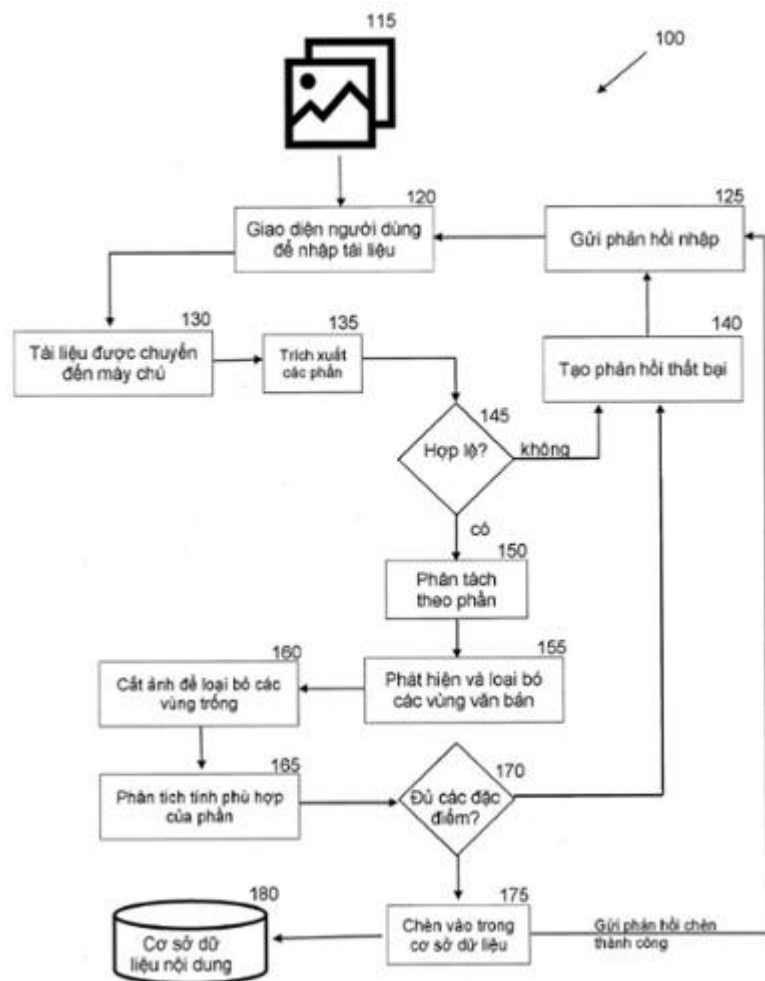
23 Greenleaf Drive, Ban Guan Park, Singapore, 279536

(72) MOORE, Stephen Maurice (SG); MOORE, Jimmy Daniel (SG); MURRAY, Larry Patrick (SG).

(74) Công ty Cổ phần Sở hữu công nghiệp INVESTIP (INVESTIP)

(54) PHƯƠNG PHÁP BIÊN DỊCH KHUNG HÌNH TỪ VIDEO, LƯU VÀ LẬP CHỈ MỤC KHUNG HÌNH TRONG CƠ SỞ DỮ LIỆU, HỆ THỐNG MÁY TÍNH ĐỂ NHẬP VIDEO, PHƯƠNG PHÁP BIÊN DỊCH HÌNH ẢNH TỪ PHƯƠNG TIỆN DẠNG BẢN IN VÀO TRONG CƠ SỞ DỮ LIỆU, HỆ THỐNG MÁY TÍNH ĐỂ NHẬP TÀI LIỆU

(57) Sáng chế đề cập đến phương pháp biên dịch các khung hình từ video, lưu và lập chỉ mục các khung hình trong cơ sở dữ liệu, hệ thống máy tính để nhập video, phương pháp biên dịch các hình ảnh từ phương tiện dạng bản in vào trong cơ sở dữ liệu, hệ thống máy tính để nhập các tài liệu. Khung phần mềm và phương pháp này cho phép chuyển đổi nội dung video thành định dạng cho phép ánh xạ và, do đó, xác định các phần riêng lẻ (phân đoạn cảnh hoặc dạng khác) của nội dung phương tiện. Trong đó bao gồm phương tiện nhập hình ảnh và các khung hình video vào cơ sở dữ liệu. Các sản phẩm tiêu dùng có thể được kết hợp với hình ảnh và/hoặc đối tượng trong hình ảnh. Người dùng có thể truy cập nội dung bằng cách gửi hình ảnh kỹ thuật số được chụp trên thiết bị di động. Một hoặc nhiều đối tượng trong hình ảnh do người dùng gửi có thể được so khớp với các sản phẩm và/hoặc tài liệu thương mại/quảng cáo.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến ứng dụng được thực hiện bằng máy tính được nối mạng, để thu thập, xử lý, phân tích và nhập nội dung phương tiện video để nội dung có thể được truy cập bởi nền tảng tìm kiếm trực quan.

Tình trạng kỹ thuật của sáng chế

Thương mại điện tử là sự giao dịch gồm mua hoặc bán trực tuyến. Thương mại điện tử đã trở thành công cụ quan trọng cho các doanh nghiệp lớn và nhỏ trên toàn thế giới, không chỉ để bán hàng cho khách hàng, mà còn để thu hút họ. Trong năm 2012, doanh số bán hàng thương mại điện tử đã vượt quá 1 nghìn tỷ đô la.

Tiếp thị qua internet đề cập đến các nỗ lực quảng cáo và tiếp thị sử dụng Web và e-mail để thúc đẩy bán hàng qua thương mại điện tử. Điều này bao gồm tiếp thị qua e-mail, tiếp thị qua các công cụ tìm kiếm (SEM), tiếp thị qua truyền thông qua mạng xã hội, nhiều loại quảng cáo hiển thị (ví dụ, quảng cáo qua biểu ngữ) và quảng cáo trên thiết bị di động. Siêu dữ liệu là thành phần không thể thiếu của tiếp thị qua internet.

Các doanh nghiệp thường lưu trữ siêu dữ liệu về các hoạt động tìm kiếm và giao dịch mà cho phép họ phân tích các xu hướng bán hàng, phát triển các kế hoạch tiếp thị và đưa ra các dự đoán. Siêu dữ liệu này cũng đồng thời có thể cho phép các doanh nghiệp cung cấp trải nghiệm mua sắm được cá nhân hóa hơn với các tính năng như lịch sử mua hàng, sở địa chỉ cho nhiều vị trí giao hàng và các đề xuất sản phẩm.

Ngày nay, hầu hết các trang web đều có siêu dữ liệu được nhúng trong chúng. Các công cụ tìm kiếm trên web xây dựng các chỉ mục rộng lớn mà sử dụng văn bản trang và siêu dữ liệu kèm theo của nó để cung cấp các kết quả tìm kiếm có liên quan cho người dùng. Siêu dữ liệu có thể được sử dụng để quảng cáo được nhắm mục tiêu. Các nhà quảng cáo có thể sử dụng các phương pháp phức tạp để nhắm mục tiêu đến các đối tượng tiếp nhận nhiều nhất với các đặc điểm nhất định, dựa trên sản phẩm mà nhà quảng cáo đang quảng bá.

Vì người mua sắm trực tuyến không thể xem từng sản phẩm riêng, họ thường tìm kiếm theo các tiêu chí như từ khóa. Ví dụ, người sử dụng có thể sử dụng trình duyệt web để tìm kiếm các chuyến bay của hãng hàng không đến New Zealand. Siêu dữ liệu dưới dạng "cookie" từ các trang web được người sử dụng truy cập được lưu trữ trên máy tính bởi trình duyệt web của người dùng. Cookie được gửi qua lại giữa máy chủ internet và trình duyệt cho phép người dùng được xác định và/hoặc theo dõi hoạt động của họ. Sau đó, người sử dụng này có thể nhận được các quảng cáo như biển quảng cáo qua biểu ngữ có liên quan đến du lịch ở New Zealand với thông tin về khách sạn, cho thuê xe hơi, chuyến tham quan và thông tin chuyến bay.

Hơn nữa, siêu dữ liệu có thể được sử dụng để nhắm đến người dùng dựa trên các thống kê về khách hàng. Doanh nghiệp có thể nhận biết được rằng sản phẩm thu hút đối tượng khách hàng nhất định và việc tiếp thị có thể được hướng đến đối tượng khách hàng đó. Ví dụ, quảng cáo qua biểu ngữ cho chứng khoán đầu tư có thể không hiệu quả đối với đối tượng vị thành niên. Việc nhắm mục tiêu quảng cáo đến đối tượng khách hàng già hơn, và cụ thể hơn là đối với những người tìm kiếm hưu trí, sẽ có hiệu quả hơn. Siêu dữ liệu có thể được biên dịch trên người dùng để xác định đối tượng khách hàng và khả năng mà khách hàng này có thể quan tâm đến chứng khoán đầu tư.

Với sự kỳ vọng của thị trường mà thể hiện sự gia tăng ngày càng tăng về các trường hợp sử dụng tìm kiếm trực quan, việc làm giàu hồ sơ đối tượng khách hàng là mối quan tâm lớn và mang lợi ích cho người dùng cũng như các doanh nghiệp. Các tổ chức có thể sử dụng các mô hình như vậy để cung cấp các dịch vụ được cá nhân hóa, tính toán thói quen sử dụng (ví dụ, dựa trên tính thời vụ), và giúp xác định hướng sản phẩm trong tương lai.

Các tìm kiếm dựa trên hình ảnh là phổ biến, đặc biệt với sự phổ biến của điện thoại thông minh và máy tính bảng. Ví dụ, người sử dụng trong ví dụ trên có thể tìm kiếm hình ảnh liên quan đến đi bộ đường dài và leo núi ở New Zealand. Vì người sử dụng nhấp vào hình ảnh, trình duyệt không thể ghi lại siêu dữ liệu dựa trên các từ khóa được nhập vào trình duyệt. Tương tự, người sử dụng có thể xem video liên quan đến các điểm tham quan ở New Zealand. Với công nghệ thông thường, cả hình ảnh lẫn video sẽ không đóng góp siêu dữ liệu cho việc tiếp thị nhắm mục tiêu.

Với sự gia tăng và phổ biến của nền tảng phân phối video dựa trên internet có các kiến trúc khác nhau, việc tìm kiếm trực quan trên nội dung video có tiềm năng khai thác tới hàng triệu người dùng dựa trên việc bao gồm người tạo nội dung, người dùng và đối tác thương mại. Nếu các đoạn video cụ thể có thể được xác định, các bên quan tâm sẽ có khả năng tăng thêm và/hoặc tích hợp các phần này với nội dung bổ sung. Điều này có thể mang hình thức làm giàu thông tin cho các phần như vậy. Các đối tác thương mại có thể muốn nhắm mục tiêu các phần có liên quan làm tuyến để cung ứng các dịch vụ sản phẩm.

Hiện tại không có phương tiện tạo và trích xuất siêu dữ liệu hiệu quả dựa trên hình ảnh mà được xem bởi người sử dụng. Do đó, hình ảnh được xem/tìm kiếm không thể được bổ sung vào hồ sơ của người đó dùng cho việc tiếp thị nhắm mục tiêu. Ngoài ra, không có phương tiện tìm kiếm dựa trên hình ảnh được chụp mà không cần mô tả nó trong truy vấn văn bản. Vì vậy, không có cách nào để người xem thể hiện sự quan tâm đến đối tượng hoặc sản phẩm mà họ xem trong một chương trình hoặc cảnh. Ví dụ, người xem có thể nhìn thấy người nổi tiếng với túi xách. Tuy nhiên, có thể không rõ nơi mua túi xách. Lựa chọn duy nhất cho người xem là mô tả các thuộc tính của túi xách trong truy vấn tìm kiếm trực quan.

Các nỗ lực đã được thực hiện để liên kết video và/hoặc bản in tới trang web. Bằng cách sử dụng công nghệ hiện tại, nhà cung cấp hoặc nhà quảng cáo có thể kèm mã QR (Mã phản hồi nhanh-Quick Response Code) trên hoặc quảng cáo qua bản in hoặc video. Người xem có thể sử dụng điện thoại thông minh để quét mã QR mà hướng người đó đến trang web và/hoặc nội dung web. Tuy nhiên, điều này đòi hỏi khối mã rõ ràng được đặt ở gần người xem. Hơn nữa, mã QR riêng biệt phải được bao gồm cho từng đối tượng quan tâm. Đối với video, mã QR phải có mặt trong toàn bộ thời lượng của video.

Có những hạn chế rõ ràng với việc sử dụng mã QR. Hơn nữa, các nhà tiếp thị đang tìm kiếm các phương pháp cải tiến để thu hút các đối tượng trẻ hơn mà được sử dụng để sử dụng phương tiện bỏ qua quảng cáo và theo yêu cầu. Vị trí sản phẩm và giải trí có thương hiệu cung cấp các khả năng “bán hàng đa kênh” để thu hút người tiêu dùng trẻ hơn và/hoặc người sử dụng có hiểu biết về công nghệ nhiều hơn. Theo đó, cần có phương pháp cho phép người xem thể hiện sự quan tâm và/hoặc có thêm thông tin liên quan đến các đối tượng trên hình ảnh và/hoặc video. Hệ thống này cần cho phép người dùng có được thông tin chi tiết và thông tin thêm về đối tượng quan tâm mà không phải thực hiện việc tìm kiếm theo từ

khóa hoặc quét mã QR. Hệ thống này phải có khả năng sử dụng phương tiện dạng bản in (ví dụ, quảng cáo trên tạp chí) cũng như phương tiện video (ví dụ, truyền hình).

Bản chất kỹ thuật của sáng chế

Khía cạnh thứ nhất của sáng chế là phương tiện nhập hình ảnh từ phương tiện dạng bản in vào trong cơ sở dữ liệu.

