



(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)



1-0033282

(51)⁷ G06K 9/00; G06F 17/30

(13) B

(21) 1-2018-05754

(22) 07/09/2017

(86) PCT/SG2017/050449 07/09/2017

(87) WO/2018/048355 15/03/2018

(30) 62/384,855 08/09/2016 US

(45) 25/09/2022 414

(43) 27/05/2019 374A

(73) GOH SOO SIAH (SG)

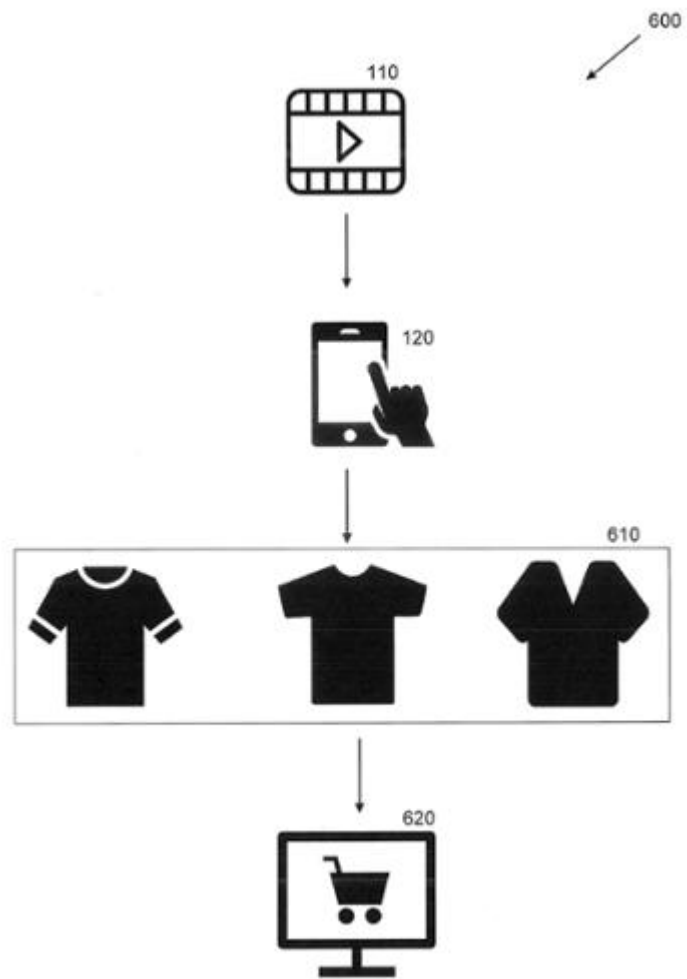
23 Greenleaf Drive, Ban Guan Park, Singapore, 279536

(72) MOORE, Stephen Maurice (SG); MURRAY, Larry Patrick (SG);
SHANMUGAMANI, Rajalingappaa (SG).

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

(54) PHƯƠNG PHÁP PHÁT HIỆN ĐỐI TƯỢNG TRONG VIDEO VÀ HỆ THỐNG
TẠO MỐI QUAN HỆ GIỮA CÁC ĐỐI TƯỢNG TRONG VIDEO VỚI CÁC SẢN
 PHẨM TRONG CƠ SỞ DỮ LIỆU CỦA CÁC SẢN PHẨM

(57) Sáng chế đề cập đến phương pháp phát hiện đối tượng trong video và hệ thống tạo mối quan hệ giữa các đối tượng trong video với các sản phẩm trong cơ sở dữ liệu của các sản phẩm. Cơ sở dữ liệu có thể được điền với sự tăng cường dữ liệu ngoại tuyến (ví dụ trình thu thập dữ liệu web) hoặc bằng cách căn chỉnh các đối tượng đã biết và các cụm siêu dữ liệu với nội dung đã được xác định. Người xem có thể truy vấn các hình ảnh từ phương tiện trực tiếp hoặc ngoại tuyến. Các đối tượng trong truy vấn người xem được liên kết với các đối tượng tương tự hoặc các sản phẩm được giới thiệu trong cơ sở dữ liệu.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến công nghệ máy tính để tiếp thị qua mạng, và cụ thể hơn, đề cập đến ứng dụng được vi tính hóa được nối mạng, để chia nhỏ video thành các phân đoạn, định danh các đối tượng trong các phân đoạn và so khớp các sản phẩm với các đối tượng.

Tình trạng kỹ thuật của sáng chế

Thương mại điện tử là giao dịch mua hoặc bán trực tuyến. Thương mại điện tử đã trở thành công cụ quan trọng cho thế giới kinh doanh, không những để bán cho các khách hàng, mà còn để thu hút họ. Năm 2012, doanh thu thương mại điện tử toàn cầu đã vượt quá 1 nghìn tỷ đô la.

Tiếp thị qua mạng là các nỗ lực quảng cáo và tiếp thị mà sử dụng Web và e-mail để thúc đẩy doanh thu thông qua thương mại điện tử. Nó bao gồm tiếp thị qua e-mail, tiếp thị qua công cụ tìm kiếm (SEM- search engine marketing), tiếp thị qua mạng xã hội, nhiều loại quảng cáo hiển thị (ví dụ quảng cáo biểu ngữ) và quảng cáo trên điện thoại di động. Siêu dữ liệu là thành phần thiết yếu của tiếp thị qua mạng.

Bởi vì những người mua sắm trực tuyến không thể nhìn thấy vật phẩm tận mắt, họ thường tìm kiếm bằng tiêu chí chẳng hạn như từ khóa. Các doanh nghiệp thường lưu trữ siêu dữ liệu về các tìm kiếm và các giao dịch mà cho phép họ phân tích xu hướng bán hàng, phát triển các kế hoạch tiếp thị và đưa ra các dự đoán. Cũng chính siêu dữ liệu này cho phép các doanh nghiệp cung cấp trải nghiệm mua sắm được cá nhân hóa hơn với các đặc trưng chẳng hạn như lịch sử mua hàng, sở địa chỉ cho nhiều địa điểm giao hàng, và đề xuất sản phẩm.

Ngày nay, hầu hết các trang web có siêu dữ liệu được nhúng trong đó. Các công cụ tìm kiếm web xây dựng các chỉ mục rộng lớn mà sử dụng văn bản trong trang và siêu dữ liệu được 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 có mục tiêu. Các nhà quảng cáo

có thể sử dụng các phương pháp phức tạp để nhắm đến các khán giả dễ tiếp thu nhất với những đặc điểm nhất định, dựa trên sản phẩm nhà quảng cáo đang quảng bá.

Ví dụ, một người có thể sử dụng trình duyệt web để tìm kiếm các chuyến bay đến New Zealand. Siêu dữ liệu dưới dạng “các tệp lưu trữ thông tin tùy chọn cá nhân” từ các trang web mà họ truy cập sẽ được lưu trữ ở máy tính bởi trình duyệt web của người dùng. Các tệp lưu trữ thông tin tùy chọn cá nhân được gửi qua lại giữa máy chủ internet và trình duyệt mà cho phép người dùng được định danh và/hoặc để theo dõi các hoạt động của họ. Sau đó, người dùng có thể nhận các quảng cáo chẳng hạn như biểu ngữ quảng cáo liên quan đến du lịch ở New Zealand chẳng hạn như thông tin khách sạn, dịch vụ cho thuê xe, chuyến du lịch và chuyến bay.

Hơn nữa, siêu dữ liệu có thể được sử dụng để định danh và nhắm đến người dùng dựa trên các đối tượng khách hàng. Doanh nghiệp có thể nhận ra là sản phẩm hấp dẫn đến một số đối tượng khách hàng nhất định và hướng các nỗ lực tiếp thị đến đối tượng khách hàng này. Ví dụ, biểu ngữ quảng cáo cho đầu tư chứng khoán có thể không hiệu quả cho khán giả là thanh thiếu niên. Nhắm quảng cáo đến đối tượng khách hàng già hơn, và cụ thể hơn là đến những người tìm kiếm sự nghỉ hưu, sẽ hiệu quả hơn. Người dùng sẽ không nhận các quảng cáo nếu siêu dữ liệu cho thấy rằng họ nằm ngoài đối tượng khách hàng mục tiêu.

Với sự tiến bộ của thị trường cho thấy sự gia tăng các trường hợp sử dụng tìm kiếm trực quan, việc làm phong phú thêm hồ sơ đối tượng khách hàng là mối quan tâm lớn và có lợi cho người tiêu dùng cũng như công ty. Các tổ chức có thể sử dụng các mô hình như này để cung cấp các dịch vụ được cá nhân hóa cao cao, ước tính các mẫu hình sử dụng (ví dụ, dựa trên mùa), và thậm chí là mở rộng phạm vi định hướng sản phẩm tương lai.

Các tìm kiếm trực quan trở nên phổ biến, đặc biệt là với sự phổ biến của điện thoại thông minh và máy tính bảng. Ví dụ, người dùng trong ví dụ nêu trên có thể tìm kiếm các hình ảnh liên quan đến leo núi và đi bộ đường dài ở New Zealand. Bởi vì họ nhấp chuột vào các bức ảnh, trình duyệt không thể ghi lại siêu dữ liệu dựa trên các từ khóa được gõ vào trong trình duyệt. Tương tự, họ có thể xem video liên quan đến các sự hấp dẫn ở New Zealand. Với công nghệ thông thường, cả hình ảnh hoặc video đều sẽ không đóng góp siêu dữ liệu cho hoạt động tiếp thị có mục tiêu.

Tuy nhiên, sự gia tăng việc phát trực tuyến và tải xuống video cũng mang đến các cơ hội mới. Ví dụ, người nổi tiếng có thể xuất hiện trong video trên mạng xã hội với một túi xách cụ thể. Doanh số của túi xách có thể tăng trưởng nhanh nếu nó có thể định danh được và có sẵn để mua.

Với sự phát triển và phổ biến của các nền tảng phân phối video dựa trên Internet với nhiều kiến trúc khác nhau, tìm kiếm trực quan trên nội dung của video có tiềm năng tiếp cận cơ sở hàng triệu người dùng bao gồm các nhà sáng tạo nội dung, người tiêu dùng và đối tác thương mại. Nếu các phân đoạn cụ thể của video có thể được định danh, các bên quan tâm có được khả năng tăng cường và/hoặc kết hợp các đoạn này với các nội dung bổ sung. Điều này có thể thực hiện dưới dạng làm giàu thông tin cho các đoạn như này. Các đối tác thương mại có thể mong muốn nhắm đến các đoạn liên quan như các con đường để phân phối các dịch vụ sản phẩm.

Hơn nữa, các nhà tiếp thị đang tìm kiếm các phương pháp được cải thiện thu hút các khán giả trẻ tuổi hơn mà đã quen với việc bỏ qua quảng cáo và sử dụng truyền thông theo yêu cầu. Sự sắp đặt sản phẩm và giải trí có thương hiệu cung cấp khả năng “tiếp cận đa kênh” để thu hút người tiêu dùng trẻ hơn và/hoặc am hiểu công nghệ cao một cách hiệu quả hơn.

Điều này gây ra vấn đề cho nhà quảng cáo. Hiện nay không có cách nào để người xem thể hiện sự quan tâm đến đối tượng hoặc sản phẩm mà họ thấy trong buổi trình diễn hoặc trên màn ảnh. Ví dụ, người xem có thể nhìn thấy người nổi tiếng với túi xách. Tuy nhiên, nó có thể không rõ nơi mua túi xách. Người xem sẽ không tìm kiếm túi xách và sẽ dần dần mất quan tâm. Đã có các nỗ lực thực hiện việc liên kết video và/hoặc ấn phẩm in vào trang web.

Bằng cách sử dụng công nghệ hiện nay, nhà cung cấp hoặc nhà quảng cáo có thể bao gồm mã QR (Mã phản hồi nhanh) lên quảng cáo dạng in hoặc video. Người xem có thể sử dụng điện thoại thông minh để quét mã QR mà sẽ hướng họ đến trang web và/hoặc nội dung web. Tuy nhiên, điều này yêu cầu khối mã dễ thấy phải được đặt ở gần người xem. Hơn nữa, mã QR riêng phải được bao gồm cho từng đối tượng quan tâm. Đối với video, mã QR phải có mặt trong toàn bộ thời lượng của nó.

Theo đó, cần phải 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 các hình ảnh và/hoặc video.

Hệ thống phải cho phép người dùng có được các chi tiết và thông tin thêm về đối tượng quan tâm mà không cần thực hiện việc tìm kiếm từ khóa hoặc quét mã QR. Nó phải có khả năng sử dụng với phương tiện in ấn (ví dụ các quảng cáo ở tạp chí) cũng như phương tiện video (ví dụ tivi).