Khía cạnh thứ hai của sáng chế là phương tiện nhập hình ảnh từ phương tiện dạng bản in vào trong cơ sở dữ liệu để các đối tượng trong các hình ảnh có thể được xác định và được ghép với các sản phẩm.

Khía cạnh thứ ba của sáng chế là phương tiện nhập hình ảnh từ phương tiện dạng bản in vào trong cơ sở dữ liệu trong đó phần văn bản và nội dung không liên quan được loại bỏ.

Khía cạnh thứ tư của sáng chế là phương tiện nhập các đoạn của phương tiện video vào trong cơ sở dữ liệu.

Khía cạnh thứ năm của sáng chế là phương tiện nhập các đoạn của video vào trong cơ sở dữ liệu trong đó các đoạn video được xác định bằng cách so sánh các cảnh giống nhau hoặc tương tự để các đoạn tương tự có thể được nhóm lại và các đoạn trùng lặp được xóa bỏ.

Khía cạnh thứ sáu của sáng chế là phương tiện nhập các đoạn của video vào trong cơ sở dữ liệu trong đó độ phân giải và/hoặc tốc độ khung hình của video được giảm xuống.

Khía cạnh thứ bảy của sáng chế là phương tiện nhập các đoạn của video vào trong cơ sở dữ liệu để các đối tượng trong các đoạn video có thể được xác định và phù hợp với các sản phẩm.

Giới thiệu

Sáng chế bao gồm phương pháp biên dịch các hình ảnh từ phương tiện dạng bản in vào trong cơ sở dữ liệu bao gồm các bước gồm (a) bước thu phương tiện dạng bản in, (b) bước chuyển đổi phương tiện dạng bản in thành phương tiện kỹ thuật số, (c) bước phát hiện văn bản trong phương tiện kỹ thuật số dựa trên các đường viền được nối của các mép, (e) bước loại bỏ các vùng có văn bản khỏi phương tiện kỹ thuật số, (f) bước phát hiện một hoặc nhiều hình ảnh trong phương tiện kỹ thuật số, (g) bước xác định liệu có đủ đặc điểm để phân loại một hoặc nhiều hình ảnh này, (h) bước phân loại một hoặc nhiều hình ảnh này và

(i) bước chèn một hoặc nhiều hình ảnh này vào trong cơ sở dữ liệu. Phương pháp này có thể bao gồm các bước bổ sung gồm bước truy cập cơ sở dữ liệu khi người sử dụng gửi truy vấn và bước khớp truy vấn với một hoặc nhiều hình ảnh trong cơ sở dữ liệu bằng cách so sánh các đặc điểm.

Sáng chế còn bao gồm phương pháp biên dịch các khung hình từ video, lưu và lập chỉ mục các khung hình trong cơ sở dữ liệu bao gồm các bước gồm (a) bước thu video, (b) bước phân tích video đối với các đặc điểm trong các khung hình, (c) bước phân tách video thành các đoạn dựa trên các đặc điểm, (d) bước phân tích các đặc điểm của các đoạn để nhóm các đoạn mà có chung các đặc điểm, (e) bước chú thích một hoặc nhiều khung hình của các đoạn với siêu dữ liệu và (f) bước lưu một hoặc nhiều khung hình của các đoạn trong cơ sở dữ liệu. Phương pháp có thể bao gồm bước bổ sung là làm giảm tỷ lệ khung hình và/hoặc độ phân giải của video. Có thể còn bao gồm các bước bổ sung gồm bước truy cập cơ sở dữ liệu khi người sử dụng gửi truy vấn và bước so khớp truy vấn với một hoặc nhiều khung hình của các đoạn trong cơ sở dữ liệu dựa trên các đặc điểm chung. Các khung hình trùng lặp có thể được xác định dựa trên nội dung chung để một hoặc nhiều khung hình trùng lặp có thể được loại bỏ.

Sáng chế còn bao gồm hệ thống máy tính để nhập các tài liệu bao gồm (a) giao diện người dùng để tải tài liệu lên máy chủ, (b) mô đun bao gồm logic để xử lý tài liệu và trích xuất các phần từ tài liệu như các hình ảnh, (c) mô đun bao gồm logic để loại bỏ văn bản khỏi các hình ảnh, (d) mô đun bao gồm logic để phát hiện các đặc điểm trong các hình ảnh và (e) mô đun bao gồm logic để gắn các ký hiệu nhận dạng duy nhất vào các hình ảnh. Hệ thống máy tính có thể còn bao gồm phương tiện để nhận truy vấn và mô đun để so khớp truy vấn với một hoặc nhiều hình ảnh trong cơ sở dữ liệu dựa trên các đặc điểm chung.

Hơn nữa, sáng chế bao gồm hệ thống máy tính để nhập video bao gồm (a) giao diện người dùng để tải video lên máy chủ, (b) chương trình máy tính để xử lý video và trích xuất các phần từ video như các hình ảnh và (c) cơ sở dữ liệu để lưu các phần được trích xuất. Các phần được trích xuất có thể được phân tích để tuân thủ các thông số phù hợp, bao gồm các đặc điểm hình ảnh và độ phân giải hình ảnh. Ký hiệu nhận dạng duy nhất có thể được gắn vào mỗi phần video được trích xuất. Hệ thống máy tính có thể còn bao gồm phương

tiện để nhận truy vấn và mô đun để so khớp truy vấn với một hoặc nhiều phần video trong cơ sở dữ liệu dựa trên các đặc điểm chung.

Mô tả vắn tắt các hình vẽ

Fig.1 mô tả tổng quan về khung phần mềm nhập hình ảnh làm ví dụ minh họa.

Fig.2 mô tả các bước được sử dụng bởi thuật toán cắt ảnh.

Fig.3 mô tả tổng quan về khung phần mềm nhập video làm ví dụ minh họa.

Fig.4 mô tả tổng quan về chương trình xử lý đoạn video.

Fig.5 mô tả chương trình nhập đoạn video.

Mô tả chi tiết sáng chế

Tham chiếu trong bản mô tả này đến "một phương án/khía cạnh" hoặc "phương án/khía cạnh" có nghĩa là tính năng, cấu trúc hoặc đặc điểm cụ thể được mô tả liên quan đến phương án/khía cạnh được bao gồm trong ít nhất một phương án/khía cạnh của sáng chế. Việc sử dụng cụm từ "theo một phương án/khía cạnh" hoặc "theo phương án/khía cạnh khác" ở các vị trí khác nhau trong bản mô tả không nhất thiết đều đề cập đến phương án/khía cạnh giống nhau, cũng không phải là các phương án/khía cạnh riêng biệt hoặc xen kẽ loại trừ lẫn nhau của các phương án/khía cạnh khác. Hơn nữa, các đặc điểm khác nhau được mô tả mà có thể được thể hiện bởi một số phương án/khía cạnh và không phải bởi các phương án/khía cạnh khác. Tương tự, các yêu cầu khác nhau được mô tả mà có thể là các yêu cầu cho một số phương án/khía cạnh nhưng không phải là phương án/khía cạnh khác. Trong một số trường hợp, phương án và khía cạnh có thể được sử dụng theo cách hoán đổi cho nhau.

Các thuật ngữ được sử dụng trong bản mô tả này thường có ý nghĩa thông thường của chúng trong lĩnh vực kỹ thuật, trong hoàn cảnh của sáng chế, và trong hoàn cảnh cụ thể mà mỗi thuật ngữ được sử dụng. Một số thuật ngữ được sử dụng để mô tả sáng chế được nêu dưới đây, hoặc ở phần khác trong bản mô tả, để cung cấp chỉ dẫn bổ sung cho người thực hiện về phần mô tả của sáng chế. Để thuận tiện, một số thuật ngữ có thể được đánh dấu, ví dụ như sử dụng chữ in nghiêng và/hoặc các dấu ngoặc kép. Việc sử dụng cách làm nổi bật không ảnh hưởng đến phạm vi và ý nghĩa của thuật ngữ; trong cùng một hoàn cảnh, phạm vi và ý nghĩa của thuật ngữ là giống nhau cho dù thuật ngữ này có được đánh dấu hay không. Cần phải hiểu là việc giống nhau có thể được nêu theo nhiều hơn một cách.

Do đó, ngôn ngữ thay thế và từ đồng nghĩa có thể được sử dụng cho một hoặc nhiều thuật ngữ được nêu trong tài liệu này. Cũng không có bất kỳ ý nghĩa đặc biệt nào được đặt ra khi có hoặc không thuật ngữ được xây dựng hoặc được nêu ở đây. Các từ đồng nghĩa với một số thuật ngữ được cung cấp. Việc nêu một hoặc nhiều từ đồng nghĩa không loại trừ việc sử dụng các từ đồng nghĩa khác. Việc sử dụng các ví dụ ở vị trí bất kỳ trong bản mô tả này bao gồm các ví dụ về các thuật ngữ bất kỳ được nêu ở đây chỉ mang tính minh họa và không nhằm hạn chế sâu hơn phạm vi và ý nghĩa của sáng chế hoặc thuật ngữ được lấy làm ví dụ bất kỳ. Tương tự như vậy, sáng chế không giới hạn ở các phương án khác nhau được đưa ra trong bản mô tả này.

Không nhằm làm giới hạn thêm phạm vi của sáng chế, ví dụ về các công cụ, thiết bị, phương pháp và kết quả liên quan của chúng theo các phương án của sáng chế được đưa ra dưới đây. Lưu ý rằng tiêu đề hoặc tiêu đề phụ có thể được sử dụng trong các ví dụ để thuận tiện cho người đọc, mà không theo bất kỳ cách nào để giới hạn phạm vi của sáng chế. Trừ khi được định nghĩa khác, tất cả các thuật ngữ kỹ thuật và khoa học được sử dụng ở đây có cùng nghĩa như thường được hiểu bởi người có trình độ trung bình trong lĩnh vực kỹ thuật mà sáng chế liên quan đến. Trong trường hợp xung đột, tài liệu hiện tại, bao gồm các định nghĩa, sẽ được ưu tiên.

Thuật ngữ "ứng dụng" đề cập đến chương trình độc lập hoặc một phần của phần mềm được thiết kế để đáp ứng mục đích cụ thể, đặc biệt là khi được tải xuống trên thiết bị di động.

Thuật ngữ "cookie", "cookie internet" hoặc "cookie HTTP" đề cập đến đoạn dữ liệu nhỏ được gửi từ trang web và được lưu trữ trên máy tính của người dùng bởi trình duyệt web của người dùng. Cookie được gửi qua lại giữa máy chủ internet và trình duyệt cho phép người dùng xác định hoặc theo dõi quá trình sử dụng. Cookie cung cấp chi tiết về những trang mà người dùng truy cập, lượng thời gian dành cho mỗi trang, các liên kết được nhấp, hoạt động tìm kiếm và tương tác đã thực hiện. Từ thông tin này, nhà phát hành cookie tập hợp sự hiểu biết về xu hướng và sở thích duyệt web của người dùng để tạo ra hồ sơ. Bằng việc phân tích hồ sơ, nhà quảng cáo có thể tạo các đoạn về đối tượng người dùng được xác định dựa trên các hoạt động sử dụng với thông tin tương tự được trả về tương tự, từ đó, tạo hồ sơ.

Thuật ngữ "phân cụm" hoặc "phân tích cụm" đề cập đến nhiệm vụ nhóm một tập hợp các đối tượng theo cách mà các đối tượng trong cùng một nhóm (được gọi là cụm) giống nhau hơn (theo nghĩa nào đó hoặc khác) với nhau hơn là những người trong các nhóm (cụm) khác. Đó là nhiệm vụ chính của khai thác dữ liệu thăm dò, và kỹ thuật phổ biến để phân tích dữ liệu thống kê, được sử dụng trong nhiều lĩnh vực, bao gồm học máy, nhận dạng mẫu, phân tích hình ảnh, truy xuất thông tin, thông tin sinh học, nén dữ liệu và đồ họa máy tính.

Thuật ngữ "học sâu" đề cập đến ứng dụng để học các nhiệm vụ của mạng nơ-ron nhân tạo (ANN) mà chứa nhiều hơn một lớp ẩn. Học sâu là một phần của hệ lớn hơn gồm các phương pháp học máy dựa trên việc học các sự biểu diễn dữ liệu, trái ngược với các thuật toán nhiệm vụ cụ thể.

Thuật ngữ "vector đặc trưng", trong nhận dạng mẫu và học máy, đề cập đến vector đặc trưng là vector n-chiều gồm các chức năng số mà biểu diễn một số đối tượng. Nhiều thuật toán trong học máy đòi hỏi sự biểu diễn số của các đối tượng, do các biểu diễn này tạo thuận lợi cho việc xử lý và phân tích thống kê. Khi biểu diễn hình ảnh, các trị số đặc điểm có thể tương ứng với các điểm ảnh của hình ảnh, khi biểu diễn các văn bản có lẽ là các tần số xuất hiện của thuật ngữ.

Thuật ngữ "chỉ số đảo ngược", "tệp đăng" hoặc "tệp ngược" là cấu trúc dữ liệu chỉ mục mà lưu trữ ánh xạ từ nội dung, chẳng hạn như từ hoặc số, đến vị trí của nó trong tệp cơ sở dữ liệu hoặc trong tài liệu hoặc bộ tài liệu (được đặt tên trái ngược với Chỉ số chuyển tiếp, mà ánh xạ từ các tài liệu đến nội dung). Mục đích của chỉ số đảo ngược là cho phép tìm kiếm văn bản đầy đủ nhanh, với chi phí xử lý gia tăng khi tài liệu được thêm vào cơ sở dữ liệu.

Thuật ngữ "vùng lân cận gần nhất" hoặc "k-NN" đề cập đến đối tượng phân loại lân cận gần nhất, trong đó cả số liệu khoảng cách ("gần nhất") và số các lân cận có thể được thay đổi. Đối tượng phân loại các quan sát mới bằng cách sử dụng phương pháp dự đoán. Đối tượng chứa dữ liệu được sử dụng để đào tạo, vì vậy có thể tính toán các dự đoán tái thay thế.

Phân tích liên kết

Thuật ngữ “mô-đun” đề cập đến bộ phận độc lập, như cụm linh kiện điện tử và hệ thống dây điện kết hợp hoặc một phần của phần mềm máy tính, mà tự thực hiện nhiệm vụ được xác định và có thể được liên kết với các bộ phận khác để tạo thành hệ thống lớn hơn.

Thuật ngữ "mạng lưới nơ-ron nhận thức nhiều lớp" hoặc "MLP" dùng để chỉ mạng lưới nơ-ron truyền thẳng với một hoặc nhiều lớp giữa các lớp đầu vào và đầu ra. Truyền thẳng có nghĩa là các luồng dữ liệu đi theo một hướng từ lớp đầu vào đến lớp đầu ra (chuyển tiếp). MLP được sử dụng rộng rãi để phân loại mẫu, nhận dạng, dự đoán và xấp xỉ. Perceptron nhiều lớp có thể giải quyết các vấn đề không thể tách rời tuyến tính.

Thuật ngữ "siêu dữ liệu" đề cập đến dữ liệu mô tả dữ liệu khác. Nó cung cấp thông tin về nội dung của mục nhất định. Hình ảnh có thể bao gồm siêu dữ liệu mô tả độ lớn hình ảnh, độ sâu màu, độ phân giải hình ảnh và khi hình ảnh được tạo. Siêu dữ liệu của tài liệu văn bản có thể chứa thông tin về tài liệu là bao lâu, tác giả là ai, khi nào tài liệu được viết và bản tóm tắt ngắn của tài liệu.

Thuật ngữ "thẻ meta" đề cập đến siêu dữ liệu được bao gồm trên các trang web. Thẻ meta mô tả và từ khóa thường được sử dụng để mô tả nội dung của trang Web. Hầu hết các công cụ tìm kiếm đều sử dụng dữ liệu này khi thêm trang vào chỉ mục tìm kiếm của chúng.

Thuật ngữ "mã QR" hoặc "mã phản hồi nhanh" đề cập đến mã vạch ma trận (hoặc mã vạch hai chiều) có chứa thông tin về mục được đính kèm. Mã QR bao gồm các hình vuông màu đen được sắp xếp trong lưới ô vuông trên nền trắng, mà có thể đọc bằng thiết bị hình ảnh như máy ảnh và được xử lý bằng cách sử dụng kỹ thuật sửa lỗi Reed-Solomon cho đến khi hình ảnh có thể được thông dịch một cách thích hợp. Dữ liệu cần thiết sau đó được trích xuất từ các mẫu có mặt ở cả thành phần ngang và dọc của hình ảnh.

Thuật ngữ "dữ liệu tổng hợp" đề cập đến dữ liệu sản xuất bất kỳ áp dụng cho tình huống cho trước mà không thu được bằng phép đo trực tiếp.

Thuật ngữ “máy vector hỗ trợ” hoặc “SVM” đề cập đến các mô hình học có giám sát có các thuật toán học tập được kết hợp mà phân tích dữ liệu được sử dụng để phân loại và phân tích hồi quy. Với tập hợp các ví dụ đào tạo, mỗi ví dụ được đánh dấu thuộc về một hoặc hai danh mục khác, thuật toán đào tạo SVM xây dựng mô hình gán các ví dụ mới cho một danh mục hoặc loại khác, làm cho nó trở thành một trình phân loại tuyến tính nhị phân không xác suất.

Thuật ngữ "quảng cáo nhắm mục tiêu" đề cập đến hình thức quảng cáo mà nhà quảng cáo trực tuyến có thể sử dụng các phương pháp phức tạp để nhắm đối tượng tiếp nhận nhiều nhất với các đặc điểm nhất định, dựa trên sản phẩm hoặc người mà nhà quảng cáo đang quảng bá. Các đặc điểm này có thể là đối tượng khách hàng mà tập trung vào chủng tộc, tình trạng kinh tế, giới tính, tuổi tác, trình độ học vấn, mức thu nhập và việc làm hoặc họ có thể tập trung vào tâm lý dựa trên giá trị, cá tính, thái độ, quan điểm, lối sống và sở thích. Chúng cũng có thể là các biến hành vi, chẳng hạn như lịch sử trình duyệt, lịch sử mua hàng và hoạt động gần đây khác.

Các thuật ngữ kỹ thuật khác được sử dụng ở đây có ý nghĩa thông thường trong lĩnh vực kỹ thuật mà chúng được sử dụng, như được thể hiện bởi nhiều từ điển kỹ thuật.

Các phương án ưu tiên

Các trị số và kết cấu cụ thể được bàn luận trong các ví dụ không làm giới hạn này có thể khác nhau và được trích dẫn chỉ để minh họa ít nhất một phương án và không nhằm giới hạn phạm vi của chúng.

Sáng chế đề cập đến việc đồng hóa nội dung video vào trong nền tảng tìm kiếm trực quan sao cho mỗi khung có thể phân biệt duy nhất có thể tìm kiếm được bằng cách sử dụng hình ảnh truy vấn của đoạn video có liên quan từ hệ thống cảm biến máy tính. Trong đó bao gồm phương tiện nhập hình ảnh và các khung hình video vào cơ sở dữ liệu. Các sản phẩm tiêu dùng có thể được kết hợp với hình ảnh và/hoặc đối tượng trong hình ảnh. Người dùng có thể truy cập cơ sở dữ liệu bằng cách gửi hình ảnh kỹ thuật số được chụp trên thiết bị di động. Một hoặc nhiều đối tượng trong hình ảnh do người dùng gửi có thể được so khớp với các sản phẩm và/hoặc tài liệu thương mại/quảng cáo.

Một trong các lợi ích của tìm kiếm trực quan (trái ngược với tìm kiếm dựa trên văn bản truyền thống) đối với việc lập hồ sơ đối tượng khách hàng là thông tin vốn đã lớn hơn về truy vấn có thể được xác định chắc chắn. Ví dụ, người dùng có thể tìm kiếm giày màu nâu trong công cụ tìm kiếm (hoặc trang web thương mại điện tử). Trong trường hợp sử dụng tìm kiếm trực quan, hình ảnh truy vấn có thể tự bộc lộ nhiều hơn về bản chất của truy vấn của người dùng. Người dùng có thể tìm hiểu về loại giày nâu rất cụ thể (hình dạng, kiểu dáng, chất liệu, thương hiệu, giày đế, có ren, vv...) bằng cách gửi hoặc nhấp vào hình ảnh

của chiếc giày cụ thể. Với việc truy cập chỉ dựa vào truy vấn tìm kiếm văn bản, thì không thể trích xuất chi tiết hơn về đối tượng tìm kiếm mà không cần thêm bất kỳ thông tin nào.

Tìm kiếm trực quan dưới dạng hình thức tương tác của người dùng đã đạt được rất nhiều sức hút gần đây do những tiến bộ trong lĩnh vực thị giác máy. Bây giờ có thể truy vấn cơ sở dữ liệu chứa hàng triệu hình ảnh với độ chính xác cao. Điều này mở ra khả năng tương tác nội dung con người. Theo một phương án này, có thể làm giàu các tài sản truyền thông tĩnh bằng cách gắn thẻ chúng với các trường hợp sử dụng thực tế được tăng cường. Để truy cập vào nội dung phong phú hơn, người dùng có thể trở thiết bị camera ở nội dung tĩnh và sử dụng công cụ kết hợp trực quan, truy xuất và hiển thị lớp phủ nội dung.

Các đoạn cụ thể của video có thể được xác định để các bên quan tâm có được khả năng tăng thêm và/hoặc kết hợp các phần này với nội dung bổ sung. Điều này có thể mang hình thức làm giàu thông tin cho các phần như vậy. Các đối tác thương mại có thể nhắm mục tiêu các phần có liên quan làm các kênh để phân phối các dịch vụ sản phẩm.

Để trích xuất siêu dữ liệu về truy vấn tìm kiếm trực quan, các thuật toán phân loại nâng cao, bao gồm nhưng không giới hạn ở học sâu, học có giám sát và học không giám sát có thể được sử dụng. Do đó, từ hình ảnh đầu vào, có thể thu được danh sách siêu dữ liệu mô tả (ví dụ, giày, nâu, ren, giày chuyên đi núi, thuộc tính, vị trí sản xuất, tài liệu và thông tin như vậy bất kỳ mà cung cấp rõ ràng về tình trạng của nội dung bên trong hình ảnh).

Theo một phương án của sáng chế, danh sách các đối tượng mà gồm hình ảnh hoặc video có thể được trích xuất và liên kết với trình tự các khung hình mà được phân tích tương ứng với "chủ đề" khác biệt về ngữ nghĩa. Ví dụ, dữ liệu từ hình ảnh in (ví dụ, tạp chí) có thể được nhập vào hệ thống. Người dùng có thể gửi hình ảnh kỹ thuật số của chiếc giày từ trang của tạp chí. Hệ thống có thể chuyển tiếp thông tin sản phẩm về chủ đề giày. Theo phương án khác, dữ liệu từ video (ví dụ, truyền hình) được nhập vào hệ thống. Người dùng có thể gửi ảnh chụp màn hình của video trong đó giày là một trong các mục trên màn hình. Hệ thống có thể chuyển tiếp thông tin sản phẩm đến người dùng về chủ đề giày.

Để xây dựng và lập chỉ mục nội dung trong hệ thống, các tài liệu kỹ thuật số trực quan phong phú, chẳng hạn như tạp chí, truyện tranh hoặc các tài liệu khác mà chuyển tải thông tin theo cách chủ yếu là phi văn bản (ví dụ, tài liệu bao gồm hình ảnh) có thể được sử dụng như được mô tả ở dưới.

Sự nhập tài liệu

Giao diện người dùng nhập tài liệu cung cấp công cụ để tải lên tài liệu chính vào cơ sở dữ liệu nội dung. Giao diện này có thể cung cấp phương tiện chú thích và tập hợp siêu dữ liệu. Tài liệu được tải lên có thể được lập chỉ mục và được lưu trữ trong cơ sở dữ liệu nội dung 180. Siêu dữ liệu có thể được gán cho mỗi tài liệu.

Các bước trong quá trình nhập dữ liệu 100 được nêu chi tiết trên Fig.1. Khung phần mềm nhập tài liệu cho phép sự xác định các phân đoạn cụ thể của tài liệu như tạp chí hoặc báo hàng ngày bằng cách sử dụng hình ảnh truy vấn trực quan để thúc đẩy nhận dạng này.

Tài liệu (ví dụ, các trang được quét từ tạp chí) 115 có thể được tải lên hoặc “được nhập” bằng cách sử dụng giao diện người dùng 120. Ví dụ, các trang của tạp chí có thể được quét và tải lên ở định dạng tài liệu di động (PDF) hoặc định dạng khác. Tài liệu được chuyển đến máy chủ 130 và các phần (ví dụ, hình ảnh riêng lẻ) được trích xuất 135. Sau đó hệ thống sẽ xác định liệu tài liệu có hợp lệ hay không 145. Ví dụ, tài liệu không có hình ảnh rõ ràng có thể không được coi là hợp lệ. Thuật toán có thể được sử dụng để xác định xem hình ảnh có phù hợp với tiêu chí trong dung sai nhất định hay không. Nếu nó không hợp lệ, hệ thống có thể tạo ra phản hồi thất bại 140. Phản hồi nhập có thể được gửi đến giao diện người dùng 125 để gửi cảnh báo hoặc thông báo.

Nếu tài liệu hợp lệ, hệ thống có thể chia tài liệu thành các phần 150. Ví dụ, các phần tương tự có thể được nhóm lại với nhau. Công cụ chú thích có thể được sử dụng để cho phép phân đoạn tài liệu. Tạp chí có thể được chia thành các bài báo, quảng cáo và khuyến mãi. Nếu không có thông tin phân đoạn nào được tạo ra, toàn bộ tài liệu có thể được xử lý như một thực thể duy nhất. Sau khi phân đoạn, mỗi phân đoạn có thể được xử lý để nhập vào cơ sở dữ liệu nội dung.

Vùng văn bản có thể được phát hiện và được loại bỏ 155. Hình ảnh có thể được cắt để tách chúng khỏi các vùng trống và/hoặc đường viền trên các trang 160. Mỗi phần có thể được phân tích để đảm bảo sự phù hợp của nó 165. Ví dụ, hệ thống có thể phát hiện và xác định các tính năng. Hình ảnh có độ phân giải thấp, mờ hoặc trừu tượng (hoặc hình ảnh không có đặc điểm nhận dạng) có thể được coi là không phù hợp. Nếu có đủ đặc điểm 170, có thể được chèn vào trong cơ sở dữ liệu 175.

Siêu dữ liệu có thể được liên kết với mỗi hình ảnh được xử lý. Các trường siêu dữ liệu có thể bao gồm nội dung được liên kết, ngôn ngữ tài liệu cũng như thông tin tác giả/nhà xuất bản của tài liệu.

Chèn cơ sở dữ liệu

Việc lưu trữ từng đoạn vào cơ sở dữ liệu hình ảnh/nội dung 180 có thể bao gồm các bước sau:

- 1) Trích xuất các đặc điểm trực quan từ hình ảnh phân đoạn;
- 2) Tạo ra phản hồi chèn bao gồm:
 - Vectơ của các đặc điểm trực quan
 - ID phân đoạn duy nhất
 - ID nội dung được liên kết
- 3) Đóng gói phản hồi chèn vào định dạng vận chuyển thích hợp (ví dụ: JSON);
- 4) Gửi phản hồi chèn đến cơ sở dữ liệu;
- 5) Nhận trạng thái chèn; và
- 6) Truyền phản hồi trạng thái chèn đến giao diện người dùng.