Sáng chế bao gồm phương pháp phát hiện đối tượng trong video và so khớp đối tượng với một hoặc nhiều sản phẩm bao gồm các bước (a) thu nhận video, (b) phân đoạn video dựa trên các cài đặt và/hoặc các sự kiện được mô tả bằng cách so sánh các nội dung của các khung hình liên tục về các điểm giống nhau và khác nhau, (c) biên dịch các phân đoạn của các cài đặt và/hoặc các sự kiện giống nhau hoặc tương tự nhau, (d) phân tích một hoặc nhiều khung hình của video để phát hiện một hoặc nhiều đối tượng từ mỗi phân đoạn của các cài đặt và/hoặc các sự kiện giống nhau hoặc tương tự nhau, (e) so sánh một hoặc nhiều đối tượng với các sản phẩm trong cơ sở dữ liệu, (f) định danh các sản phẩm có liên quan đến một hoặc nhiều đối tượng này và (g) thông báo cho một hoặc nhiều người xem sản phẩm.

Phương pháp này có thể sử dụng mạng nơ-ron tích chập (CNN-convolutional neural network) để định danh các sản phẩm có liên quan đến một hoặc nhiều đối tượng. Cơ sở dữ liệu có thể được điền bằng cách tăng cường dữ liệu ngoại tuyến (trình thu thập dữ liệu web) và/hoặc bằng cách căn chỉnh các đối tượng đã biết và các cụm siêu dữ liệu với nội dung đã được xác định. Sự tăng cường nội dung màn hình thứ hai có thể được sử dụng cho video trực tiếp hoặc video truyền tải nhanh. Bước thông báo cho một hoặc nhiều người xem hoặc các sản phẩm có thể bao gồm việc hiển thị quảng cáo và/hoặc thông báo cho một hoặc nhiều người xem hoặc các sản phẩm bằng cách cung cấp siêu liên kết đến trang web hoặc video.

Sáng chế cũng bao gồm phương pháp phát hiện một hoặc nhiều đối tượng trên ảnh chụp màn hình và so khớp một hoặc nhiều đối tượng với tài liệu quảng cáo bao gồm các bước (a) nhận sự truy vấn từ người xem dưới dạng hình ảnh kỹ thuật số hoặc ảnh chụp màn hình, (b) định danh một hoặc nhiều vật phẩm trên ảnh chụp màn hình, (c) so sánh một hoặc nhiều vật phẩm với các sản phẩm trong cơ sở dữ liệu, (d) định danh các sản phẩm có liên quan đến một hoặc nhiều vật phẩm này và (e) liên hệ người xem với tài liệu quảng cáo liên quan đến các sản phẩm được định danh.

Cơ sở dữ liệu có thể được điền với các sản phẩm bằng cách tăng cường dữ liệu ngoại tuyến (trình thu thập dữ liệu web) và/hoặc bằng cách căn chỉnh các đối tượng đã biết và các cụm siêu dữ liệu với nội dung đã được xác định. Sự tăng cường nội dung màn hình thứ hai có thể được sử dụng cho video trực tiếp hoặc video truyền tải nhanh. Bước liên hệ người xem với tài liệu quảng cáo có thể bao gồm việc hiển thị quảng cáo và/hoặc cung cấp đường dẫn siêu liên kết đến trang web hoặc video.

Sáng chế cũng bao gồm hệ thống tạo ra các mối quan hệ giữa các đối tượng trong video với các sản phẩm trong cơ sở dữ liệu và phân phối thông tin về các sản phẩm bao gồm (a) mạng và hệ thống được vi tính hóa để được trình bày cục bộ hoặc từ xa với người dùng hoặc các nhóm người dùng thông qua ứng dụng giao diện người dùng chẳng hạn như điện thoại di động, trình duyệt hoặc hệ thống được điều khiển bằng máy tính tương tự bất kỳ, (b) môđun để phát hiện và lưu trữ nội dung phương tiện truyền thông cục bộ hoặc trên máy chủ, (c) môđun để truyền tải nội dung phương tiện truyền thông đến bộ xử lý, điều khiển từ xa hoặc được dựa trên máy chủ, để nhập siêu dữ liệu và/hoặc các đặc trưng trực quan, (d) môđun để truyền tải nội dung phương tiện truyền thông đến bộ xử lý, điều khiển từ xa hoặc được dựa trên máy chủ, để trích xuất siêu dữ liệu và/hoặc các đặc trưng trực quan, (e) phương tiện để nhận đầu vào từ một hoặc nhiều người xem dưới dạng hình ảnh kỹ thuật số mà bao gồm các đặc trưng trực quan, (f) môđun được định cấu hình để định danh các đặc trưng trực quan và tương quan các đặc trưng trực quan với các đối tượng và/hoặc các nhóm sản phẩm liên quan và (g) dịch vụ mạng mà phân phối thông tin lên các đối tượng và/hoặc các nhóm sản phẩm liên quan cho người dùng và/hoặc các nhóm người dùng.

Mạng nơ-ron tích chập (CNN) có thể được sử dụng để phân tích các đặc trưng trực quan và siêu dữ liệu để tương quan các đặc trưng trực quan với các đối tượng và/hoặc các nhóm các sản phẩm liên quan. Cơ sở dữ liệu có thể được điền với các đối tượng đã biết bằng cách sử dụng quá trình tăng cường dữ liệu ngoại tuyến (trình thu thập dữ liệu web) và/hoặc bằng cách căn chỉnh các đối tượng đã biết và các cụm siêu dữ liệu với nội dung đã được xác định. Thông tin trên các sản phẩm liên quan bao gồm các quảng cáo và/hoặc siêu liên kết để truy cập nội dung thông qua internet.

Bản chất kỹ thuật của sáng chế

Khía cạnh thứ nhất của sáng chế là hệ thống tạo ra các mối quan hệ giữa các truy vấn trực quan của người dùng và các đối tượng được phát hiện từ trong kho ngữ liệu của các cơ sở dữ liệu của đối tượng.

Khía cạnh thứ hai của sáng chế là mạng và hệ thống được vi tính hóa cục bộ hoặc từ xa được trình bày với người dùng hoặc các nhóm người dùng thông qua ứng dụng giao diện người dùng chẳng hạn như điện thoại di động, trình duyệt hoặc hệ thống được điều khiển bằng máy tính tương tự bất kỳ.

Khía cạnh thứ ba của sáng chế là môđun được định cấu hình để phát hiện và lưu trữ nội dung phương tiện truyền thông cục bộ hoặc trên máy chủ.

Khía cạnh thứ tư của sáng chế là môđun được định cấu hình để truyền tải nội dung phương tiện truyền thông đến bộ xử lý, điều khiển từ xa hoặc được dựa trên máy chủ, cho việc nhập và xuất siêu dữ liệu có liên quan và/hoặc các đặc trưng trực quan.

Khía cạnh thứ năm của sáng chế là mô hình máy tính để phân tích các đặc trưng trực quan và siêu dữ liệu để tương quan với các đối tượng và/hoặc các nhóm các đối tượng cụ thể.

Khía cạnh thứ sáu của sáng chế là môđun được đặt máy chủ cục bộ hoặc ở phía máy chủ, được định cấu hình để liên kết các đối tượng được phát hiện đến các nhóm các đối tượng liên quan.

Khía cạnh thứ bảy của sáng chế là môđun được đặt máy chủ cục bộ hoặc ở phía máy chủ, được định cấu hình để căn chỉnh các đối tượng và các cụm siêu dữ liệu đã biết với nội dung đã được xác định trước.

Khía cạnh thứ tám của sáng chế là dịch vụ mạng mà phân phối nội dung đến người dùng và/hoặc các nhóm người dùng.

Mô tả vắn tắt các hình vẽ

Fig.1 mô tả quy trình tổng thể của một phương án của sáng chế.

Fig.2 mô tả sự phân đoạn cảnh trong video.

Fig.3 mô tả tổng quan về tìm kiếm trực quan.

Fig.4 mô tả quá trình tăng cường dữ liệu sản phẩm ngoại tuyến.

Fig.5 mô tả bộ khung đề xuất sản phẩm cho nội dung ngoại tuyến được nhập trước.

Fig.6 mô tả sự truy vấn của người dùng và đề xuất sản phẩm.

Mô tả chi tiết sáng chế

Các định nghĩa

Trong bản mô tả kỹ thuật này, khi đề cập đến "một phương án/khía cạnh" hoặc "phương án/khía cạnh" nghĩa là đặc trưng, cấu trúc, hoặc đặc tính 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. Cách sử dụng cụm từ "theo một phương án/khía cạnh" hoặc "theo phương án/khía cạnh khác" ở nhiều chỗ khác nhau trong bản mô tả kỹ thuật không nhất thiết nói tới phương án/khía cạnh giống nhau, cũng không phải là các phương án/khía cạnh riêng hoặc thay thế lẫn nhau mà không có các phương án/khía cạnh khác. Hơn nữa, các đặc trưng khác nhau được mô tả mà có thể được bộc lộ bởi một số các phương án/khía cạnh và không bộc lộ trong 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 các yêu cầu cho các phương án/khía cạnh khác. Trong các trường hợp nhất định, phương án và khía cạnh có thể được sử dụng thay thế cho nhau.

Các thuật ngữ được sử dụng trong bản mô tả kỹ thuật này thường có các nghĩa thông thường của chúng trong lĩnh vực kỹ thuật này, trong ngữ cảnh của sáng chế, và trong ngữ cảnh cụ thể mà mỗi thuật ngữ được sử dụng. Các thuật ngữ nhất định mà được sử dụng để mô tả sáng chế được nêu dưới đây, hoặc ở chỗ nào đó trong bản mô tả kỹ thuật, để cung cấp sự hướng dẫn bổ sung cho người thực hiện về bản mô tả của sáng chế. Để thuận tiện, các thuật ngữ nhất định có thể được làm nổi bật, ví dụ sử dụng các dấu ngoặc kép và/hoặc in nghiêng. Việc làm nổi bật không ảnh hưởng đến phạm vi và ý nghĩa của thuật ngữ; phạm vi và ý nghĩa của thuật ngữ là giống nhau, trong cùng ngữ cảnh, dù nó có được làm nổi bật hoặc không. Tốt hơn là cùng một điều có thể được nêu theo nhiều hơn một cách.

Kết quả là, ngôn ngữ thay thế và các từ đồng nghĩa có thể được sử dụng cho một hoặc nhiều thuật ngữ bất kỳ được nêu ở đây. Cũng không có bất kỳ ý nghĩa đặc biệt nào được đặt ra khi có thuật ngữ được trình bày rõ hoặc được nêu ở đây hoặc không. Các từ đồng nghĩa với các thuật ngữ nhất định được cung cấp. Việc nêu chi tiết một hoặc nhiều

từ đồng nghĩa không loại trừ việc sử dụng các từ đồng nghĩa khác. Việc sử dụng các ví dụ ở bất kỳ chỗ nào trong bản mô tả kỹ thuật này bao gồm các ví dụ về bất kỳ thuật ngữ nào được nêu ở đây chỉ mang tính chất minh họa, và không có ý định giới hạn hơn phạm vi và ý nghĩa của sáng chế hoặc bất kỳ thuật ngữ được minh họa bằng ví dụ. Cũng như vậy, sáng chế không bị giới hạn ở các phương án khác nhau được nêu trong bản mô tả kỹ thuật này.

Không có ý định giới hạn thêm phạm vi của sáng chế, các ví dụ về các dụng cụ, các thiết bị, phương pháp và các kết quả liên quan của chúng theo các phương án của sáng chế được nêu dưới đây. Lưu ý rằng các tiêu đề hoặc phụ đề có thể được sử dụng trong các ví dụ để thuận tiện cho người đọc, mà không 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 ý nghĩa khi được hiểu một cách thông thường bởi những người có trình độ bình thường trong lĩnh vực kỹ thuật này mà sáng chế liên quan đến. Trong trường hợp mâu thuẫn, tài liệu này, bao gồm các định nghĩa, sẽ kiểm soát.

Thuật ngữ “app” hoặc “ứng dụng” là chỉ chương trình độc lập hoặc phần mềm được thiết kế để thực hiện mục đích cụ thể, đặc biệt là được tải về thiết bị điện thoại di động.

Thuật ngữ “túi từ” hoặc “mô hình túi từ” đề cập đến sự phân loại hình ảnh, bằng cách xử lý các đặc trưng hình ảnh như các từ ngữ. Trong phân loại tài liệu, túi từ là véc-tơ thưa đếm số lần xuất hiện của các từ ngữ; nghĩa là, biểu đồ thưa trên từ vựng. Trong thị giác máy tính, túi từ trực quan là véc-tơ đếm số lần xuất hiện của từ vựng về các đặc trưng hình ảnh cục bộ.