Trích xuất phần tài liệu

Khi tài liệu đã được tải lên máy chủ, chương trình máy tính có thể xử lý tài liệu này và phân vùng nó thành một loạt các phân đoạn.

Ví dụ, bản sao kỹ thuật số của tạp chí có thể là tài liệu được tải lên ban đầu dưới dạng PDF. Các phân đoạn có thể được chỉ định là từng trang của tệp PDF. Giai đoạn trích xuất cũng có thể đánh giá sự phù hợp của từng phần của tài liệu để xử lý tiếp. Điều này có thể có dạng kiểm tra tính toàn vẹn để đánh giá xem tệp có ở định dạng thích hợp để nhập hay không, hoặc liệu mỗi phần nằm trong các tham số an toàn (chẳng hạn như độ phân giải tối thiểu).

Phân tích tài liệu

Đối với mỗi phân đoạn được trích xuất từ tài liệu được tải lên ban đầu, khung phần mềm là cần thiết để xử lý và xác thực chất lượng của hình ảnh cần được nhập vào hệ thống thị giác. Khung phần mềm này bao gồm việc xóa văn bản dựa trên các đường viền mép được kết nối.

Một vấn đề với việc nhập hình ảnh từ tài liệu là xu hướng của các đặc điểm văn bản hoặc dạng văn bản được nhúng trong hình ảnh. Trong nhiều hệ thống tìm kiếm trực quan, điều này có thể gây ra sự cố trong việc đảm bảo tính duy nhất trong các hình ảnh có nhiều chữ. Theo đó, điều quan trọng là phát hiện, cô lập và loại bỏ các vùng giàu văn bản ra khỏi hình ảnh tổng hợp mà bao gồm hình ảnh và văn bản.

Fig.2 mô tả các bước trong hàm của thuật toán phát hiện và cắt văn bản có thể được sử dụng để tách biệt các vùng của hình ảnh phân đoạn có chứa văn bản. Hình ảnh có thể được cắt để các vùng văn bản gần tối ưu được xóa trước khi chèn. Hình ảnh thông thường 185 từ tạp chí hoặc báo hàng ngày sẽ chứa các đoạn văn bản. Máy tính có thể xác định các phần văn bản 190 dựa trên, ví dụ như nhận dạng ký tự và sắp xếp thứ tự các chữ cái thành các dòng và đoạn (được mô tả bên dưới). Các khu vực nhiều văn bản được loại bỏ để lại hình ảnh tách biệt 195. Sau đó, hình ảnh có thể được cắt 205 để loại bỏ các vùng thừa.

Văn bản cũng có thể được xác định dựa trên các đường viền mép được kết nối. Văn bản được chuẩn bị để in có các cạnh rất rõ nét, thường được làm rõ nét hơn hình ảnh thông thường. Các cạnh mức độ cao có thể được bảo toàn và được chuẩn bị cho việc xử lý đường nét. Sự phát hiện các đường nét sẽ tách biệt các từ, dòng và văn bản gần đó. Nếu các đường viền chứa nhiều dòng hơn, chúng có thể được phân tách thành các dòng đơn, và sau đó, việc lọc được thực hiện. Các dòng và từ đơn lẻ được kết nối và dán nhãn thành các đoạn văn hoặc các dòng đơn lớn. Các nguyên tắc sau đây có thể được áp dụng cho quá trình xác định văn bản.

Các cạnh có thể được phát hiện trên ảnh màu xám bằng cách sử dụng phép toán gradient hình thái học.

Sự nhị phân có thể đạt được bằng cách sử dụng việc tạo ngưỡng của hình ảnh có mép được phát hiện bằng cách sử dụng phương pháp Otsu (tạo ngưỡng hình ảnh dựa trên việc phân cụm).

Việc đóng hình thái học có thể đạt được để kết nối các đối tượng nhỏ (chữ cái hoặc từ).

Các viền có thể được phát hiện bằng cách sử dụng việc phân tích các thành phần được kết nối.

Bước lọc đầu tiên có thể đạt được bằng cách sử dụng chiều rộng, chiều cao và tỷ lệ các điểm ảnh khác 0 trong vùng đường bao.

Phần nhô dọc của vùng đường viền có thể được sử dụng để phát hiện nếu đường viền được tìm thấy chứa nhiều hơn một dòng. Trong trường hợp có nhiều dòng, ngưỡng phần nhô dọc có thể được sử dụng để xác định nơi để tách các dòng.

Các vùng đường viền có thể được xác thực và được lọc bằng cách sử dụng khu vực, kích thước, tỷ lệ các điểm khác 0 trong vùng đường viền và một số đặc điểm được thực hiện tùy chỉnh diện tích các lần tỷ lệ khung hình tương đối (RARA) và kích cỡ cạnh vùng trung bình các lần tỷ lệ khung hình tương đối (RARAS).

Việc kết nối và dán nhãn các từ và dòng thành các đoạn văn có thể được thực hiện trên các hình chữ nhật mà biểu diễn các hộp giới hạn của vùng đường bao. Các hình chữ nhật đầu tiên được liên kết trái và phải, để các dòng được xác định. Sau đó, hình chữ nhật không thể xóa được. Cuối cùng, liên kết lên và xuống có thể được thực hiện. Kết quả cuối cùng bao gồm các hình chữ nhật lớn được xác định có chứa các vùng được coi là chứa đoạn văn bản.

Sự nhập video

Tương tự, nội dung video có thể được biên dịch và được nhập vào cơ sở dữ liệu nội dung. Phần này phác thảo hệ thống cho phép người dùng tải nội dung video lên chương trình máy tính được thiết kế để xử lý trước video thành định dạng phù hợp cho việc phân tích và đồng hóa sâu hơn vào trong nền tảng tìm kiếm trực quan. Giao diện người dùng (UI) cho phép lựa chọn nguồn video mà người dùng muốn đưa vào nền tảng. Nguồn có thể là tệp có thể truy cập từ xa hoặc cục bộ từ thiết bị máy tính của người dùng.

Giao diện cũng có thể cho phép chú thích các đoạn tệp video với siêu dữ liệu có thể tùy chỉnh được liên kết với đoạn. Giao diện cũng có thể cung cấp tổng quan trạng thái của từng phân đoạn được xác định trong video và chỉ số liên quan đến tất cả phân tích và kiểm tra được thực hiện trên phân đoạn, từ đó cung cấp phản hồi ngay lập tức về tất cả trạng thái giai đoạn cho phân đoạn đó. Fig.3 mô tả cách giao diện này có thể tạo thuận lợi cho việc vận chuyển và chú thích video và các đoạn video đến nền tảng tìm kiếm trực quan.

Các bước gồm việc nhập video 300 được nêu chi tiết trên Fig.3. Khung phần mềm nhập video cho phép xác định các phân đoạn cụ thể của video bằng cách sử dụng hình ảnh truy vấn trực quan.

Nội dung video (ví dụ, các chương trình truyền hình) 215 có thể được tải lên hoặc "được nhập" bằng cách sử dụng giao diện người dùng 220. Video được chuyển đến máy chủ 230 và các phần được phân tích 235. Video có thể được phân đoạn thêm 245. Ví dụ, các đối tượng khác nhau trong cảnh có thể chỉ ra sự thay đổi về thiết đặt. Sau đó, hệ thống sẽ xác định xem đoạn video có hợp lệ hay không 250. Ví dụ, video không có các đối tượng có thể nhận dạng có thể được coi là không hợp lệ. Thuật toán có thể được sử dụng để xác định xem video có phù hợp với tiêu chí trong dung sai nhất định hay không. Nếu không hợp lệ, hệ thống có thể tạo ra phản hồi thất bại 240. Phản hồi nhập có thể được gửi đến giao diện người dùng 225 để thông báo người dùng.

Nếu đoạn video hợp lệ, hệ thống có thể chia video thành các phân đoạn 245. Ví dụ, các phần tương tự (dựa trên nội dung cảnh) có thể được nhóm lại với nhau. Các khung riêng lẻ có thể được trích xuất 255. Các đặc điểm trực quan có thể được tạo ra cho mỗi khung hình 260. Mỗi khung có thể được phân tích để đảm bảo sự phù hợp của nó 265. Ví dụ, hệ thống có thể cố gắng phát hiện và xác định các đối tượng và đặc điểm trong mỗi khung hình. Khung có độ phân giải thấp hoặc bị mờ có thể được coi là không phù hợp. Nếu có đủ các đặc điểm 270, khung có thể được chèn vào cơ sở dữ liệu nội dung (275, 180).

Phân tích đoạn

Phân tích đoạn video là thành phần của khung phần mềm nhập video cho phép người dùng tách biệt và cung cấp siêu dữ liệu cụ thể theo phân đoạn cho các phân đoạn đã xác định.

Fig.4 mô tả cách khung phần mềm phân tích phân đoạn 350 có thể hoạt động. Video tải lên 110 được chia thành các phân đoạn. Các đoạn 310 có thể được xác định và sau đó được nhóm. Các phân đoạn có tính duy nhất nội phân khúc có thể được kết hợp hoặc nhóm lại với nhau 330. Ví dụ, các đoạn video từ việc thiết đặt tương tự với các đối tượng tương tự có khả năng là từ tình huống tương tự trong video. Hơn nữa, các đoạn lặp lại có thể được xác định 180. Khi phân đoạn trùng lặp được xác định, nó có thể được loại bỏ/xóa.

Theo phương án ưu tiên, video được xử lý theo cách toàn diện để xác định những điều sau đây:

Phân đoạn hoặc cảnh có thể phân biệt trong video -> Điều này xử lý video để phát hiện các nhóm khung hình mà mỗi chúng tương ứng với sự kiện logic duy nhất.

Các đoạn trùng lặp hiển thị trên video -> Việc này so sánh tất cả các đoạn được phát hiện từ quá trình trên để xác định cặp hoặc nhóm phân đoạn bất kỳ có thể không thể phân tách trực quan với nhau (tức là có khả năng gây ra sự mơ hồ khi tìm kiếm hình ảnh).

Các đoạn trùng lặp đã có trong cơ sở dữ liệu nội dung của nền tảng tìm kiếm trực quan -> Việc kiểm tra này đánh giá liệu phân đoạn hoặc khung hình trong đoạn có thể đã có trong cơ sở dữ liệu nội dung tìm kiếm trực quan.

Sau khi phân đoạn video thành công, mỗi đoạn có thể được chuyển đổi thêm thành định dạng phù hợp cho việc nhập vào cơ sở dữ liệu nội dung 180.

Trích xuất đặc điểm đoạn

Khi xác định và tách biệt các đoạn video hợp lệ, cần phải đóng gói từng khung hình trong đoạn vào bản ghi chèn.

Điều này đòi hỏi mô-đun được thực hiện bởi chương trình máy tính hoạt động trên hệ thống để xử lý từng khung hình trong đoạn và trích xuất dữ liệu biểu diễn cho quy mô và các đặc điểm trực quan bất biến quay thành vector (chẳng hạn như ORB). Hơn nữa, mỗi vector đặc trưng có thể được kết hợp với nhiều siêu dữ liệu và các ký hiệu nhận dạng khác nhau để tạo ra cấu trúc dữ liệu tổng hợp để chèn vào cơ sở dữ liệu nội dung.

Phương án mẫu về cấu trúc dữ liệu tổng hợp như vậy có thể bao gồm như sau:

- Ký hiệu nhận dạng khung hình duy nhất
- Ký hiệu nhận dạng đoạn duy nhất
- Các đặc điểm trực quan khung hình
- Siêu dữ liệu (nhà xuất bản, thể mô tả, v.v.)

Chèn cơ sở dữ liệu

Khi bản ghi chèn khung hình đã được tạo ra, đối tượng sẵn sàng để chèn vào cơ sở dữ liệu nội dung. Trước khi chèn bản ghi khung vào cơ sở dữ liệu nội dung, cơ sở dữ liệu nội dung được truy vấn bằng cách sử dụng ID nội dung được cung cấp. Nếu nội dung không tồn tại, có thể có quy trình xử lý để cho phép bổ sung nội dung khi phù hợp.

Bản ghi chèn sau đó có thể được vận chuyển đến cơ sở dữ liệu để duy trì. Sau khi duy trì thành công, cơ sở dữ liệu sẽ cung cấp sự xác nhận yêu cầu cập nhật này và trả về phản hồi cho chương trình máy tính có thể chịu trách nhiệm để chuyển bản ghi chèn. Bản cập nhật trạng thái này sau đó có thể được truyền thông qua khung phần mềm, chỉ ra sự hoàn thành của chuỗi quy trình của khung hình cụ thể đó.

Fig.5 cho thấy phương án làm ví dụ về cách khung hình có thể được xử lý và được chèn vào cơ sở dữ liệu nội dung 180. Đoạn video được xử lý 410 được phân tích để trích xuất các đặc điểm trực quan 415 và tạo ra bản ghi nhập đoạn 435.

Hệ thống có thể phân tích cú pháp siêu dữ liệu 420 và xác định xem nội dung được liên kết có tồn tại không 440. Ví dụ, nội dung tương tự có thể đã có trong cơ sở dữ liệu. Nếu có, hệ thống có thể xác định ID nội dung được liên kết 450 để tạo bản ghi nhập đoạn 435 và chèn khung vào trong cơ sở dữ liệu nội dung 180. Nếu không có nội dung liên kết tồn tại, hệ thống có thể tạo nội dung 445 để chèn vào cơ sở dữ liệu nội dung 180. Cách tiếp cận này có thể được sử dụng cho nội dung video (ví dụ, các khung được xử lý) cũng như nội dung hình ảnh (tức là hình ảnh được quét).

Các trường hợp sử dụng

Sự thêm vào nội dung

Sáng chế có thể được sử dụng để làm giàu và/hoặc tăng cường phương tiện truyền thông dưới dạng video. Ngoài ra, tùy thuộc vào phân tích và phân lập các phân đoạn có thể phân biệt trong video, có thể liên kết từng phân đoạn với tập hợp nội dung khác.