Thuật ngữ “tập lưu trữ thông tin tùy chọn cá nhân,” “tập lưu trữ thông tin tùy chọn cá nhân internet” hoặc “tập lưu trữ thông tin tùy chọn cá nhân HTTP” đề cập đến phần nhỏ dữ liệu được gửi từ trang web và được lưu trữ trên máy tính của người dùng bởi trình duyệt web của người dùng. Các tập lưu trữ thông tin tùy chọn cá nhân được gửi qua lại giữa máy chủ internet và trình duyệt mà cho phép người dùng được định danh hoặc để theo dõi tiến trình của họ. Các tập lưu trữ thông tin tùy chọn cá nhân cung cấp chi tiết về các trang khách hàng truy cập, lượng thời gian dành để xem mỗi trang, các liên kết đã nhấn vào, các tìm kiếm và các sự tương tác được thực hiện. Từ các thông tin này, nhà phát hành tập lưu trữ thông tin tùy chọn cá nhân tập hợp sự hiểu biết về các

sở thích và xu hướng duyệt của người dùng để tạo ra hồ sơ. Phân tích hồ sơ, các nhà quảng cáo có thể tạo ra các phân khúc khán giả được xác định dựa trên những người dùng tương tự được trả về thông tin tương tự, tạo nên các hồ sơ.

Thuật ngữ “tạo cụm” hoặc “phân tích cụm” đề cập đến tác vụ nhóm tập hợp các đối tượng theo cách mà các đối tượng trong cùng nhóm (được gọi là cụm) tương đồng với nhau hơn (theo một vài ý nghĩa hoặc khác) so với các đối tượng trong các nhóm khác (các cụm). Nó là tác vụ chính của việc khai thác dữ liệu thăm dò, và kỹ thuật thông thường để phân tích dữ liệu thống kê, được sử dụng trong nhiều lĩnh vực, bao gồm học máy, nhận dạng mẫu hình, phân tích hình ảnh, truy hồi thông tin, tin sinh học, nén dữ liệu, và đồ họa máy tính.

Thuật ngữ “tăng cường dữ liệu” đề cập đến việc gia tăng số lượng điểm dữ liệu. Xét về hình ảnh, điều đó có nghĩa là làm tăng số lượng hình ảnh trong bộ dữ liệu. Xét về dữ liệu dạng hàng/cột truyền thống, nó có nghĩa là làm tăng số lượng hàng hoặc các đối tượng.

Thuật ngữ “học sâu” đề cập đến ứng dụng cho các tác vụ học của mạng nơ-ron nhân tạo (ANNs-artificial neural networks) mà chứa nhiều hơn một lớp ẩn. Học sâu là phần của họ lớn hơn của các phương pháp học máy dựa trên việc học các biểu diễn dữ liệu, trái ngược với tác vụ thuật toán cụ thể.

Thuật ngữ “véc-tơ đặc trưng,” trong sự nhận dạng mẫu hình và học máy, đề cập đến véc-tơ đặc trưng là véc-tơ n chiều của các đặc trưng dạng số mà biểu diễn một số đối tượng. Nhiều thuật toán trong học máy yêu cầu sự biểu diễn dạng số của các đối tượng, vì các biểu diễn như này dễ dàng cho việc xử lý và phân tích thống kê. Khi biểu diễn các hình ảnh, các giá trị đặc trưng có thể tương ứng với các điểm ảnh của hình ảnh, khi biểu diễn các văn bản có thể là tần suất xuất hiện cụm từ.

Thuật ngữ “tập dữ liệu không cân bằng” đề cập đến trường hợp đặc biệt cho vấn đề phân loại mà trong đó sự phân bố lớp là không đồng nhất giữa các lớp. Thông thường, chúng bao gồm hai lớp: lớp đa số (âm) và lớp thiểu số (dương). Việc cân bằng lớp có thể cần thiết để đặt dữ liệu dưới dạng có thể sử dụng được.

Thuật ngữ “chỉ mục ngược,” “tệp đăng” hoặc “tệp tin ngược” là cấu trúc dữ liệu chỉ mục lưu trữ ánh xạ từ nội dung, chẳng hạn như các từ hoặc các số, vào các vị trí của nó trong tệp tin cơ sở dữ liệu, hoặc trong tài liệu hoặc bộ tài liệu (được đặt tên trái ngược

với chỉ mục thuận, mà ánh xạ từ các tài liệu vào nội dung). Mục đích của chỉ mục ngược là để cho phép các tìm văn bản đầy đủ nhanh chóng, với chi phí của việc xử lý được gia tăng khi một tài liệu được bổ sung vào cơ sở dữ liệu.

Thuật ngữ “k-láng giềng gần nhất” hoặc “k-NN” đề cập đến đối tượng phân loại láng giềng gần nhất, mà cả số đo khoảng cách ("gần nhất") và số lượng láng giềng 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 huấn luyện, vì vậy có thể tính toán các dự đoán thay thế lại.

Phân tích liên kết

Thuật ngữ “môđun” đề cập đến đơn vị độc lập, chẳng hạn như cụm linh kiện điện tử và nối dây liên kết hoặc phân đoạn phần mềm máy tính, mà tự thực hiện tác vụ xác định và có thể được liên kết với các đơn vị khác thành để tạo thành hệ thống lớn hơn.

Thuật ngữ “Mạng nơ-ron truyền thẳng nhiều lớp” hoặc “MLP” đề cập đến Mạng nơ-ron truyền thẳng với một hoặc nhiều lớp giữa các lớp đầu vào và đầu ra. Truyền thẳng nghĩa là dữ liệu chảy theo một hướng từ đầu vào đến lớp đầu ra (hướng về phía trước). Các MLP được sử dụng rộng rãi để phân loại mẫu hình, nhận dạng, dự đoán và tính xấp xỉ. Truyền thẳng nhiều lớp có thể giải quyết các vấn đề mà không thể tách rời tuyến tính.

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

Thuật ngữ “siêu thể” đề cập đến siêu dữ liệu có trên các trang web. Các siêu thể mô tả và các từ khóa thường được sử dụng để mô tả nội dung của trang Web. Hầu hết các công cụ tìm kiếm sử dụng dữ liệu này khi bổ sung các 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) mà chứa thông tin về vật phẩm mà nó được gắn vào. Mã QR bao gồm các hình vuông màu đen được sắp xếp thành lưới ô vuông trên nền trắng, mà có thể

được đọc bởi thiết bị hình ảnh chẳng hạn như máy ghi hình, và được xử lý bằng cách sử dụng kỹ thuật sửa lỗi Reed–Solomon cho đến khi hình ảnh có thể được diễn dịch hợp lý. Sau đó, dữ liệu được yêu cầu được trích xuất từ các mẫu hình mà có mặt trong cả các thành phần theo hướng ngang và hướng dọc của hình ảnh.

Thuật ngữ “dữ liệu tổng hợp” đề cập đến bất kỳ dữ liệu sản phẩm nào có thể áp dụng được vào các tình huống được đưa ra mà không thu được bằng phép đo trực tiếp.

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

Thuật ngữ “quảng cáo có mục tiêu” đề cập đến dạng quảng cáo mà các nhà quảng cáo trực tuyến có thể sử dụng các phương pháp phức tạp để nhắm mục tiêu đến các khán giả dễ tiếp thu nhất với những đặc điểm nhất định, dựa trên sản phẩm hoặc cá nhân mà nhà quảng cáo đang quảng bá. Những đặc điểm này có thể là đối tượng khách hàng mà được tập trung vào chủng tộc, trạng thái kinh tế, giới tính, độ tuổi, trình độ học, mức thu nhập và nghề nghiệp hoặc chúng có thể được tập trung vào biểu đồ tâm lý mà dựa trên giá trị, tính cách, thái độ, các quan điểm, phong cách sống và sở thích của khách hàng. 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 “các từ trực quan” hoặc “các cụm từ trực quan” như được sử dụng trong các hệ thống truy hồi hình ảnh, đề cập đến các phần nhỏ của hình ảnh mà mang một số loại thông tin liên quan đến các đặc trưng (chẳng hạn như màu sắc, hình dạng hoặc kết cấu bề mặt), hoặc thay đổi xảy ra trong các điểm ảnh chẳng hạn như lọc, các bộ mô tả đặc trưng cấp thấp (SIFT, SURF, ...v.v.).

Thuật ngữ “sự biến đổi làm trắng” hoặc “sự biến đổi hình cầu” đề cập đến sự biến đổi tuyến tính mà biến đổi véc-tơ của các biến ngẫu nhiên với ma trận hiệp phương sai đã biết thành tập hợp các biến mới mà hiệp phương sai của nó là ma trận đơn vị nghĩa là chúng không tương quan với nhau và tất cả có phương sai là 1. Sự biến đổi được gọi là "làm trắng" bởi vì nó thay đổi véc-tơ đầu vào thành véc-tơ nhiễu trắng.

Các thuật ngữ kỹ thuật khác được sử dụng ở đây có ý nghĩa thông thường của nó trong lĩnh vực kỹ thuật mà chúng được sử dụng, như được minh họa bằng ví dụ bởi nhiều từ điển kỹ thuật.

Các phương án ưu tiên

Các cấu hình và giá trị cụ thể được nêu các ví dụ không giới hạn này có thể được thay đổi và chỉ được trích dẫn để minh họa ít nhất một phương án và không nhằm giới hạn phạm vi của nó.

Một trong những thuận lợi lớn nhất của tìm kiếm trực quan (trái ngược với tìm kiếm dựa trên văn bản theo kiểu truyền thống) cho việc lập hồ sơ đối tượng khách hàng là thông tin vốn có lớn hơn về truy vấn có thể được xác định chắc chắn. Ví dụ, người dùng có thể tìm kiếm giấy màu nâu trong công cụ tìm kiếm (hoặc trang web thương mại điện tử). người dùng có thể sau đó chọn mua hoặc hỏi về loại giấy màu nâu rất cụ thể (giấy lười, giấy có dây, v.v.).

Chỉ với truy cập chỉ vào truy vấn tìm kiếm văn bản, thì không thể trích xuất chi tiết hơn về đối tượng tìm kiếm mà không có thêm thông tin. Tuy nhiên, với trường hợp sử dụng tìm kiếm trực quan, truy vấn hình ảnh tự có thể nói với chúng ta nhiều điều hơn về bản chất truy vấn của người dùng.

Để trích xuất siêu dữ liệu về truy vấn tìm kiếm trực quan, thì có thể sử dụng các thuật toán phân loại nâng cao, bao gồm nhưng không bị giới hạn ở Học sâu, Học được giám sát và Học không được giám sát. Do đó, từ hình ảnh đầu vào, có thể thu được danh sách siêu dữ liệu mô tả (ví dụ giấy, màu nâu, có dây, giấy tây, bối cảnh, địa điểm sản xuất, vật liệu và thông tin bất kỳ mà cung cấp rõ ràng về trạng thái các nội dung nằm trong hình ảnh).

Theo một phương án của sáng chế được mô tả ở đây, danh sách các đối tượng mà bao gồm hình ảnh có thể được trích xuất từ khung hình, được liên kết với chuỗi khung hình mà được phân tích tương ứng với “chủ đề” khác biệt về mặt ngữ nghĩa.

Fig.1 mô tả dòng quy trình tổng thể của sáng chế. Nội dung từ video 110 được thu thập và được biên dịch để xây dựng cơ sở dữ liệu nội dung 170. Người xem video có thể truy cập cơ sở dữ liệu nội dung 170 thông qua truy vấn 120 bằng cách, ví dụ, gửi ảnh chụp màn hình từ video lên hệ thống.

Tập tin video 110 được xử lý để tự động xác định và trích xuất siêu dữ liệu và các thuộc tính liên quan đến các đối tượng nằm trong khung hình, hoặc nhóm khung hình tương tự về mặt ngữ nghĩa. Phân tích thời gian (được mô tả dưới đây) bao gồm sự phân đoạn 140 của video. Theo công cụ phát hiện đối tượng 160, khung hình chính có thể được phân tích cho các đối tượng. Bản ghi chèn 180 được tạo ra để giải thích vị trí thời gian của các đối tượng. Các đối tượng được định danh và thông tin định danh được bổ sung vào cơ sở dữ liệu nội dung 170.

Người dùng có thể truy vấn khung hình 120 bằng cách lấy hình ảnh của video nhờ sử dụng Giao diện ảo 130. Hệ thống bao gồm công cụ tìm kiếm trực quan 150 mà có thể truy cập cơ sở dữ liệu nội dung 170.

Phân tích thời gian

Các bước liên quan đến việc phân tích thời gian 200 được mô tả trên Fig.2. Video 110 được đưa vào trong bộ khung và chuỗi khung thời gian được phân tích để phân đoạn nội dung phương tiện. Mục đích của sự phân đoạn này là để định danh, phân lập và gắn thẻ các chuỗi khung sao cho mỗi phân đoạn tương ứng với chủ đề hoặc sự kiện đơn lẻ. Sau đó, các phân đoạn này có thể được phân tích cho các đối tượng.

Video 110 được duyệt tuần tự 210 để phát hiện cặp, hoặc chuỗi, khung hình mà vi phạm ngưỡng tương tự. Nghĩa là, sự so sánh khung hình 220 có thể cho thấy sự thay đổi đáng kể trong thành phần khung hình, ngụ ý sự thay đổi của cảnh hoặc chủ đề.

Các khung hình mà mô tả các sự kiện/các cảnh giống nhau sẽ không vi phạm ngưỡng tương tự. Trong trường hợp này, cặp khung kế tiếp sẽ được đánh giá 250. Khung hình mà mô tả các sự kiện/các cảnh khác nhau thường sẽ vi phạm ngưỡng tương tự, trong trường hợp này, phân đoạn sẽ được định danh 240 làm sự kiện/cảnh. Quá trình này có thể được lặp lại sao cho mỗi khung trong video 110 được bao gồm trong phân đoạn. Sau đó, công cụ phát hiện đối tượng 160 có thể phân tích các phân đoạn video cho các đối tượng.

Tầm quan trọng của sự phân đoạn được tăng lên khi xử lý các đối tượng mà có thể thể hiện tính đa hình theo thời gian (nghĩa là thay đổi hình dạng của chúng qua các khung hình). Sau khi hoàn tất sự định danh của các cảnh hoặc chuỗi khung hình trong video, có thể liên kết với các đối tượng mà có thể biến đổi thành các hình dạng không xác định từ khung hình chính được phát hiện. Phương án thay thế là để huấn luyện mô

hình phát hiện đối tượng với các ví dụ về tất cả các sự biến dạng hình dạng đã biết cho đối tượng cụ thể.

Siêu dữ liệu có thể được liên kết với nội dung trực quan của chính khung hình bằng mã định danh phân đoạn và dữ liệu có thể được nhập vào trong cơ sở dữ liệu tìm kiếm trực quan song song. Khi người dùng truy vấn hình ảnh của khung hình đã nhập, nó có thể được gửi đến máy chủ so khớp. Mã định danh phân đoạn có thể được sử dụng để định danh chuỗi khung hình mà khung hình này có liên quan đến. Thông tin này được sử dụng để truy hồi danh sách các đối tượng và bất kỳ nội dung tăng cường được liên kết nào cho các đối tượng nêu trên để được trả về người dùng 120. Lưu ý là bản thân truy vấn khung hình không được phân tích cho các đối tượng. Thay vào đó, nó được so khớp với mã định danh phân đoạn mà liên kết với kho ngữ liệu của đối tượng được phân tích trước cho phân đoạn.

Sự phát hiện đối tượng

Khung hình chính được định danh thông qua phân tích thời gian có thể được phân tích để nhận dạng và định danh các đối tượng đơn lẻ. Theo một phương án, mạng nơ-ron tích chập sâu (CNN) bao gồm nhiều lớp có thể được sử dụng để thực hiện tác vụ này.

Fig.3 mô tả cách công cụ tìm kiếm trực quan dựa trên các tiếp cận “túi từ trực quan” có thể được sử dụng để tìm kiếm cho các mã định danh phân đoạn từ hình ảnh truy vấn trực quan. Công cụ tìm kiếm sử dụng cơ sở dữ liệu hình ảnh 385. Các hình ảnh huấn luyện 310 (nghĩa là các hình ảnh với các đặc trưng đã biết) được gửi cùng với các đối tượng được định ra. Các hình ảnh được sử dụng để tạo ra các đặc trưng 320. Các hình ảnh huấn luyện 365 có thể được đưa vào trong cơ sở dữ liệu 375.

Để huấn luyện mô hình CNN, cần phải đưa vào số lượng hình ảnh đáng kể các ví dụ hình ảnh cho mỗi đối tượng được nhập vào. Kho ngữ liệu lớn của dữ liệu hình ảnh có thể được tổ chức cho nhiều đối tượng với nhiều thuộc tính và vị trí của chúng. Dữ liệu này còn được trải qua biến đổi làm trắng, tăng cường dữ liệu và cân bằng lớp. Cơ sở dữ liệu này của các đối tượng được sử dụng như đầu vào để huấn luyện mạng lưới tích chập sâu.

Mô hình có thể được huấn luyện sao cho các đối tượng và các thuộc tính của chúng được học với nhau. Hơn nữa, mô hình cung cấp các giá trị từ các lớp ẩn mà là sự

mô tả véc-tơ có giá trị thực của các hình ảnh qua các sự trừu tượng về mặt ngữ nghĩa khác nhau. Suy luận của mô hình có thể cung cấp các nhãn đối tượng, điểm tin cậy, các véc-tơ lớp ẩn và các thuộc tính.

Ví dụ, các đối tượng có thể là vật phẩm trong miền hàng may mặc (túi, quần bò, v.v..) cùng với các thuộc tính của chúng (màu sắc, hoa văn, chiều dài v.v..). Ngoài ra, suy luận về hình ảnh trên nhiều khu vực vị trí của đối tượng và có thể được sử dụng để xác định vị trí có khả năng nhất của hình ảnh trong khung hình trước hoặc sau mỗi khung hình chính.

Một khi mô hình được đánh giá để đạt tới trạng thái được huấn luyện thành công (được đánh giá dựa trên việc đáp ứng các chỉ số lỗi tối thiểu cho một tập hợp thử nghiệm), nó được truyền đến môđun nhập trực tiếp trong bộ khung. Sau khi phân tích thời gian của phân đoạn video, mỗi khung hình chính được phân tích để thu được vị trí của các đối tượng có mặt trong hình ảnh. Mọi đối tượng được phát hiện trong khung hình có thể được theo dõi theo thời gian theo cả hai hướng trong phân đoạn. Đối với mỗi vị trí được tìm thấy trong phân đoạn, sự suy luận về các nhãn, các thuộc tính và các lớp bên trong có thể được thực hiện và được tính theo thời gian trung bình có trọng số bởi phép đo tin cậy. Điều này lên đến cực điểm trong việc tạo ra véc-tơ đặc trưng theo đối tượng cho mỗi phân đoạn và/hoặc khung hình. Các mô tả véc-tơ được tạo ra bởi các lớp ẩn có thể thường được bao gồm chiều cao. Số lượng lớn các hình ảnh có thể được thu thập để học kỹ thuật nén cho sự phân phối cụ thể đó. Theo phương án ví dụ về bộ khung được mô tả ở đây, bộ mã hóa tự động sâu cung cấp độ nén tốt nhất với ít sự hao tổn nhất đối với độ chính xác tìm kiếm.

Tìm kiếm trực quan

Khi người dùng muốn tương tác với video cụ thể, họ có thể cho thấy sự quan tâm của mình bằng cách chụp khung hình với máy ghi hình của thiết bị điện thoại di động của họ, từ trong chương trình máy tính thường hoạt động trên thiết bị điện thoại di động của người dùng, được gọi là “Ứng dụng.” Ứng dụng này có thể tải lên hình ảnh truy vấn để xử lý tiếp.

Sau khi hình ảnh truy vấn trực quan được gửi đến máy chủ từ thiết bị của người dùng, nó được sử dụng để tìm kiếm thông qua cơ sở dữ liệu hình ảnh đã biết để định danh các kết quả so khớp có thể có. Sau đó, kết quả xếp hạng đầu được sử dụng để định

đánh tiếp phân đoạn nào mà khung hình trong câu hỏi có liên quan nhiều nhất. Sau đó, mã định danh phân đoạn có thể được sử dụng để truy hồi các đối tượng được liên kết với phân đoạn cụ thể đó. Sau khi phản hồi truy vấn được tổng hợp, tất cả nội dung được tăng cường được chuyển tiếp trở lại về thiết bị điện thoại di động của người dùng.

Ví dụ, người xem có thể gửi truy vấn về túi xách mà họ đã để ý thấy trong video. Túi xách có thể được so khớp với các hình ảnh trong cơ sở dữ liệu dựa trên một số tiêu chí, bao gồm hình dạng, hoa văn, nhãn hiệu, hình dạng, kích thước, nhãn hiệu và các chi tiết khác. Hệ thống có thể trả về số lượng các hình ảnh phù hợp được xếp hạng.

Quá trình này được mô tả trên Fig.3 mà người dùng truy vấn hình ảnh 120. Các cụm từ trực quan được gán 340. Từ các cụm từ trực quan, danh sách tệp tin ngược được điền. Danh sách tệp tin ngược được truy vấn 350 và được lưu trữ 355. Các hình ảnh tiêu biểu đứng đầu được lọc 360 mà có thể được bổ sung vào cơ sở dữ liệu hình ảnh 385. Sự xác minh về không gian có thể được thực hiện cho các kết quả so khớp đứng đầu 380 và các kết quả so khớp đứng đầu có thể được trả về cho người xem 390.

Tăng cường dữ liệu ngoại tuyến

Các hình ảnh mà không được định rõ, chẳng hạn như những hình ảnh thu được bằng trình thu thập dữ liệu web cũng có thể được sử dụng để điền cơ sở dữ liệu nội dung. Fig.4 mô tả sự tăng cường dữ liệu sản phẩm ngoại tuyến 400. Trong bộ khung ví dụ, sự tăng cường dữ liệu ngoại tuyến được sử dụng để điền cơ sở dữ liệu của các đối tượng và siêu dữ liệu được kết hợp 170.

Quá trình thu thập dữ liệu có thể được sử dụng để truy hồi các hình ảnh và siêu dữ liệu được chú giải của chúng từ nhiều nguồn trực tuyến (ví dụ các danh sách sản phẩm từ các nền tảng thương mại điện tử hoặc các hình ảnh từ các mạng truyền thông xã hội) 420. Các thuộc tính hình ảnh và siêu dữ liệu được thu thập này được đi qua giai đoạn làm sạch dữ liệu 430 mà biến đổi dữ liệu được thu thập, thô thành các bản ghi mà thích hợp cho việc đưa vào trong cơ sở dữ liệu 440.

Định dạng đưa cơ sở dữ liệu có thể liên kết mỗi bản ghi hình ảnh được thu thập với mã định danh đối tượng. Điều này cho phép liên kết các đối tượng được phát hiện từ các phân đoạn đến các đối tượng trong cơ sở dữ liệu sản phẩm, do đó cung cấp giao diện để cung cấp dữ liệu được tăng cường cho đối tượng, hoặc nhóm các đối tượng được định danh từ hình ảnh truy vấn trực quan.

Các trường hợp sử dụng

Đề xuất sản phẩm

Sáng chế có thể được sử dụng để nâng cao dịch vụ tăng cường nội dung màn hình thứ hai cho nội dung phương tiện truyền thông trực quan hiện có.

Ví dụ, các chương trình ti vi (hoặc phim ảnh) phổ biến có thể được nhập vào trong nền tảng này để phân tích danh mục sản phẩm có thể truy cập được thông qua từng phương tiện truyền thông. Người xem có thể được quảng bá, hoặc nhận biết về các cơ chế tương tác có thể có diễn ra trong lúc, hoặc sau khi, chương trình được phát sóng (hoặc được truyền tải nhanh trong trường hợp video được truyền trực tuyến).

Bất kỳ truy vấn trực quan nào của người dùng đối với các khung hình từ được nhập các video có thể được làm giàu sau đó bằng bộ khung phát hiện đối tượng. Điều này cung cấp nền tảng duy nhất cho các nhà sáng tạo nội dung cũng như người tiêu dùng tương tác với các sản phẩm/dịch vụ được cung cấp. Phạm vi của các đối tượng phát hiện trong bộ khung có thể bao gồm vật phẩm vô tri vô giác trong khung hình (vật phẩm quần áo, đồ đạc, các cơ hội du lịch v.v.), hoặc được mở rộng đến các thực thể được phát hiện hoặc tương ứng với các thực thể trong các phân đoạn (nghĩa là các diễn viên, dàn diễn viên v.v.).

Ví dụ, nhà quảng cáo hoặc nhà bán lẻ có thể sử dụng vị trí sản phẩm để quảng bá túi xách trong video hoặc phim. Người xem gửi truy vấn về cảnh mà bao gồm túi xách. Quảng cáo khuyến mãi cho sản phẩm được so khớp (nghĩa là túi xách) có thể được phát trên thiết bị của người dùng. Thông tin bổ sung, bao gồm các hướng dẫn mua cũng có thể được cung cấp đến người xem.

Fig.5 mô tả bộ khung ví dụ để giải quyết trường hợp sử dụng được mô tả ở đây. Video 110 có thể được sử dụng để tăng cường cơ sở dữ liệu nội dung thông qua việc nhập ngoại tuyến. Video trải qua sự phát hiện sự thay đổi cảnh quay 510. Các phân đoạn video 540 được định danh bằng các khung hình chính 560, định vị đối tượng 590 và theo dõi 630. Bước tiếp theo là nhận dạng đối tượng, gán nhãn các thuộc tính và trích xuất đặc trưng 660 sau đó lấy bình quân thời gian 690.