Ví dụ, người tạo ra nội dung có thể tải nội dung video lên nền tảng tìm kiếm trực quan bằng cách sử dụng khung phần mềm. Nhiều đoạn được xác định từ việc phân tích video:

Trình tự tiêu đề đưa vào. Điều này có thể được liên kết với URL trang chủ của người tạo hoặc các vị trí khác.

Nhiều đoạn đưa vào từng thành phần trong nhóm: Mỗi đoạn video của từng thành phần có thể được liên kết với URL của các trang sinh học tương ứng của họ.

Đoạn mô tả sản phẩm. Đoạn này có thể được ánh xạ tới lớp phủ thực tế tăng cường, mà cung cấp cơ cấu để tương tác thực sự với sản phẩm.

Đoạn kết luận về lợi thế và bất lợi của sản phẩm. Đoạn này có thể được ánh xạ tới danh sách các nhà bán lẻ trực tuyến mà sản phẩm có thể được mua từ đó.

Theo cách này, nội dung video hiện tại có thể được tăng cường và làm giàu bằng cách cung cấp nền tảng cho sự tương quan với nội dung bổ sung dựa trên tìm kiếm trực quan.

Môi trường hoạt động:

Hệ thống này thường bao gồm máy chủ trung tâm được kết nối bởi mạng dữ liệu đến máy tính của người dùng. Máy chủ trung tâm có thể bao gồm một hoặc nhiều máy tính được kết nối với một hoặc nhiều thiết bị lưu trữ khối. Kiến trúc chính xác của máy chủ trung tâm không giới hạn sáng chế được yêu cầu bảo hộ. Hơn nữa, máy tính của người dùng có thể là loại máy tính xách tay hoặc máy tính để bàn. Cũng có thể là điện thoại di động, điện thoại thông minh hoặc thiết bị cầm tay khác, bao gồm cả máy tính bảng. Yếu tố hình thức chính xác của máy tính của người dùng không giới hạn sáng chế được yêu cầu bảo hộ. Ví dụ về các hệ thống máy tính, môi trường và/hoặc cấu hình nổi tiếng có thể phù hợp để sử dụng với sáng chế bao gồm, nhưng không giới hạn ở, máy tính cá nhân, máy tính chủ, máy cầm tay, máy tính xách tay hoặc máy tính di động hoặc các thiết bị liên lạc như điện thoại di động và PDA, các hệ thống đa xử lý, hệ thống trên cơ sở vi xử lý, bộ giải mã set-top-box, thiết bị điện tử tiêu dùng có thể lập trình, máy tính mạng, máy tính mini, máy tính lớn, các môi trường điện toán phân tán mà bao gồm hệ thống bất kỳ trong số các hệ thống hoặc các thiết bị ở trên, và tương tự. Yếu tố hình thức chính xác của máy tính của người dùng không giới hạn sáng chế được yêu cầu bảo hộ. Theo một phương án, máy tính của người dùng được bỏ qua, và thay vào đó là chức năng máy tính riêng biệt được tạo ra mà làm việc với máy chủ trung tâm. Trong trường hợp này, người dùng sẽ đăng nhập vào máy chủ từ máy tính khác và truy cập vào hệ thống thông qua môi trường người dùng.

Môi trường người dùng có thể được đặt trong máy chủ trung tâm hoặc được kết nối với nó. Hơn nữa, người dùng có thể nhận và truyền dữ liệu đến máy chủ trung tâm bằng mạng internet, nhờ đó người dùng truy cập tài khoản bằng trình duyệt web internet và trình duyệt hiển thị trang web tương tác được kết nối hoạt động với máy chủ trung tâm. Máy chủ trung tâm truyền và nhận dữ liệu để đáp ứng với dữ liệu và các lệnh được truyền từ trình duyệt để đáp ứng với hành động của khách hàng trên giao diện người dùng của trình duyệt.

Một số bước của sáng chế có thể được thực hiện trên máy tính của người dùng và kết quả tạm thời được truyền tới máy chủ. Các kết quả tạm thời này có thể được xử lý tại máy chủ và kết quả cuối cùng được trả lại cho người dùng.

Phương pháp được mô tả ở đây có thể được thực thi trên hệ thống máy tính, thường bao gồm bộ xử lý trung tâm (CPU) được kết nối hoạt động với thiết bị bộ nhớ, mạch đầu vào và đầu ra dữ liệu (I/O) và mạch giao tiếp mạng dữ liệu máy tính. Máy tính do CPU thực thi có thể lấy dữ liệu được nhận bởi mạch giao tiếp dữ liệu và lưu trữ nó trong thiết bị bộ nhớ. Ngoài ra, CPU có thể lấy dữ liệu từ mạch I/O và lưu trữ nó trong thiết bị bộ nhớ. Hơn nữa, CPU có thể lấy dữ liệu từ thiết bị bộ nhớ và xuất nó qua mạch I/O hoặc mạch giao tiếp dữ liệu. Dữ liệu được lưu trữ trong bộ nhớ có thể còn được thu hồi từ thiết bị bộ nhớ, được xử lý hoặc biến đổi thêm bởi CPU theo cách được mô tả ở đây và được khôi phục trong cùng thiết bị bộ nhớ hoặc thiết bị bộ nhớ khác được kết nối hoạt động với CPU bao gồm nhờ mạch mạng dữ liệu. Thiết bị bộ nhớ có thể là loại mạch lưu trữ dữ liệu bất kỳ hoặc lưu trữ từ tính hoặc thiết bị quang học, bao gồm đĩa cứng, đĩa quang hoặc bộ nhớ trạng thái rắn. Các thiết bị I/O có thể bao gồm màn hình hiển thị, loa, micrô và chuột di động mà chỉ báo đến máy tính vị trí tương đối của vị trí con trỏ trên màn hình và một hoặc nhiều nút có thể được kích hoạt để chỉ báo lệnh.

Máy tính có thể hiển thị trên màn hình hiển thị mà được kết nối hoạt động với mạch I/O sự xuất hiện của giao diện người dùng. Các hình dạng, văn bản và các dạng đồ họa khác nhau được hiển thị trên màn hình do máy tính tạo ra dữ liệu mà làm các điểm ảnh bao gồm hành động của khách hàng trên màn hình hiển thị của giao diện người dùng trình duyệt. Một số bước của sáng chế có thể được thực hiện trên máy tính của người dùng và kết quả tạm thời được truyền tới máy chủ. Các kết quả tạm thời này có thể được xử lý tại máy chủ và kết quả cuối cùng được trả lại cho người dùng.

Phương pháp được mô tả ở đây có thể được thực thi trên hệ thống máy tính, thường bao gồm bộ xử lý trung tâm (CPU) được kết nối hoạt động với thiết bị bộ nhớ, mạch đầu vào và đầu ra dữ liệu (I/O) và mạch giao tiếp mạng dữ liệu máy tính. Máy tính do CPU thực thi có thể lấy dữ liệu được nhận bởi mạch giao tiếp dữ liệu và lưu trữ nó trong thiết bị bộ nhớ. Ngoài ra, CPU có thể lấy dữ liệu từ mạch I/O và lưu trữ nó trong thiết bị bộ nhớ. Hơn nữa, CPU có thể lấy dữ liệu từ thiết bị bộ nhớ và xuất nó qua mạch I/O hoặc mạch

giao tiếp dữ liệu. Dữ liệu được lưu trữ trong bộ nhớ có thể còn được thu hồi từ thiết bị bộ nhớ, được xử lý hoặc biến đổi thêm bởi CPU theo cách được mô tả ở đây và được khôi phục trong cùng thiết bị bộ nhớ hoặc thiết bị bộ nhớ khác được kết nối hoạt động với CPU bao gồm nhờ mạch mạng dữ liệu. Thiết bị bộ nhớ có thể là loại mạch lưu trữ dữ liệu bất kỳ hoặc lưu trữ từ tính hoặc thiết bị quang học, bao gồm đĩa cứng, đĩa quang hoặc bộ nhớ trạng thái rắn. Các thiết bị I/O có thể bao gồm màn hình hiển thị, loa, micrô và chuột di động mà chỉ báo đến máy tính vị trí tương đối của vị trí con trỏ trên màn hình và một hoặc nhiều nút có thể được kích hoạt để chỉ báo lệnh.

Máy tính có thể hiển thị trên màn hình hiển thị mà được kết nối hoạt động với mạch I/O sự xuất hiện của giao diện người dùng. Các hình dạng, văn bản và các hình thức đồ họa khác nhau được hiển thị trên màn hình do máy tính tạo ra dữ liệu khiến cho các pixel bao gồm màn hình hiển thị có nhiều màu và sắc thái khác nhau. Giao diện người dùng cũng hiển thị đối tượng đồ họa được đề cập đến trong lĩnh vực kỹ thuật làm con trỏ. Vị trí của đối tượng trên màn hình chỉ báo đến người dùng sự lựa chọn đối tượng khác trên màn hình. Con trỏ có thể được di chuyển bởi người dùng bằng thiết bị khác được kết nối bằng mạch I/O đến máy tính. Thiết bị này phát hiện chuyển động vật lý nhất định của người sử dụng, ví dụ, vị trí của bàn tay trên bề mặt phẳng hoặc vị trí của ngón tay trên bề mặt phẳng. Các thiết bị này có thể được đề cập đến trong lĩnh vực kỹ thuật như chuột hoặc bàn di chuột cảm ứng. Theo một số phương án, màn hình hiển thị tự có thể hoạt động như một bàn di chuột cảm ứng bằng cách cảm nhận sự hiện diện và vị trí của một hoặc nhiều ngón tay trên bề mặt của màn hình hiển thị. Khi con trỏ nằm trên đối tượng đồ họa mà xuất hiện như nút hoặc nút chuyển, người dùng có thể kích hoạt nút hoặc nút chuyển bằng cách tương tác với nút chuyển vật lý trên chuột hoặc bàn di chuột cảm ứng hoặc thiết bị máy tính hoặc chạm vào bàn di chuột cảm ứng hoặc màn hình hiển thị cảm ứng. Khi máy tính phát hiện thấy công tắc vật lý đã được tương tác (hoặc là việc chạm vào bàn di chuột cảm ứng hoặc màn hình cảm ứng chạm đã xảy ra), sẽ có vị trí rõ ràng của con trỏ (hoặc trong trường hợp của màn hình cảm ứng chạm, vị trí được phát hiện của ngón tay) trên màn hình và thực hiện quá trình liên kết với vị trí đó. Ví dụ, không nhằm hạn chế phạm vi của sáng chế, đối tượng đồ họa mà xuất hiện như hộp 2 chiều với từ "nhập" bên trong nó có thể được hiển thị trên màn hình. Nếu máy tính phát hiện thấy nút chuyển đã được nhấn trong khi vị trí con trỏ (hoặc

vị trí ngón tay cho màn hình cảm ứng chạm) nằm trong ranh giới của đối tượng đồ họa, ví dụ, hộp được hiển thị, máy tính sẽ thực thi quá trình được liên kết với lệnh "nhập". Bằng cách này, các đối tượng đồ họa trên màn hình tạo ra giao diện người dùng cho phép người dùng kiểm soát các hoạt động xử lý trên máy tính.

Sáng chế cũng có thể được thực hiện hoàn toàn trên một hoặc nhiều máy chủ. Máy chủ có thể là máy tính bao gồm bộ xử lý trung tâm có thiết bị lưu trữ dung lượng lớn và kết nối mạng. Ngoài ra, máy chủ có thể bao gồm nhiều máy tính được kết nối cùng với mạng dữ liệu hoặc kết nối truyền dữ liệu khác, hoặc nhiều máy tính trên mạng với bộ nhớ truy cập mạng, theo cách đó tạo ra chức năng như một nhóm. Những người thực hiện có trình độ trung bình sẽ nhận ra rằng các chức năng được thực hiện trên một máy chủ có thể được phân chia và được thực hiện trên nhiều máy chủ mà được kết nối hoạt động bởi mạng máy tính bằng phương tiện giao tiếp giữa các quá trình thích hợp. Ngoài ra, việc truy cập trang web có thể bằng trình duyệt internet mà truy cập trang an toàn hoặc công cộng hoặc bằng chương trình máy khách chạy trên máy tính cục bộ được kết nối qua mạng máy tính với máy chủ. Thông báo dữ liệu và việc tải lên hoặc tải xuống dữ liệu có thể được gửi qua internet bằng cách sử dụng các giao thức điển hình, bao gồm TCP/IP, HTTP, TCP, UDP, SMTP, RPC, FTP hoặc các loại giao thức giao tiếp dữ liệu khác cho phép các tiến trình đang chạy trên hai máy tính từ xa trao đổi thông tin bằng phương tiện giao tiếp mạng kỹ thuật số. Kết quả là thông báo dữ liệu có thể là gói dữ liệu được truyền từ hoặc nhận bởi máy tính mà chứa địa chỉ mạng đích, quá trình đích hoặc mã định danh ứng dụng và các trị số dữ liệu có thể được phân tích cú pháp tại máy tính đích được đặt ở địa chỉ mạng đích bởi ứng dụng đích để các trị số dữ liệu có liên quan được trích xuất và được sử dụng bởi ứng dụng đích. Kiến trúc chính xác của máy chủ trung tâm không giới hạn sáng chế được yêu cầu bảo hộ. Ngoài ra, mạng dữ liệu có thể hoạt động với một số cấp độ, sao cho máy tính của người dùng được kết nối thông qua tường lửa tới máy chủ, mà định tuyến các giao tiếp tới máy chủ khác mà thực hiện các phương pháp được bộc lộ.