Việc nhập video có thể cũng bao gồm việc lấy mẫu khung hình 550 và chức năng xử lý hình ảnh 570. Nội dung cơ sở dữ liệu có thể bao gồm các khung hình 610, các sản

phẩm 670 và các đối tượng 710. Việc nhập sản phẩm (thu thập dữ liệu web, tổng hợp và tăng cường) xảy ra tại 640.

Người dùng có thể truy vấn khung hình 120. Hình ảnh được so khớp bằng cách xếp hạng nó với các hình ảnh trong cơ sở dữ liệu mà tương tự nhau 580. Số phân đoạn có thể được tạo ra 620 cũng như việc xếp hạng tìm kiếm trực quan 650. Các sản phẩm được so khớp 680 có thể được chuyển tiếp đến người dùng để phản hồi lại truy vấn của họ. Các sản phẩm mà các nhà quảng cáo và/hoặc các nhà tiếp thị mong muốn quảng bá, có thể nhận được xếp hạng tìm kiếm trực quan cao hơn 650.

Màn hình thứ hai của ti vi phát trực tiếp

Ngoài nội dung ngoại tuyến, có sẵn, cũng có thể mở rộng bộ khung này để giải quyết các luồng phát video trực tiếp. Thử thách trong bối cảnh như này là để đảm bảo mỗi khung hình mà được truyền vào trong nền tảng phát hiện đối tượng sẽ hoàn thành cơ chế nhập trước bất kỳ truy vấn nào của cùng khung hình.

Để giải thích cho video trực tiếp, mô đun phân tích thời gian có thể được thay thế bằng cơ sở dữ liệu nội dung giao dịch mà duy trì lịch sử thời gian của “N” phút cuối cùng (hoặc là giờ nếu cần). Trong bộ khung được biến đổi này, mọi khung hình đến từ luồng video trực tiếp được nhập vào trong các cơ sở dữ liệu phát hiện đối tượng và tìm kiếm trực quan, với cơ chế “thời gian trực tiếp” đảm bảo rằng dữ liệu sẽ hết hạn sau khi độ trễ có thể định cấu hình. Theo cách này, kích cỡ của cơ sở dữ liệu (và cụm máy tính) được kết nối với trạng thái hiệu năng cao có khả năng duy trì các hoạt động độ trễ thấp về khối lượng.

Bộ khung như này có thể được sử dụng để đóng vai trò là nội dung màn hình thứ hai cho:

- các sự kiện bóng đá trực tiếp, hiển thị số liệu thống kê về sự nghiệp và/hoặc trò chơi của vận động viên
- phát sóng bản tin trực tiếp, hiển thị các đồ họa thông tin cho các đối tượng /các vị trí được phát hiện trong các bản tin v.v.
- phát sóng tiếp thị qua điện thoại, hiển thị các sự so sánh giá cho các sản phẩm được phát hiện.

Lợi ích của sáng chế cho việc mua sắm trực tuyến

Như được mô tả trên Fig.6, sáng chế cho phép người xem video 110 để truy vấn ảnh chụp màn hình của video để biết thêm thông tin về sản phẩm mà họ đề ý. Trong ví dụ này 600, người xem đang xem phim truyền hình trên ti vi. Người xem nhận thấy thấy diễn viên đang mặc áo đặc biệt. Người xem có thể chụp ảnh màn hình để gửi vào hệ thống 120. Đối với chương trình trên ti vi, họ có thể chụp ảnh màn hình bằng cách sử dụng ứng dụng. Nếu người xem đang phát trực tuyến video trên điện thoại, máy tính bảng hoặc máy tính, họ có thể chụp ảnh màn hình để gửi đi. Các hình ảnh của ảnh tĩnh (ví dụ các quảng cáo tạp chí) cũng có thể được gửi đi.

Hệ thống phát hiện các đối tượng trên ảnh chụp màn hình. Trong trường hợp này, hệ thống phát hiện áo phong ngắn tay tối màu. Các sản phẩm đang bán sẵn trên thị trường mà phù hợp với tiêu chí này được hiển thị đến người xem 610. Họ có thể gửi tiêu chí bổ sung để làm cho truy vấn cụ thể hơn. Ví dụ, việc tìm kiếm đối tượng có thể được thu hẹp hơn nữa để chỉ bao gồm thiết kế cụ thể (ví dụ cổ chữ v) hoặc các áo phong từ nhà thiết kế cụ thể.

Sau đó, người xem có thể mua sản phẩm trực tuyến 620 thông qua nhà cung cấp cụ thể. Điều này cho phép các nhà cung cấp tiếp thị các sản phẩm thông qua việc đặt sản phẩm mà không yêu cầu thời gian và/hoặc các quảng cáo thương mại bổ sung.

Phần mô tả ở trên chỉ bộc lộ các phương án ví dụ về sáng chế. Các biến thể của các phương pháp và thiết bị được bộc lộ ở trên mà nằm trong phạm vi của sáng chế sẽ dễ dàng được nhận thấy đối với những người có hiểu biết trung bình trong lĩnh vực kỹ thuật này. Theo đó, mặc dù sáng chế được bộc lộ liên quan đến các phương án được lấy làm ví dụ của nó, nhưng cần phải hiểu rằng các phương án khác có thể nằm trong bản chất và phạm vi của sáng chế như được định rõ bởi phần yêu cầu bảo hộ dưới đây.

Môi trường điều hành:

Hệ thống thường bao gồm máy chủ trung tâm mà được kết nối bằng 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 đến một hoặc nhiều thiết bị lưu trữ chung. 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 kiểu máy tính xách tay hoặc máy tính để bàn cá nhân. Nó cũng có thể điện thoại di động, điện thoại thông minh hoặc thiết bị cầm tay khác, bao gồm máy tính bảng. Yếu tố hình thức chính xác của máy tính của người dùng không làm

giới hạn sáng chế được yêu cầu bảo hộ. Các ví dụ về các hệ thống máy tính, các môi trường, và/hoặc các cấu hình thông dụng mà có thể thích hợp để sử dụng với sáng chế bao gồm, nhưng không giới hạn ở, máy tính cá nhân, máy tính dạng máy chủ, máy cầm tay, máy tính xách tay hoặc máy tính dạng điện thoại di động hoặc các thiết bị giao tiếp chẳng hạn như điện thoại di động và thiết bị hỗ trợ cá nhân kỹ thuật số (PDA- personal digital assistant), các hệ thống đa xử lý, các hệ thống được dựa trên bộ vi xử lý, thiết bị giải mã tín hiệu, các thiết bị điện tử tiêu dùng có thể lập trình được, các máy tính cá nhân (PCs- personal computers) nối mạng, máy vi tính, các máy tính lớn, các môi trường điện toán được phân tán mà bao gồm thiết bị hoặc hệ thống bất kỳ nêu trên và thiết bị 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 bị bỏ qua, và thay vào đó là chức năng máy tính riêng biệt miễn là nó hoạt động với máy chủ trung tâm. Trong trường hợp này, người dùng sẽ đăng nhập vào máy chủ từ máy tính khác và truy cập hệ thống thông qua môi trường người dùng.

Môi trường người dùng có thể được chứa trong máy chủ trung tâm hoặc được kết nối hoạt động vào máy chủ trung tâm. Hơn nữa, người dùng có thể nhận từ và truyền tải dữ liệu đến máy chủ trung tâm bằng Internet, theo đó người dùng truy cập vào tài khoản sử dụng trình duyệt Internet web và trình duyệt hiển thị trang web tương tác được kết nối hoạt động đến máy chủ trung tâm. Máy chủ trung tâm truyền và nhận dữ liệu theo sự phản hồi lại dữ liệu và các lệnh được truyền từ trình duyệt theo sự phản hồi lại hành động của khách hàng đối với giao diện sử dụng trình duyệt. Một số bước của sáng chế có thể được thực hiện trên máy tính của người dùng và các kết quả tạm thời được truyền đến máy chủ. Các kết quả tạm thời này có thể được xử lý tại máy chủ và các kết quả cuối cùng được trả về 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- central processing unit) mà được kết nối hoạt động đến thiết bị nhớ, mạch đầu vào và đầu ra dữ liệu (I/O) và mạch giao tiếp dữ liệu mạng máy tính. Mã máy tính được thực thi bởi CPU có thể lấy dữ liệu được nhận bởi mạch giao tiếp dữ liệu và lưu trữ nó trong thiết bị nhớ. Ngoài ra, CPU có thể lấy dữ liệu từ mạch I/O và lưu trữ nó trong thiết bị nhớ. Hơn nữa, CPU có thể lấy dữ liệu từ thiết bị nhớ và xuất dữ liệu thông qua mạch I/O hoặc mạch giao tiếp dữ liệu. Dữ liệu được chứa trong bộ nhớ có thể còn được gọi lại từ thiết bị nhớ, còn được xử lý hoặc được biến đổi

bằng CPU theo cách được mô tả ở đây và được khôi phục trong cùng một thiết bị nhớ hoặc thiết bị nhớ khác được kết nối hoạt động đến CPU kể cả bằng mạch mạng dữ liệu. Thiết bị nhớ có thể là loại mạch lưu trữ dữ liệu hoặc ổ lưu trữ từ tính hoặc thiết bị quang, bao gồm ổ cứng, ổ quang hoặc bộ nhớ thể rắn bất kỳ. Các thiết bị I/O có thể bao gồm màn hình hiển thị, loa ngoài, micro và chuột di chuyển được mà cho máy tính biết vị trí tương đối của con trỏ trên màn hình và một hoặc nhiều nút mà có thể được kích hoạt để biểu thị lệnh.

Máy tính có thể hiển thị trên màn hình hiển thị mà được kết nối hoạt động đến mạch I/O về bề ngoài của giao diện sử dụng. Các hình dạng, văn bản khác nhau và các dạng đồ họa khác được hiển thị trên màn hình do máy tính tạo ra dữ liệu mà gây ra 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 các kết quả tạm thời được truyền đến máy chủ. These các kết quả tạm thời có thể được xử lý tại máy chủ và các kết quả cuối cùng trả về người dùng.

Máy tính có thể hiển thị trên màn hình hiển thị mà được kết nối hoạt động đến mạch I/O về bề ngoài của giao diện sử dụng. Các hình dạng, văn bản khác nhau và các dạng đồ họa khác được hiển thị trên màn hình do máy tính tạo ra dữ liệu mà gây ra các điểm ảnh bao gồm màn hình hiển thị có các màu sắc và sắc thái khác nhau. Giao diện sử dụng cũng hiển thị đối tượng đồ họa được gọi là con trỏ trong lĩnh vực kỹ thuật này. Vị trí của đối tượng trên màn hình hiển thị biểu thị cho người dùng sự lựa chọn đối tượng khác trên màn hình. Con trỏ có thể được di chuyển bởi người dùng bằng phương tiện của thiết bị khác được kết nối bằng mạch I/O đến máy tính. Thiết bị này phát hiện các chuyển động vật lý nhất định của người dùng, ví dụ, vị trí tay trên bề mặt phẳng hoặc vị trí của ngón tay trên bề mặt phẳng. Các thiết bị như này có thể được gọi trong lĩnh vực kỹ thuật này là chuột hoặc bàn di chuột cảm ứng. Trong một số phương án, bản thân màn hình hiển thị có thể hoạt động như bàn di chuột cảm ứng bằng các cảm ứng sự có mặt 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 dưới dạng nút bấm hoặc công tắc, người dùng có thể kích hoạt nút bấm hoặc công tắc bằng cách nhấn công tắc vật lý trên chuột hoặc bàn di chuột cảm ứng hoặc thiết bị máy tính hoặc ấn bàn di chuột cảm ứng hoặc chạm vào màn hình cảm ứng. Khi máy tính phát hiện là công tắc vật lý được nhấn (hoặc là việc ấn bàn di chuột cảm ứng hoặc chạm màn hình cảm ứng được thực hiện), nó lấy

vị trí thể hiện của con trỏ (hoặc trong trường hợp chạm màn hình cảm ứng, vị trí được phát hiện của ngón tay) trên màn hình và thực thi quá trình xử lý có liên quan đến vị trí đó. Ví dụ, không có ý định giới hạn phạm vi của sáng chế, đối tượng đồ họa mà xuất hiện là hộp 2 chiều với từ "đi vào" ở bên trong đó nó có thể được hiển thị trên màn hình. Nếu máy tính phát hiện là công tắc được nhấn trong khi con trỏ vị trí (hoặc vị trí ngón tay để chạm màn hình cảm ứng) mà nằm trong các đường biên của đối tượng đồ họa, ví dụ, hộp được hiển thị, máy tính sẽ thực thi quá trình xử lý có liên quan đến lệnh "đi vào". Theo cách này, các đối tượng đồ họa trên màn hình tạo ra giao diện sử dụng mà cho phép người dùng điều khiển các quá trình xử lý đang hoạt động trên máy tính.

Sáng chế cũng có thể được thực thi toàn bộ trên một hoặc nhiều máy chủ. Máy chủ có thể máy tính bao gồm bộ xử lý trung tâm với thiết bị lưu trữ chung và kết nối mạng. Ngoài ra máy chủ có thể bao gồm nhiều máy tính như này được kết nối với mạng dữ liệu hoặc kết nối chuyên dữ liệu khác, hoặc, nhiều máy tính trên mạng với bộ nhớ được truy cập mạng, theo cách cung cấp chức năng thành nhóm. Những người thực hiện có hiểu biết trung bình sẽ nhận ra là các chức năng mà được thực hiện trên một máy chủ có thể được phân vùng và được thực hiện trên nhiều máy chủ mà được kết nối hoạt động bởi mạng máy tính bằng giao tiếp xử lý thích hợp. Ngoài ra, việc truy cập trang web có thể bằng trình duyệt Internet mà truy cập trang công cộng hoặc có bảo mật hoặc bằng chương trình khách chạy trên máy tính cục bộ mà được kết nối trên mạng máy tính vào máy chủ. Tin nhắn dữ liệu và việc tải lên hoặc tải xuống dữ liệu có thể được chuyển qua Internet bằng cách sử dụng các giao thức đặc biệt, bao gồm TCP/IP, HTTP, TCP, UDP, SMTP, RPC, FTP hoặc các loại các giao thức giao tiếp dữ liệu khác mà cho phép các quá trình xử lý chạy trên hai máy tính điều khiển từ xa để trao đổi thông tin bằng giao tiếp mạng kỹ thuật số. Kết quả là tin nhắn dữ liệu có thể là gói dữ liệu được truyền từ hoặc được nhận bởi máy tính chứa địa chỉ mạng đích, quá trình xử lý đích hoặc mã định danh ứng dụng, và các giá trị dữ liệu mà có thể được phân tích cú pháp tại máy tính đích được nằm tại địa chỉ mạng đích bằng ứng dụng đích để các giá trị 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 vài cấp độ, sao cho máy tính của người dùng được kết nối thông qua tường lửa đến một máy chủ, mà giao tiếp định tuyến đến máy chủ khác thực thi các phương pháp của sáng chế.

Máy tính của người dùng có thể hoạt động chương trình mà nhận từ máy chủ điều khiển từ xa tệp tin dữ liệu mà được chuyển đến chương trình diễn dịch dữ liệu trong tệp tin dữ liệu và lệnh cho thiết bị hiển thị thể hiện cụ thể văn bản, các hình ảnh, video, âm thanh và các đối tượng khác. Chương trình có thể phát hiện vị trí tương đối của con trỏ khi nút chuột được kích hoạt, và diễn dịch câu lệnh để được thực thi dựa trên vị trí ở trên vị trí tương đối được biểu thị trên màn hình khi nút được ấn. Tệp tin dữ liệu có thể là tài liệu HTML, chương trình là chương trình trình duyệt web và lệnh là siêu liên kết mà làm cho trình duyệt yêu cầu tài liệu HTML mới từ vị trí địa chỉ mạng dữ liệu điều khiển từ xa khác. HTML cũng có thể có các tham chiếu mà dẫn đến các mô đun mã khác được gọi lên và được thực thi, ví dụ, Flash hoặc mã riêng khác.

Những người có hiểu biết trung bình trong lĩnh vực kỹ thuật này sẽ hiểu rõ là sáng chế có thể được áp dụng với các giao tiếp, quá trình xử lý dữ liệu, hoặc kết cấu hệ thống máy tính khác, bao gồm: các thiết bị không dây, các ứng dụng Internet, các thiết bị cầm tay (bao gồm các trợ giúp số cá nhân (PDAs)), máy tính có thể mang theo được, tất cả các thiết bị di động hoặc điện thoại di động, các hệ thống đa xử lý, các hệ thống được dựa trên vi xử lý hoặc các thiết bị điện tử tiêu dùng có thể lập trình được, thiết bị giải mã tín hiệu, các máy tính cá nhân nối mạng, máy vi tính, các máy tính lớn, và tương tự. Thật vậy, các thuật ngữ "máy tính," "máy chủ," và thuật ngữ tương tự được sử dụng có thể thay thế cho nhau ở đây, và có thể là thiết bị và hệ thống bất kỳ nêu trên.

Trong một số trường hợp, đặc biệt khi máy tính của người dùng là thiết bị điện toán di động được sử dụng để truy cập dữ liệu thông qua mạng, mạng này có thể là bất kỳ loại mạng di động, mạng viễn thông hội tụ hoặc dựa trên giao thức Internet (IP-Internet Protocol) nào, bao gồm nhưng không bị giới hạn ở Hệ thống thông tin di động toàn cầu (GSM- Global System for Mobile Communications), Đa truy cập phân chia theo thời gian (TDMA- Time Division Multiple Access), Đa truy cập phân chia theo mã (CDMA- Code Division Multiple Access), Đa truy cập phân chia theo tần số trực giao (OFDM- Orthogonal Frequency Division Multiple Access), Dịch vụ vô tuyến gói tổng hợp (GPRS- General Packet Radio Service), Môi trường GSM dữ liệu được nâng cao (EDGE- Enhanced Data GSM Environment), Hệ thống điện thoại di động tiên tiến (AMPS- Advanced Mobile Phone System), khả năng khai thác liên mạng trên toàn cầu đối với truy cập vi ba (WiMAX- Worldwide Interoperability for Microwave Access), Hệ thống viễn thông di động toàn cầu (UMTS- Universal Mobile Telecommunications

System), Phát triển – Tối ưu hóa Dữ liệu (EVDO- Evolution-Data Optimized), Tiến hóa dài hạn (LTE- Long Term Evolution), Mạng thông tin di động siêu băng rộng (UMB- Ultra Mobile Broadband), dịch vụ thoại thông qua giao thức Internet (VoIP- Voice over Internet Protocol), hoặc Truy cập di động không được đăng ký (UMA- Unlicensed Mobile Access).

Internet là mạng máy tính mà cho phép các khách hàng vận hành máy tính cá nhân để được tương tác với các máy chủ máy tính được đặt từ xa và để xem nội dung mà được chuyển từ các máy chủ đến máy tính cá nhân như các tệp tin dữ liệu qua mạng. Trong một loại giao thức, các máy chủ hiển thị các trang web mà được trả lại về máy tính cá nhân của khách hàng bằng cách sử dụng chương trình cục bộ đã được biết là trình duyệt. Trình duyệt nhận một hoặc nhiều tệp tin dữ liệu từ máy chủ mà được hiển thị trên màn hình máy tính cá nhân của khách hàng. Trình duyệt tìm kiếm các tệp tin dữ liệu này từ địa chỉ cụ thể, mà được biểu diễn bởi chuỗi chữ số được gọi là định vị tài nguyên thống nhất (URL- Universal Resource Locator). Tuy nhiên, trang web có thể chứa các thành phần mà được tải từ nhiều URL hoặc địa chỉ IP khác nhau. Trang web là tập hợp các URL có liên quan, thường chia sẻ cùng địa chỉ gốc hoặc dưới sự điều khiển của một vài thực thể. Theo một phương án, các vùng khác của các không gian được mô phỏng có các URL khác nhau. Nghĩa là, các không gian được mô phỏng có thể là cấu trúc dữ liệu đơn nhất, nhưng các URL khác nhau tham chiếu các vị trí khác nhau trong cấu trúc dữ liệu. Điều này khiến nó có thể mô phỏng vùng rộng và có những người tham gia bắt đầu sử dụng nó trong vùng lân cận ảo của họ.

Logic chương trình máy tính triển khai tất cả hoặc một phần chức năng được mô tả ở trên trong bản mô tả có thể được thể hiện dưới các dạng khác nhau, bao gồm, nhưng không bị giới hạn, dạng mã nguồn, dạng có thể thực thi được của máy tính, và các dạng trung gian khác nhau (ví dụ, các dạng được tạo ra trình hợp dịch, trình biên dịch, bộ liên kết, hoặc bộ định vị). Mã nguồn có thể bao gồm chuỗi các chỉ lệnh chương trình máy tính được triển khai bằng các ngôn ngữ lập trình khác nhau bất kỳ (ví dụ, mã đối tượng, hợp ngữ, hoặc ngôn ngữ cấp cao chẳng hạn 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 điều hành khác nhau. Mã nguồn có thể định ra và sử dụng các cấu trúc dữ liệu và tin nhắn giao tiếp khác nhau. Mã nguồn có thể dưới dạng có thể thực thi được trên máy tính (ví dụ, thông qua bộ diễn dịch), hoặc mã nguồn có thể được chuyển đổi

(ví dụ, thông qua bộ chuyển ngữ, trình hợp dịch, hoặc trình biên dịch) thành dạng có thể thực thi được trên máy tính.

Sáng chế có thể được mô tả trong ngữ cảnh chung của các chỉ lệnh có thể thực thi được trên máy tính, chẳng hạn như các môđun chương trình, được thực thi bởi máy tính. Nói chung, các môđun chương trình bao gồm các chương trình con, các chương trình, các đối tượng, các thành phần, các cấu trúc dữ liệu, v.v., mà thực hiện các tác vụ cụ thể hoặc triển khai các kiểu dữ liệu trừu tượng cụ thể. Chương trình máy tính và dữ liệu có thể được cố định dưới dạng bất kỳ (ví dụ, dạng mã nguồn, dạng có thể thực thi được trên máy tính, hoặc dạng trung gian) dài hạn hoặc ngắn hạn trong phương tiện lưu trữ hữu hình, chẳng hạn như thiết bị nhớ bán dẫn (ví dụ, RAM, ROM, PROM, EEPROM, hoặc RAM có thể lập trình nhanh), thiết bị nhớ từ tính (ví dụ, đĩa mềm hoặc ổ cứng cố định), thiết bị nhớ quang (ví dụ, CD-ROM hoặc DVD), thẻ PC (ví dụ, thẻ PCMCIA), hoặc thiết bị nhớ khác. Chương trình máy tính và dữ liệu có thể được cố định dưới bất kỳ dạng tín hiệu nào mà có thể truyền đến máy tính bằng cách sử dụng bất kỳ công nghệ giao tiếp khác nhau nào, bao gồm, nhưng không bị giới hạn ở, các công nghệ điện tử tương tự, các công nghệ số, các công nghệ quang, các công nghệ không dây, các công nghệ nối mạng, và các công nghệ liên mạng. Chương trình máy tính và dữ liệu có thể được phân phối dưới bất kỳ hình thức nào như phương tiện lưu trữ di động với tài liệu điện tử hoặc tài liệu in đi kèm (ví dụ, phần mềm nén gọn lại hoặc băng từ), được tải trước với hệ thống máy tính (ví dụ, lên hệ thống ROM hoặc ổ đĩa cố định), hoặc được phân phối từ máy chủ hoặc bảng thông báo điện tử trên hệ thống giao tiếp (ví dụ, Internet hoặc World Wide Web). Tốt hơn là bất kỳ thành phần phần mềm nào của sáng chế, nếu muốn, có thể được triển khai trong dạng ROM (bộ nhớ chỉ đọc). Nói chung, các thành phần phần mềm có thể được triển khai trong ổ cứng, nếu muốn, bằng cách sử dụng các kỹ thuật thông thường.