Máy tính người dùng có thể thực hiện chương trình mà nhận từ máy chủ từ xa tệp dữ liệu được truyền tới một chương trình mà biên dịch dữ liệu trong tệp dữ liệu và lệnh cho thiết bị hiển thị văn bản, hình ảnh, video, âm thanh và các đối tượng khác cụ thể. Chương trình có thể phát hiện vị trí tương đối của con trỏ khi nút chuột được kích hoạt và biên dịch

lệnh cần được thực thi dựa trên vị trí trên vị trí tương đối được chỉ báo trên màn hình khi nhấn nút. Tập dữ liệu có thể là tài liệu HTML, chương trình, chương trình trình duyệt web và lệnh là siêu liên kết mà khiến trình duyệt yêu cầu tài liệu HTML mới từ vị trí địa chỉ mạng dữ liệu từ xa khác. HTML cũng có thể có các tham chiếu dẫn đến các mô-đun mã khác được gọi lên và được thực thi, ví dụ, Flash hoặc mã gốc khác.

Những người có trình độ trong lĩnh vực kỹ thuật có liên quan sẽ hiểu là sáng chế có thể được thực hiện với các giao tiếp, xử lý dữ liệu hoặc cấu hình hệ thống máy tính khác, bao gồm: thiết bị không dây, thiết bị internet, thiết bị cầm tay (bao gồm thiết bị kỹ thuật số hỗ trợ cá nhân (PDA)), máy tính đeo được, tất cả các cách thức của điện thoại di động hoặc di động, hệ thống đa bộ xử lý, thiết bị điện tử người dùng có thể lập trình hoặc trên cơ sở vi xử lý, các bộ giải mã set-top-box, các máy tính mạng, máy tính mini, máy tính lớn và các loại tương tự. Thật vậy, các thuật ngữ "máy tính", "máy chủ" và các loại tương tự được sử dụng hoán đổi cho nhau ở đây và có thể đề cập đến thiết bị và hệ thống bất kỳ trong số các thiết bị và hệ thống ở trên.

Trong một số trường hợp, đặc biệt là máy tính người dùng là thiết bị điện toán di động được sử dụng để truy cập dữ liệu qua mạng, mạng có thể là loại bất kỳ trong số mạng viễn thông tập trung di động hoặc trên cơ sở IP, bao gồm nhưng không giới hạn ở hệ thống thông tin di động toàn cầu (GSM), đa truy nhập phân chia theo thời gian (TDMA), đa truy nhập phân chia theo mã (CDMA), đa truy nhập phân tần trực giao (OFDM), dịch vụ vô tuyến gói tổng hợp (GPRS), môi trường GSM dữ liệu nâng cao (EDGE), hệ thống điện thoại di động tiên tiến (AMPS), mạng khả năng tương tác toàn cầu với truy nhập vi ba (WiMAX), hệ thống viễn thông di động toàn cầu (UMTS), mạng Cải tiến-Tối ưu hóa dữ liệu (EVDO), công nghệ mạng phát triển lâu dài (LTE), mạng băng thông rộng siêu di động (UMB), công nghệ truyền giọng nói trên giao thức IP (VoIP) hoặc công nghệ truy cập di động không cấp phép (UMA).

Internet là mạng máy tính cho phép khách hàng vận hành máy tính cá nhân để tương tác với máy chủ máy tính được đặt từ xa và xem nội dung được phân phối từ máy chủ đến máy tính cá nhân dưới dạng tập dữ liệu qua mạng. Trong một loại giao thức, các máy chủ hiển thị các trang web được hiển thị trên máy tính cá nhân của khách hàng bằng cách sử dụng chương trình cục bộ được gọi là trình duyệt. Trình duyệt sẽ nhận một hoặc nhiều tệp

dữ liệu từ máy chủ được hiển thị trên màn hình máy tính cá nhân của khách hàng. Trình duyệt tìm kiếm các tệp dữ liệu đó từ địa chỉ cụ thể, được biểu diễn bằng một chuỗi ký tự chữ và số được gọi là định vị tài nguyên đa năng (URL). Tuy nhiên, trang web có thể chứa các thành phần được tải xuống từ nhiều địa chỉ URL hoặc địa chỉ IP khác nhau. Trang web là tập hợp các URL có liên quan, thường là tất cả chia sẻ cùng một địa chỉ gốc hoặc dưới sự kiểm soát của một số thực thể. Theo một phương án, các vùng khác nhau có không gian mô phỏng có các URL khác nhau. Nghĩa là, không gian mô phỏng có thể là cấu trúc dữ liệu đơn nhất, nhưng các vị trí khác nhau tham chiếu của URL khác nhau trong cấu trúc dữ liệu. Điều này làm cho nó có thể mô phỏng khu vực rộng lớn và có người tham gia bắt đầu sử dụng nó trong khu vực lân cận ảo của họ.

Logic chương trình máy tính thực hiện tất cả hoặc một phần chức năng được mô tả ở đây có thể được thể hiện dưới nhiều hình thức, bao gồm, nhưng không giới hạn, biểu mẫu mã nguồn, biểu mẫu thực thi máy tính và các biểu mẫu trung gian khác nhau (ví dụ, biểu mẫu được tạo bởi một trình biên dịch bậc thấp, trình biên dịch, trình liên kết hoặc trình định vị). Mã nguồn có thể bao gồm một loạt các lệnh chương trình máy tính được thực hiện bằng ngôn ngữ lập trình bất kỳ khác nhau (ví dụ, mã đối tượng, hợp ngữ, hoặc ngôn ngữ cấp cao như C, C-HF, C #, Action Script, PHP, EcmaScript, JavaScript, JAVA hoặc 5 HTML) để sử dụng với các hệ điều hành hoặc môi trường hoạt động khác nhau. Mã nguồn có thể xác định và sử dụng các cấu trúc dữ liệu khác nhau và thông điệp truyền thông. Mã nguồn có thể ở dạng thực thi máy tính (ví dụ, thông qua trình thông dịch) hoặc mã nguồn có thể được chuyển đổi (ví dụ, thông qua trình dịch, trình biên dịch bậc thấp hoặc trình biên dịch) thành dạng thực thi trên máy tính.

Sáng chế có thể được mô tả trong trường hợp chung của các lệnh thực thi máy tính, chẳng hạn như các mô-đun chương trình, được thực thi bởi máy tính. Nói chung, các mô-đun chương trình bao gồm các đoạn chương trình, chương trình, đối tượng, thành phần, cấu trúc dữ liệu, v.v., thực hiện các tác vụ cụ thể hoặc thực hiện các kiểu dữ liệu trừu tượng cụ thể. Chương trình máy tính và dữ liệu có thể được sửa theo bất kỳ hình thức nào (ví dụ, biểu mẫu mã nguồn, biểu mẫu thực thi trên máy tính hoặc biểu mẫu trung gian) hoặc vĩnh viễn hoặc tạm thời trong môi trường lưu trữ hữu hình, chẳng hạn như thiết bị nhớ bán dẫn (ví dụ, RAM, ROM, PROM, EEPROM hoặc RAM có thể lập trình Flash), thiết bị nhớ từ

(ví dụ, đĩa mềm hoặc đĩa cứng cố định), thiết bị nhớ quang (ví dụ, CD-ROM hoặc DVD), thẻ PC (ví dụ, thẻ PCMCIA) hoặc thiết bị bộ nhớ khác. Chương trình máy tính và dữ liệu có thể được cố định dưới bất kỳ hình thức nào trong tín hiệu mà có thể truyền đến máy tính bằng cách sử dụng công nghệ giao tiếp bất kỳ, bao gồm, nhưng không giới hạn ở, công nghệ tương tự, công nghệ số, công nghệ quang học, công nghệ không dây, công nghệ mạng và các công nghệ liên mạng. Chương trình máy tính và dữ liệu có thể được phân phối dưới hình thức bất kỳ như phương tiện lưu trữ có thể tháo rời kết hợp với tài liệu in hoặc điện tử đi kèm (ví dụ, phần mềm được đóng gói thu nhỏ hoặc băng từ), được tải sẵn với hệ thống máy tính (ví dụ, trên ROM hệ thống hoặc đĩa cố định), hoặc được phân phối từ máy chủ hoặc bảng thông báo điện tử trên hệ thống truyền thông (ví dụ, internet hoặc mạng lưới toàn cầu). Cần hiểu là thành phần bất kỳ trong số các thành phần phần mềm của sáng chế có thể, nếu muốn, được thực hiện trong biểu mẫu ROM (chỉ đọc bộ nhớ). Các thành phần phần mềm có thể, nói chung, được thực thi trong phần cứng, nếu muốn, bằng cách sử dụng các kỹ thuật thông thường.

Sáng chế cũng có thể được thực hành trong các môi trường điện toán phân tán, nơi các tác vụ được thực hiện bởi các thiết bị xử lý từ xa được liên kết thông qua mạng truyền thông. Trong môi trường điện toán phân tán, các mô đun chương trình có thể được đặt trong cả phương tiện lưu trữ máy tính cục bộ và từ xa bao gồm các thiết bị lưu trữ bộ nhớ. Những người thực hiện có kỹ năng trung bình sẽ nhận ra rằng sáng chế có thể được thực hiện trên một hoặc nhiều bộ vi xử lý máy tính mà được liên kết bằng cách sử dụng mạng dữ liệu, bao gồm, ví dụ, internet. Theo phương án khác, các bước khác nhau của quy trình có thể được thực thi bởi một hoặc nhiều máy tính và thiết bị lưu trữ được phân tách theo địa lý nhưng được kết nối bởi mạng dữ liệu theo cách để chúng hoạt động cùng nhau để thực hiện các bước quy trình. Theo một phương án, máy tính của người dùng có thể chạy ứng dụng làm cho máy tính của người dùng truyền luồng của một hoặc nhiều gói dữ liệu trên mạng dữ liệu sang một máy tính thứ hai, ở đây được gọi là máy chủ. Các máy chủ, lần lượt, có thể được kết nối với một hoặc nhiều thiết bị lưu trữ dữ liệu hàng loạt, nơi cơ sở dữ liệu được lưu trữ. Máy chủ có thể thực thi chương trình nhận gói tin được truyền và biên dịch các gói dữ liệu được truyền để trích xuất thông tin truy vấn cơ sở dữ liệu. Sau đó, máy chủ có thể thực hiện các bước còn lại của sáng chế bằng cách truy cập các thiết bị lưu trữ lớn để rút ra

kết quả truy vấn mong muốn. Ngoài ra, máy chủ có thể truyền thông tin truy vấn đến một máy tính khác mà được kết nối với thiết bị lưu trữ khối và máy tính đó có thể thực hiện sáng chế để rút ra kết quả mong muốn. Kết quả sau đó có thể được truyền lại cho máy tính của người dùng bằng luồng khác gồm một hoặc nhiều gói dữ liệu được giải quyết thích hợp cho máy tính của người dùng. Theo một phương án, cơ sở dữ liệu quan hệ có thể được đặt trong một hoặc nhiều máy chủ kết nối hoạt động được kết nối với bộ nhớ máy tính, ví dụ, ổ đĩa. Theo phương án khác nữa, việc khởi tạo cơ sở dữ liệu quan hệ có thể được chuẩn bị trên bộ máy chủ và sự tương tác với máy tính của người dùng xảy ra ở vị trí khác trong quy trình tổng thể.

Cần lưu ý rằng các lưu đồ được sử dụng ở đây để chứng minh các khía cạnh khác nhau của sáng chế, và không nên được hiểu là hạn chế sáng chế vào luồng logic hoặc sự thực hiện logic cụ thể bất kỳ. Logic được mô tả có thể được phân đoạn thành các khối logic khác nhau (ví dụ, các chương trình, mô-đun, hàm hoặc các đoạn chương trình con) mà không thay đổi kết quả tổng thể hoặc nằm ngoài phạm vi thực sự của sáng chế. Thông thường, các phần tử logic có thể được bổ sung, sửa đổi, bỏ qua, được thực hiện theo thứ tự khác hoặc được thực hiện bằng cách sử dụng các cấu trúc logic khác nhau (ví dụ, cổng logic, vòng lặp nguyên thủy, logic điều kiện và các cấu trúc logic khác) mà không thay đổi kết quả tổng thể hoặc phạm vi thực sự của sáng chế.

Các phương án được mô tả của sáng chế được nhằm làm ví dụ và nhiều biến thể và các biến đổi sẽ rõ ràng đối với những người có trình độ trong lĩnh vực kỹ thuật. Tất cả các biến thể và biến đổi như vậy được dự định nằm trong phạm vi của sáng chế như được xác định trong các yêu cầu bảo hộ kèm theo. Mặc dù sáng chế đã được mô tả và minh họa chi tiết, cần phải được hiểu rõ ràng là các mô tả này chỉ nhằm làm minh họa và ví dụ, và không nhằm làm giới hạn sáng chế. Cần phải hiểu là các đặc điểm khác nhau của sáng chế, cho mục đích rõ ràng, được mô tả trong trường hợp của các phương án riêng biệt cũng có thể được đề xuất ở dạng kết hợp trong một phương án duy nhất.