Sáng chế có thể cũng có thể được thực hiện trong môi trường điện toán được phân tán mà các tác vụ được thực hiện bởi các thiết bị xử lý từ xa mà được liên kết thông qua mạng giao tiếp. Trong môi trường điện toán được phân tán, các môđun chương trình có thể được nằm ở 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 hiểu biết trung bình sẽ nhận ra rằng sáng chế có thể được thực thi trên một hoặc nhiều bộ 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 quá trình xử lý có thể được thực thi bởi một hoặc nhiều máy tính và thiết bị lưu trữ tách biệt về mặt địa lý như được kết nối bởi mạng dữ liệu theo cách sao cho chúng hoạt động với nhau để thực thi các bước xử lý. Theo một phương án, máy tính của người dùng có thể chạy ứng dụng mà làm cho máy tính của người dùng truyền luồng dữ liệu của một hoặc nhiều gói dữ liệu qua mạng dữ liệu đến máy tính thứ hai, ở đây được gọi là máy chủ. Đến lượt máy chủ này có thể được kết nối đến một hoặc nhiều thiết bị lưu trữ dữ liệu chung, nơi mà cơ sở dữ liệu được lưu trữ. Máy chủ có thể thực thi chương trình mà nhận gói được truyền đến và diễn dịch các gói dữ liệu được truyền đến để trích xuất thông tin truy vấn cơ sở dữ liệu. Sau đó máy chủ có thể thực thi các bước còn lại của sáng chế bằng cách truy cập các thiết bị lưu trữ chung để rút ra kết quả mong muốn của truy vấn. Theo cách khác, máy chủ có thể truyền tải thông tin truy vấn đến máy tính khác mà được kết nối đến các thiết bị lưu trữ chung, và mà máy tính có thể thực thi sáng chế để rút ra kết quả mong muốn. Sau đó, kết quả có thể được truyền lại đến máy tính của người dùng bằng dòng dữ liệu khác của một hoặc nhiều gói dữ liệu được đánh địa chỉ thích hợp đến máy tính của người dùng. Theo một phương án, cơ sở dữ liệu liên quan có thể được chứa trong một hoặc nhiều máy chủ được kết nối hoạt động được kết nối hoạt động đến bộ nhớ máy tính, ví dụ, vác ổ đĩa. Theo phương án khác nữa, sự khởi tạo của cơ sở dữ liệu liên quan có thể được chuẩn bị trên tập hợp các máy chủ và sự tương tác với máy tính của người dùng xảy ra tại một nơi khác trong toàn bộ quá trình.

Cần phải lưu ý rằng các lưu đồ được sử dụng ở đây để minh họa các khía cạnh khác nhau của sáng chế, và không nên được hiểu là giới hạn sáng chế ở bất kỳ dòng logic hoặc triển khai logic cụ thể nào. Logic được mô tả có thể được phân vùng vào trong các khối logic khác nhau (ví dụ, các chương trình, các môđun, các chức năng, hoặc các đoạn chương trình con) mà không làm thay đổi các kết quả tổng thể hoặc không xa rời khỏi phạm vi của sáng chế. Thông thường, các phần tử logic có thể được bổ sung, được biến đổi, được bỏ qua, được thực hiện theo thứ tự khác, hoặc được triển khai bằng cách sử dụng các cấu trúc logic khác (ví dụ, các cổng logic, vòng lặp nguyên thủy, logic có điều kiện, và các cấu trúc logic khác) mà không làm thay đổi các kết quả tổng thể hoặc không xa rời khỏi phạm vi của sáng chế.

Các phương án được mô tả của sáng chế chỉ mang tính chất ví dụ và rất nhiều biến thể và thay đổi sẽ rõ ràng đối với những người có hiểu biết trung bình trong lĩnh

vực kỹ thuật này. Tất cả các biến thể và thay đổi như này đều nằm trong phạm vi của sáng chế như được định rõ trong các yêu cầu bảo hộ. Mặc dù sáng chế được mô tả và được minh họa chi tiết, nhưng cần phải hiểu rằng chúng chỉ mang tính chất ví dụ và minh họa, và không được xem là giới hạn của sáng chế. Tốt hơn là các dấu hiệu khác nhau của sáng chế mà, để rõ ràng, được mô tả trong ngữ cảnh của các phương án riêng cũng có thể được đề xuất kết hợp với một phương án đơn lẻ.

YÊU CẦU BẢO HỘ

1. Phương pháp phát hiện đối tượng trong video và so khớp đối tượng với một hoặc nhiều sản phẩm bao gồm các bước:

- a) thu nhận video và tự động trích xuất siêu dữ liệu và các thuộc tính của các đối tượng trong khung hình và/hoặc các phần của khung hình trong video;
- b) phân đoạn video dựa trên các cài đặt được mô tả và/hoặc các sự kiện bằng cách so sánh các nội dung của các khung hình liên tục về các điểm giống nhau và khác nhau, trong đó video được dẫn đi ngang lần lượt để phát hiện cặp khung hình hoặc chuỗi khung hình mà vi phạm ngưỡng đồng dạng, trong đó khung hình chính được định danh cho mỗi phân đoạn;
- c) biên dịch các phân đoạn của các cài đặt và/hoặc các sự kiện giống nhau hoặc tương tự nhau, trong đó mỗi phân đoạn được gắn thẻ với mã định danh phân đoạn;
- d) phân tích một hoặc nhiều phân đoạn để phát hiện một hoặc nhiều đối tượng, trong đó khung hình và/hoặc các phần của khung hình được so sánh với nội dung đã được xác định trong cơ sở dữ liệu được điền bằng cách căn chỉnh các đối tượng và các cụm siêu dữ liệu đã biết, trong đó siêu dữ liệu được liên kết với khung hình và/hoặc các phần của khung hình bởi mã định danh phân đoạn, trong đó vị trí của một hoặc nhiều đối tượng được phát hiện được thu trong mỗi khung hình chính, trong đó véc-tơ đặc trưng theo đối tượng được tạo ra cho mỗi phân đoạn;
- e) so sánh một hoặc nhiều đối tượng với các sản phẩm;
- f) định danh các sản phẩm có liên quan đến một hoặc nhiều đối tượng này, trong đó mạng nơ-ron tích chập (CNN- convolutional neural network) được sử dụng;
- g) thông báo cho một hoặc nhiều người xem các sản phẩm;

trong đó sự tăng cường màn hình thứ hai được sử dụng cho video trực tiếp hoặc trực tuyến.

2. Phương pháp theo điểm 1, trong đó cơ sở dữ liệu được điền với nội dung đã được xác định bằng cách sử dụng trình thu thập dữ liệu web.

3. Phương pháp theo điểm 1, trong đó bước thông báo cho một hoặc nhiều người xem các sản phẩm bao gồm việc hiển thị quảng cáo.

4. Phương pháp theo điểm 1, trong đó bước thông báo cho một hoặc nhiều người xem các sản phẩm bao gồm việc cung cấp đường dẫn siêu liên kết đến trang web hoặc video.

5. Phương pháp phát hiện một hoặc nhiều đối tượng trong ảnh chụp màn hình kỹ thuật số từ video và so khớp một hoặc nhiều đối tượng với tài liệu quảng cáo bao gồm các bước:

a) nhận sự truy vấn từ người xem dưới dạng ảnh chụp màn hình kỹ thuật số;

b) định danh một hoặc nhiều đối tượng trong ảnh chụp màn hình kỹ thuật số bằng cách so sánh ảnh chụp màn hình kỹ thuật số và/hoặc các phần của ảnh chụp màn hình kỹ thuật số với nội dung đã được xác định trong cơ sở dữ liệu được điền bằng cách căn chỉnh các đối tượng đã biết và các cụm siêu dữ liệu, trong đó ảnh chụp màn hình kỹ thuật số được so khớp với mã định danh phân đoạn từ video để truy hồi danh sách các đối tượng được liên kết với phân đoạn của video, trong đó phương pháp tiếp cận mô hình túi từ trực quan được sử dụng và N hình ảnh tiêu biểu hàng đầu được lọc và được bổ sung vào cơ sở dữ liệu;

c) so sánh một hoặc nhiều đối tượng với các sản phẩm;

d) so khớp các sản phẩm có liên quan đến một hoặc nhiều đối tượng, trong đó mạng nơ-ron tích chập (CNN) được sử dụng, trong đó sự xác minh về không gian được thực hiện lên các sản phẩm so khớp để xác nhận; và

e) liên hệ người xem với tài liệu quảng cáo liên quan đến các sản phẩm được so khớp;

trong đó sau khi nhận sự truy vấn, các cụm từ trực quan được gán.

6. Phương pháp theo điểm 5, trong đó cơ sở dữ liệu được điền với nội dung đã được xác định bằng cách sử dụng trình thu thập dữ liệu web.

7. Phương pháp theo điểm 5, trong đó sự tăng cường nội dung trên màn hình thứ hai được sử dụng cho video trực tiếp hoặc trực tuyến.

8. Phương pháp theo điểm 5, trong đó bước liên hệ người xem với tài liệu quảng cáo liên quan đến các sản phẩm được định danh bao gồm việc hiển thị quảng cáo và/hoặc cung cấp đường dẫn siêu liên kết đến trang web hoặc video.

9. Hệ thống tạo các mối quan hệ giữa các đối tượng trong video với các sản phẩm trong cơ sở dữ liệu của các sản phẩm bao gồm:

mạng và hệ thống được vi tính hóa để được trình bày cục bộ hoặc từ xa với người dùng hoặc các nhóm người dùng thông qua ứng dụng giao diện người dùng;

môđun phát hiện và lưu trữ nội dung phương tiện truyền thông cục bộ hoặc trên máy chủ;

môđun truyền tải nội dung phương tiện truyền thông đến bộ xử lý, điều khiển từ xa hoặc dựa trên máy chủ, để nhập siêu dữ liệu và/hoặc các đặc trưng trực quan;

môđun truyền tải nội dung phương tiện truyền thông đến bộ xử lý, điều khiển từ xa hoặc dựa trên máy chủ, để trích xuất siêu dữ liệu và các đặc trưng trực quan;

phương tiện nhận đầu vào từ một hoặc nhiều người dùng dưới dạng hình ảnh kỹ thuật số được rút ra từ video và bao gồm các đặc trưng trực quan;

môđun được định cấu hình để triển khai các việc sau:

định danh các đặc trưng trực quan trong hình ảnh kỹ thuật số và tương quan các đặc trưng trực quan với các đối tượng và/hoặc các nhóm sản phẩm liên quan trong cơ sở dữ liệu, trong đó cơ sở dữ liệu được điền với nội dung đã được xác định bằng cách căn chỉnh các đối tượng đã biết và/hoặc các nhóm sản phẩm liên quan và các cụm siêu dữ liệu, trong đó phương pháp tiếp cận mô hình túi từ trực quan được sử dụng để định danh các đặc trưng trực quan và tương quan các đặc trưng trực quan với các đối tượng và/hoặc các nhóm sản phẩm liên quan trong cơ sở dữ liệu;

trong đó hình ảnh kỹ thuật số được so khớp với mã định danh phân đoạn để truy hồi danh sách các đối tượng và/hoặc các nhóm sản phẩm liên quan được liên kết với phân đoạn của video; và

phân tích các đặc trưng trực quan và siêu dữ liệu của hình ảnh kỹ thuật số để tương quan các đặc trưng trực quan với các đối tượng và/hoặc các nhóm sản phẩm liên quan bằng cách sử dụng mạng nơ-ron tích chập (CNN), trong đó sự xác minh về không gian được thực hiện trên các đối tượng so khớp và/hoặc các nhóm sản phẩm liên quan để xác minh, và

dịch vụ mạng mà phân phối thông tin trên các đối tượng và/hoặc các nhóm sản phẩm liên quan cho người dùng và/hoặc các nhóm người dùng;

trong đó sau khi nhận đầu vào, các cụm từ trực quan được gán.

10. Hệ thống theo điểm 9, trong đó thông tin trên các đối tượng và/hoặc các nhóm sản phẩm liên quan bao gồm các quảng cáo.

11. Hệ thống theo điểm 9, trong đó thông tin trên các đối tượng và/hoặc các nhóm sản phẩm liên quan bao gồm đường dẫn siêu liên kết hoặc nội dung có thể truy cập thông qua internet.