YÊU CẦU BẢO HỘ

1. Phương pháp biên dịch các khung hình từ video, lưu và lập chỉ mục các khung hình trong cơ sở dữ liệu bao gồm các bước:

- a) bước thu video;
- b) bước phân tích video cho các đặc điểm;
- c) bước phân tách video thành các đoạn dựa trên các đặc điểm, trong đó các đoạn được xác định bằng cách nhóm các khung hình mà tương ứng với các cảnh giống nhau hoặc tương tự, trong đó các đoạn trùng lặp trong video được loại bỏ, trong đó các đoạn trong video mà đã có trong cơ sở dữ liệu được loại bỏ, trong đó các khung hình riêng lẻ trong các đoạn được trích xuất và các đặc điểm trực quan được tạo ra cho mỗi khung hình và trong đó các đoạn được chú thích với siêu dữ liệu;
- d) bước phân tích các đặc điểm trực quan của các đoạn thành các đoạn nhóm mà chia sẻ các đặc điểm trực quan;
- e) bước phân tích sự phù hợp của mỗi khung hình;
- f) bước tạo ra cảnh báo nếu xác định thấy sự không phù hợp;
- g) bước chú thích một hoặc nhiều khung hình của các đoạn với siêu dữ liệu nếu xác định thấy sự phù hợp, trong đó các đặc điểm trực quan được trích xuất thành vector đặc trưng, trong đó mỗi vector đặc trưng được kết hợp với siêu dữ liệu và các ký hiệu nhận dạng duy nhất để tạo ra cấu trúc dữ liệu tổng hợp bao gồm ký hiệu nhận dạng khung, ký hiệu nhận dạng đoạn, các đặc điểm trực quan khung hình và siêu dữ liệu; và
- h) bước lưu một hoặc nhiều khung hình của các đoạn vào trong cơ sở dữ liệu, trong đó cơ sở dữ liệu được truy vấn bằng cách sử dụng cấu trúc dữ liệu tổng hợp để xác định liệu nội dung được liên kết có tồn tại trong cơ sở dữ liệu và hoặc i) tạo ra bản ghi chèn đoạn và chèn các khung hình vào trong cơ sở dữ liệu hoặc ii) tạo ra nội dung để chèn vào trong cơ sở dữ liệu.

2. Phương pháp theo điểm 1, bao gồm bước làm giảm bổ sung tốc độ khung hình và/hoặc độ phân giải của video.

3. Phương pháp theo điểm 1, bao gồm hai bước bổ sung gồm bước truy cập cơ sở dữ liệu khi người sử dụng gửi truy vấn và bước so khớp truy vấn với một hoặc nhiều khung hình của các đoạn trong cơ sở dữ liệu dựa trên các đặc điểm chung.

4. Hệ thống máy tính để nhập video bao gồm: 2

giao diện người dùng để tải video lên máy chủ;

chương trình máy tính để xử lý video và trích xuất các đoạn từ video làm các khung hình, trong đó các đặc điểm trực quan được tạo ra cho mỗi khung hình và trong đó các đoạn được trích xuất được chú thích với siêu dữ liệu,

trong đó các đoạn được xác định bằng cách nhóm các khung hình mà tương ứng với các cảnh giống nhau hoặc tương tự, trong đó các đoạn trùng lặp trong video được loại bỏ, trong đó các đoạn trong video mà đã có trong cơ sở dữ liệu được loại bỏ;

mô đun để xử lý mỗi khung trong các đoạn và trích xuất các đặc điểm trực quan thành vectơ đặc trưng, trong đó mỗi vectơ đặc trưng được kết hợp với siêu dữ liệu và các ký hiệu nhận dạng duy nhất để tạo ra cấu trúc dữ liệu tổng hợp bao gồm ký hiệu nhận dạng khung, ký hiệu nhận dạng đoạn, các đặc điểm trực quan khung hình và siêu dữ liệu;

cơ sở dữ liệu để lưu các đoạn được trích xuất, trong đó cơ sở dữ liệu được truy vấn bằng cách sử dụng cấu trúc dữ liệu tổng hợp để xác định liệu nội dung được liên kết có tồn tại trong cơ sở dữ liệu và hoặc i) tạo ra bản ghi chèn đoạn và chèn các khung hình vào trong cơ sở dữ liệu hoặc ii) tạo ra nội dung để chèn vào trong cơ sở dữ liệu;

trong đó các đoạn được trích xuất được phân tích để tuân thủ các thông số phù hợp, bao gồm các đặc điểm hình ảnh và độ phân giải; và

trong đó sự phù hợp của mỗi khung hình được phân tích và phản hồi được tạo ra nếu phát hiện sự không phù hợp của khung hình.

5. Hệ thống máy tính để nhập video theo điểm 4, bao gồm phương tiện để nhận truy vấn và mô đun để so khớp truy vấn với một hoặc nhiều phần video trong cơ sở dữ liệu dựa trên các đặc điểm chung.

6. Phương pháp biên dịch các hình ảnh từ phương tiện dạng bản in vào trong cơ sở dữ liệu bao gồm các bước:

- a) bước thu phương tiện dạng bản in;
- b) bước chuyển đổi phương tiện dạng bản in thành phương tiện kỹ thuật số;
- c) bước phân đoạn phương tiện kỹ thuật số thành một hoặc nhiều đoạn;
- d) bước phát hiện văn bản trong một hoặc nhiều đoạn dựa trên các đường viền được nối của các mép, trong đó các mép được phát hiện bằng phép toán gradient hình thái học, trong đó nhị phân hóa được thực hiện bằng phương pháp Otsu, trong đó các viền được phát hiện bằng sự phân tích các thành phần được nối để cô lập các từ, các đường và các văn bản gần kề, trong đó các phần nhô dọc của các đường viền phát hiện liệu các đường viền chứa nhiều hơn một dòng, trong đó các dòng đơn và các từ được nối và được dán nhãn thành các đoạn và các dòng đơn lớn trong các hình chữ nhật mà biểu diễn các khung đường bao hoặc các vùng có viền trong đó các vùng viền được lọc và xác nhận sử dụng diện tích vùng, kích thước, tỷ lệ của pixel khác không trong các vùng viền
- e) bước loại bỏ các vùng có văn bản từ một hoặc nhiều đoạn;
- f) bước phát hiện một hoặc nhiều hình ảnh trong một hoặc nhiều đoạn;
- g) bước phát hiện liệu có đủ các đặc điểm để phân loại một hoặc nhiều hình ảnh;
- h) bước tạo ra cảnh báo nếu các đặc điểm là không đủ;
- i) bước phân loại một hoặc nhiều hình ảnh nếu các đặc điểm là đủ; và
- j) bước chèn một hoặc nhiều hình ảnh vào trong cơ sở dữ liệu, trong đó việc chèn bao gồm sự trích xuất của các đặc điểm trực quan từ một hoặc nhiều hình ảnh và tạo ra phản hồi nhập bao gồm vectơ của các đặc điểm trực quan, ID đoạn và ID nội dung được liên kết.

7. Phương pháp theo điểm 6, bao gồm hai bước bổ sung gồm bước truy cập cơ sở dữ liệu khi người sử dụng gửi truy vấn và bước so khớp truy vấn với một hoặc nhiều hình ảnh trong cơ sở dữ liệu bằng cách so sánh các đặc điểm.

8. Phương pháp theo điểm 6, bao gồm bước chú thích bổ sung một hoặc nhiều hình ảnh với siêu dữ liệu.

9. Hệ thống máy tính để nhập các tài liệu vào trong cơ sở dữ liệu bao gồm:
giao diện người dùng để tải tài liệu lên máy chủ;

mô đun bao gồm logic để xử lý tài liệu và phân đoạn tài liệu thành một hoặc nhiều các hình ảnh dạng đoạn;

mô đun bao gồm logic để loại bỏ văn bản khỏi các hình ảnh dựa trên các đường viền được nối của các mép, trong đó các mép được phát hiện bằng phép toán gradient hình thái học, trong đó nhị phân hóa được thực hiện bằng phương pháp Otsu, trong đó các viền được phát hiện bằng sự phân tích các thành phần được nối để cô lập các từ, các dòng và các văn bản gần kề, trong đó phần nhô dọc của các đường viền phát hiện liệu các đường viền chứa nhiều hơn một dòng, trong đó các dòng đơn và các từ được nối và được dán nhãn thành các đoạn và các dòng đơn lớn trong các hình chữ nhật mà biểu diễn các khung đường bao hoặc các vùng có viền;

trong đó các vùng viền được lọc và xác nhận sử dụng diện tích vùng, kích thước, tỷ lệ của pixel khác không trong các vùng viền; mô đun bao gồm logic để phát hiện các đặc điểm trong các hình ảnh;

mô đun bao gồm logic để trích xuất các đặc điểm trực quan từ các hình ảnh và tạo ra phản hồi nhập bao gồm vectơ của các đặc điểm trực quan, ID đoạn và ID nội dung được liên kết; và

cơ sở dữ liệu để chèn hình ảnh, trong đó các đặc điểm được phát hiện được phân tích về tính hợp lệ của các hình ảnh;

và trong đó phản hồi được tạo ra nếu tính không hợp lệ được xác định.

10. Hệ thống máy tính để nhập các tài liệu theo điểm 9, bao gồm phương tiện để nhận truy vấn và mô đun để so khớp truy vấn với một hoặc nhiều hình ảnh trong cơ sở dữ liệu dựa trên các đặc điểm chung.

1/5

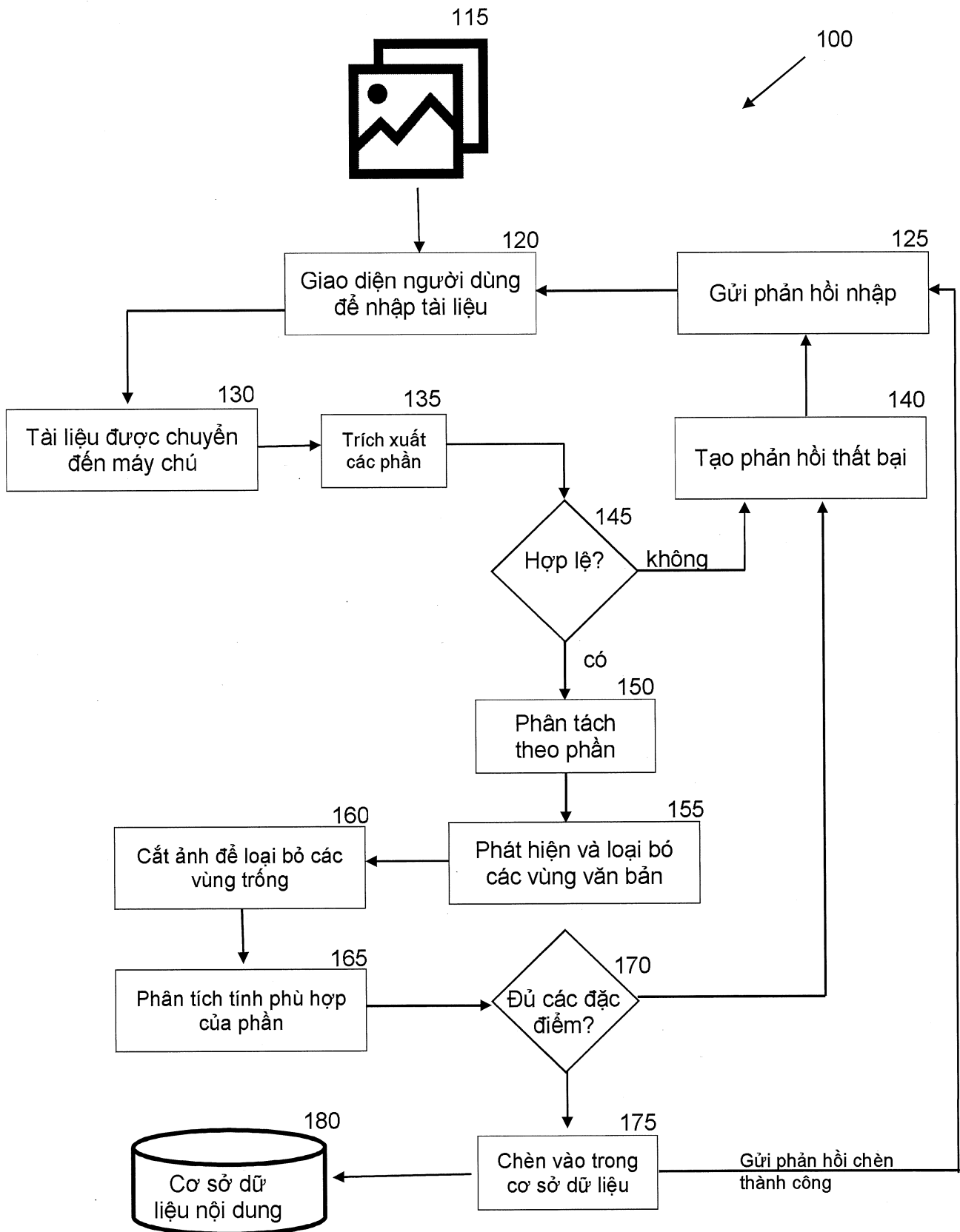


FIG. 1

2/5

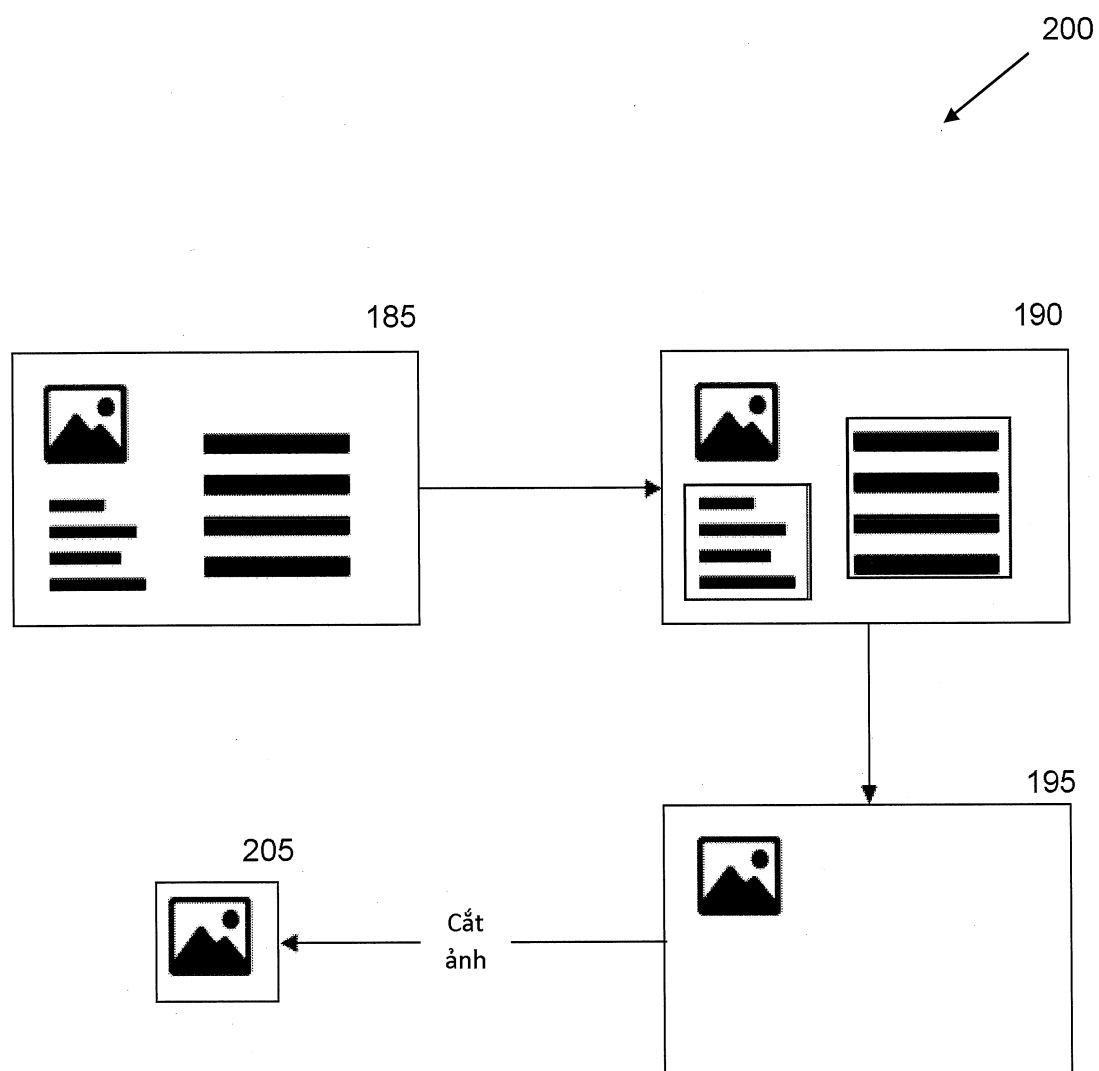


FIG. 2

3/5

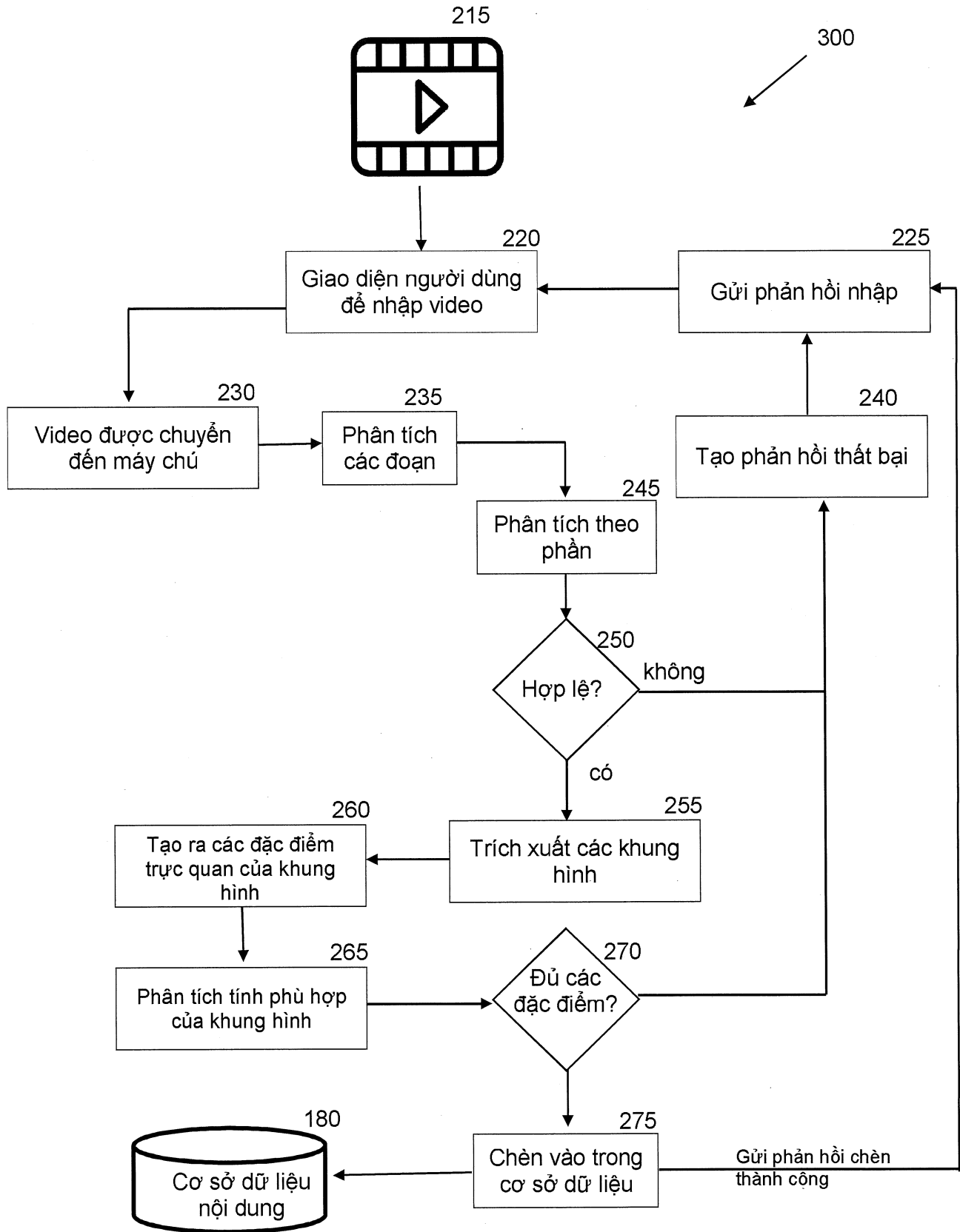


FIG. 3

4/5

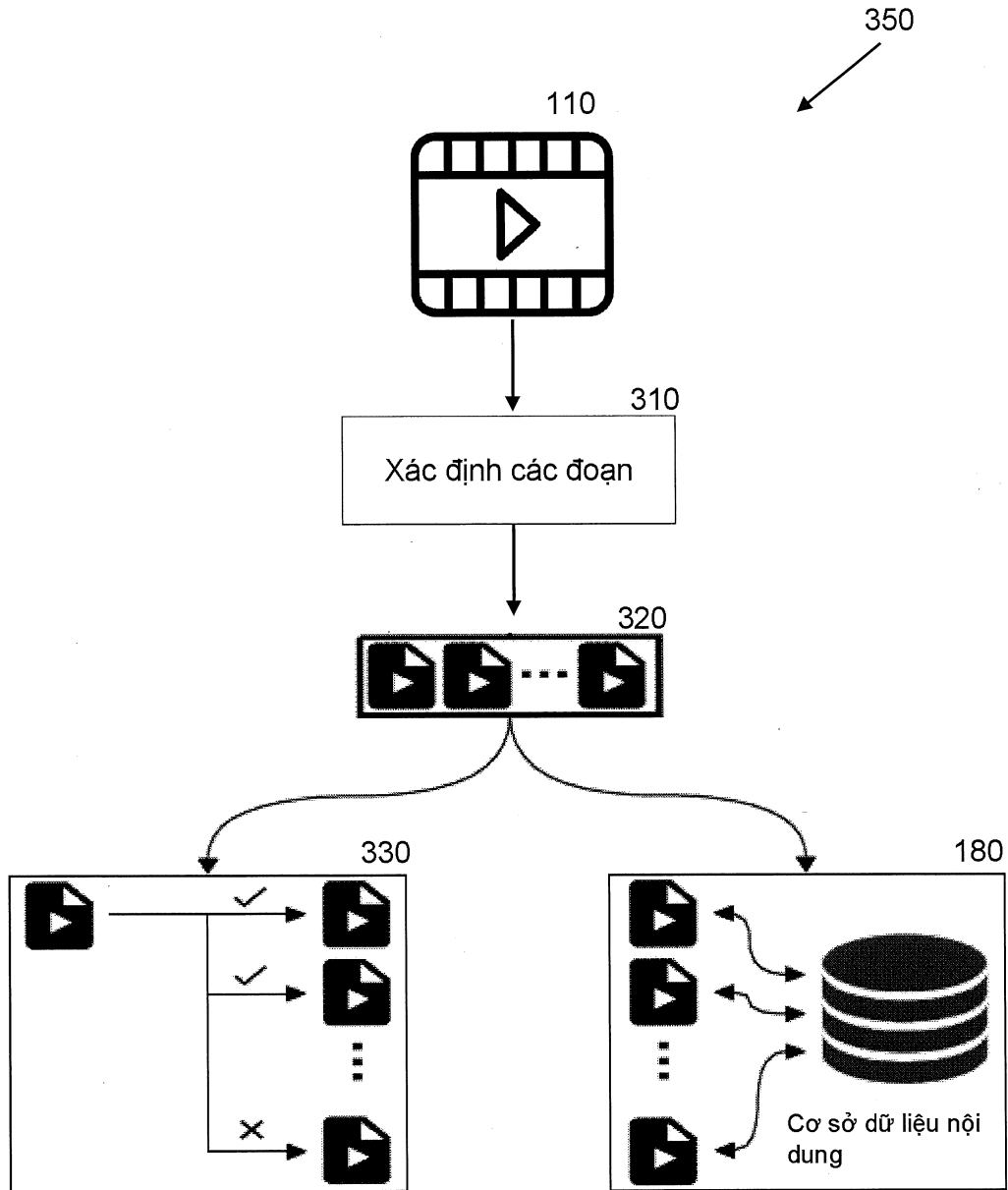


FIG. 4

5/5

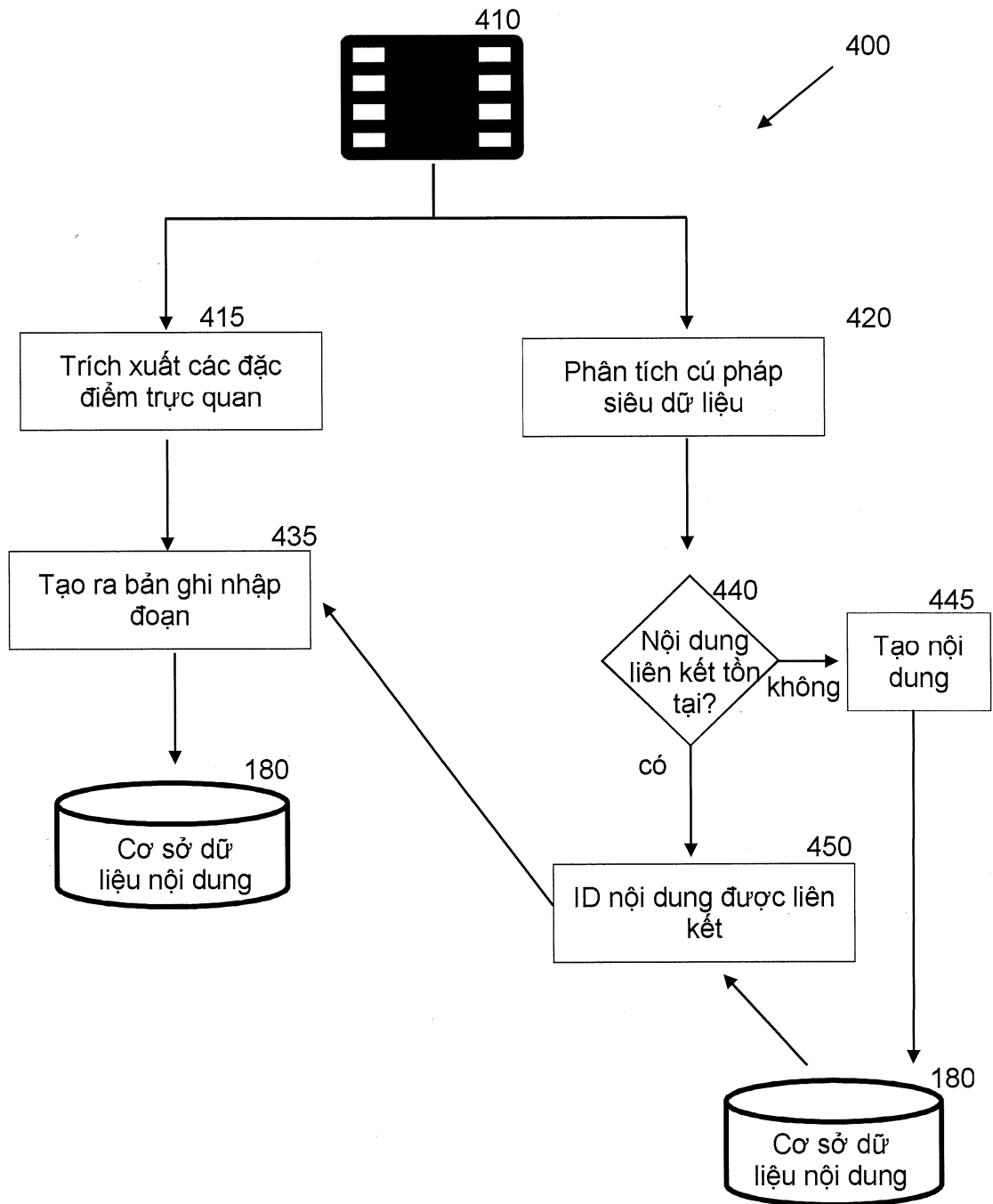


FIG. 5