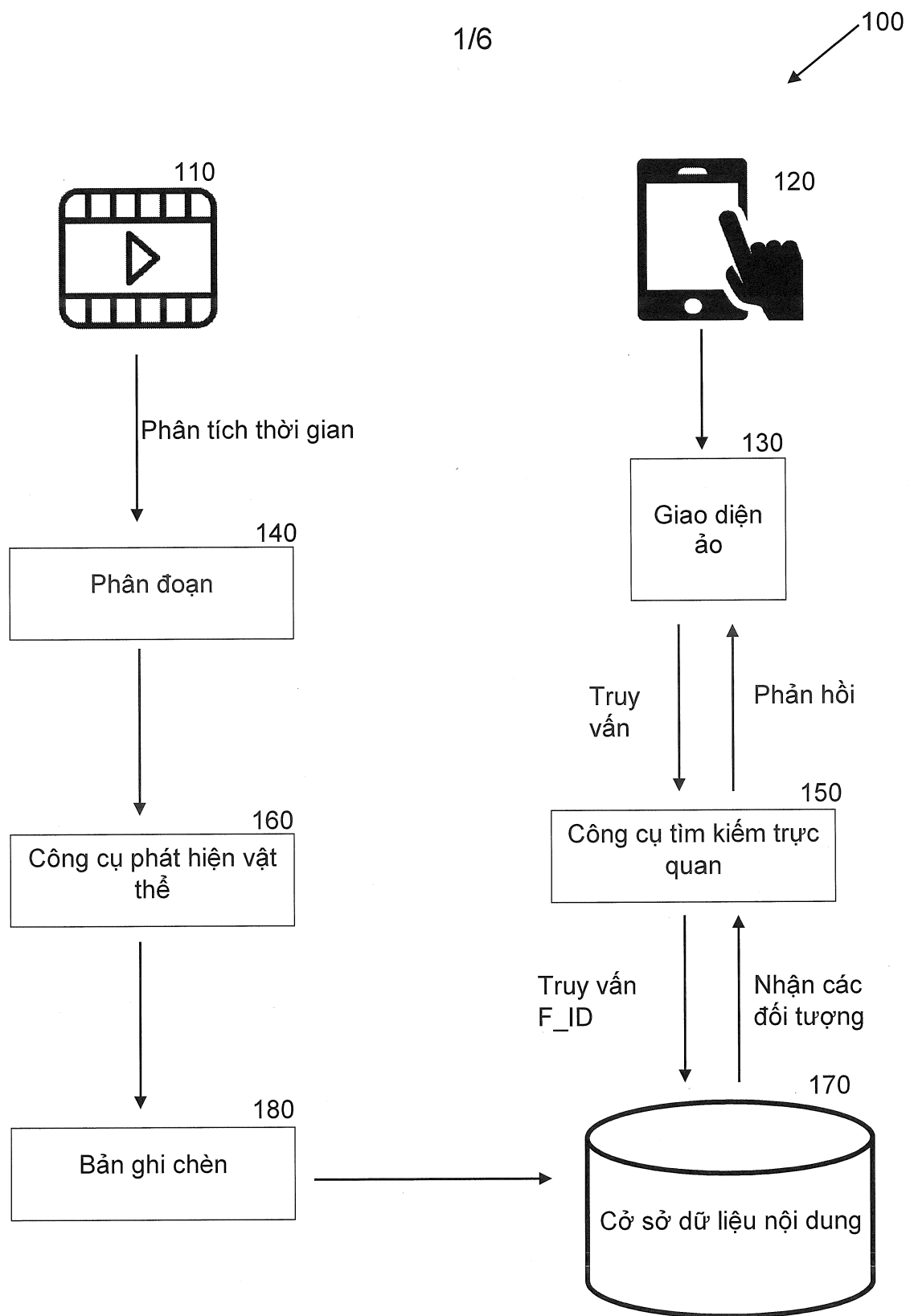


FIG. 1

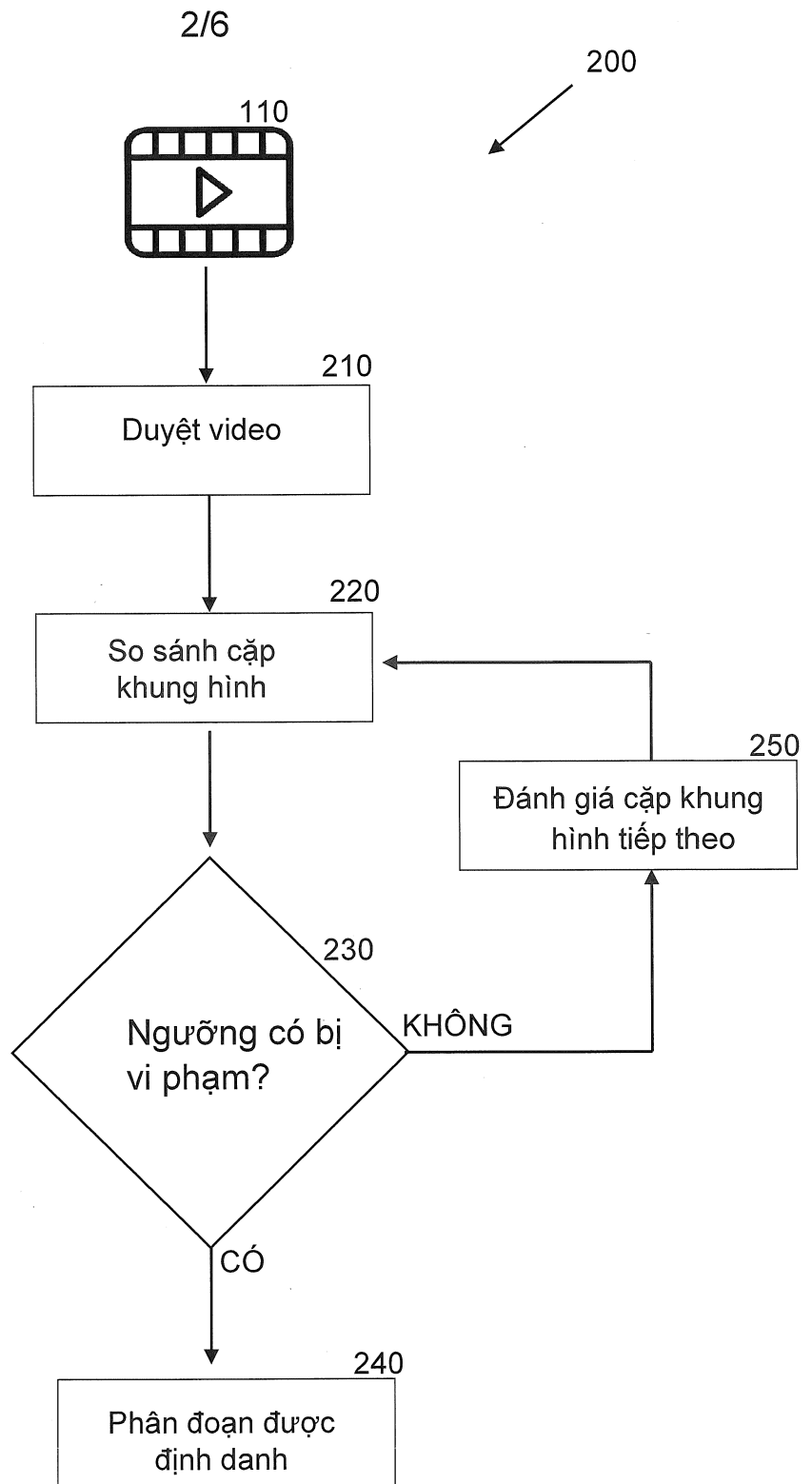


FIG. 2

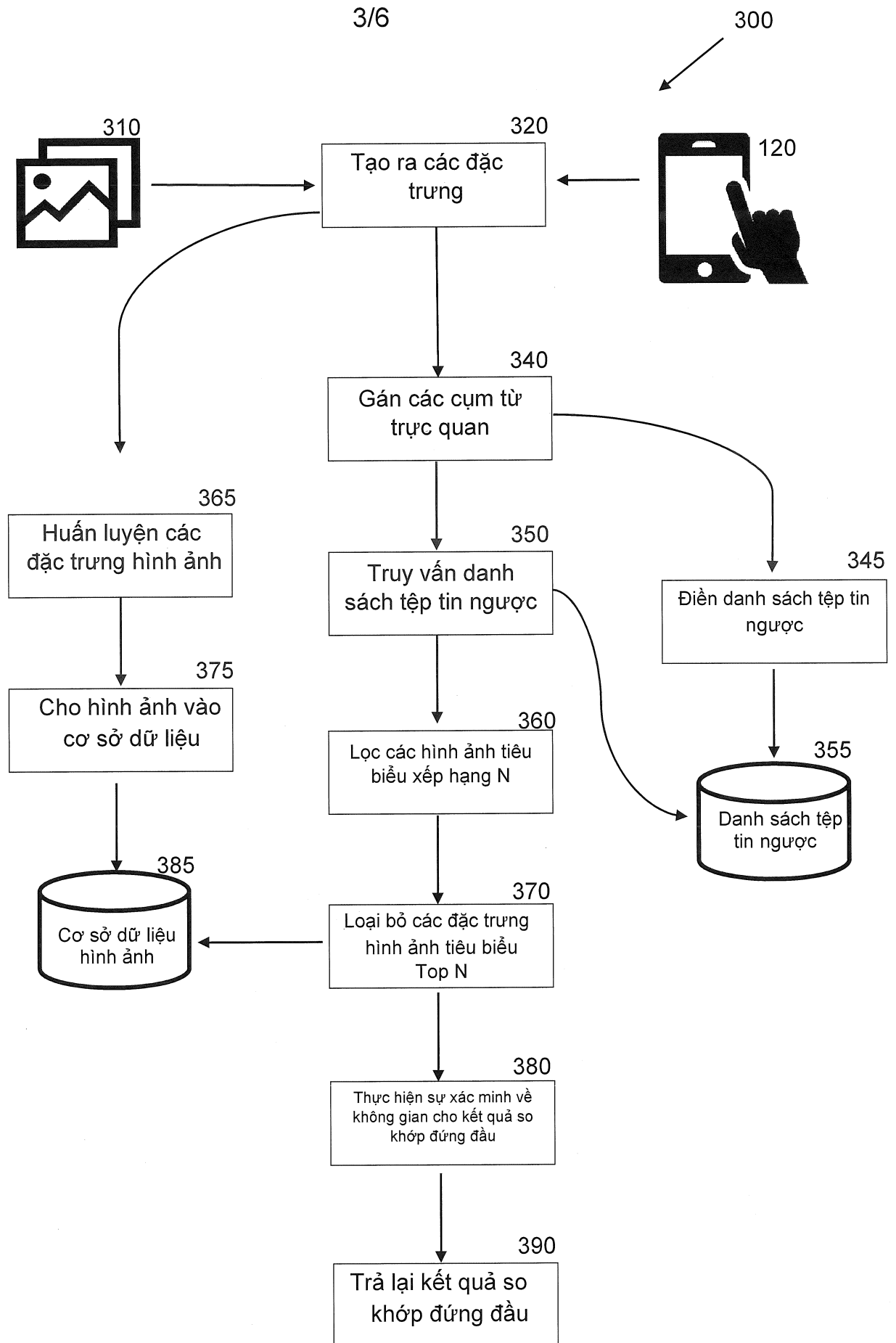


FIG. 3

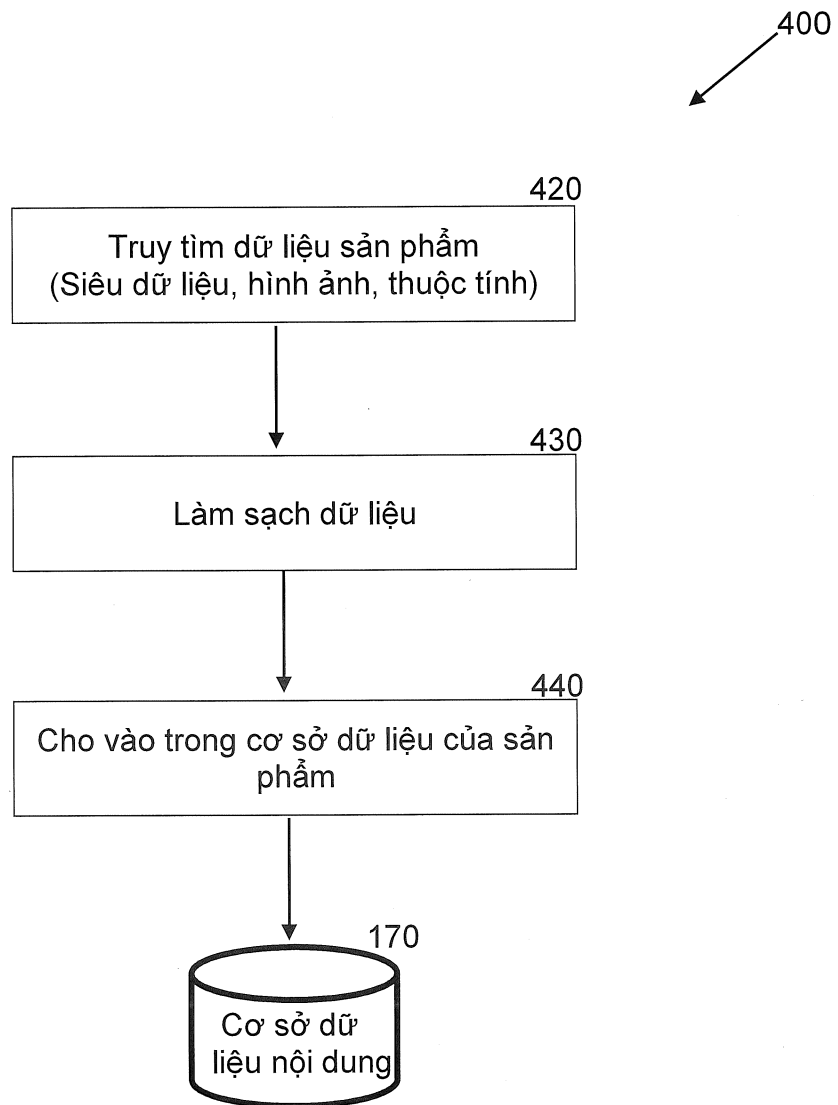


FIG. 4

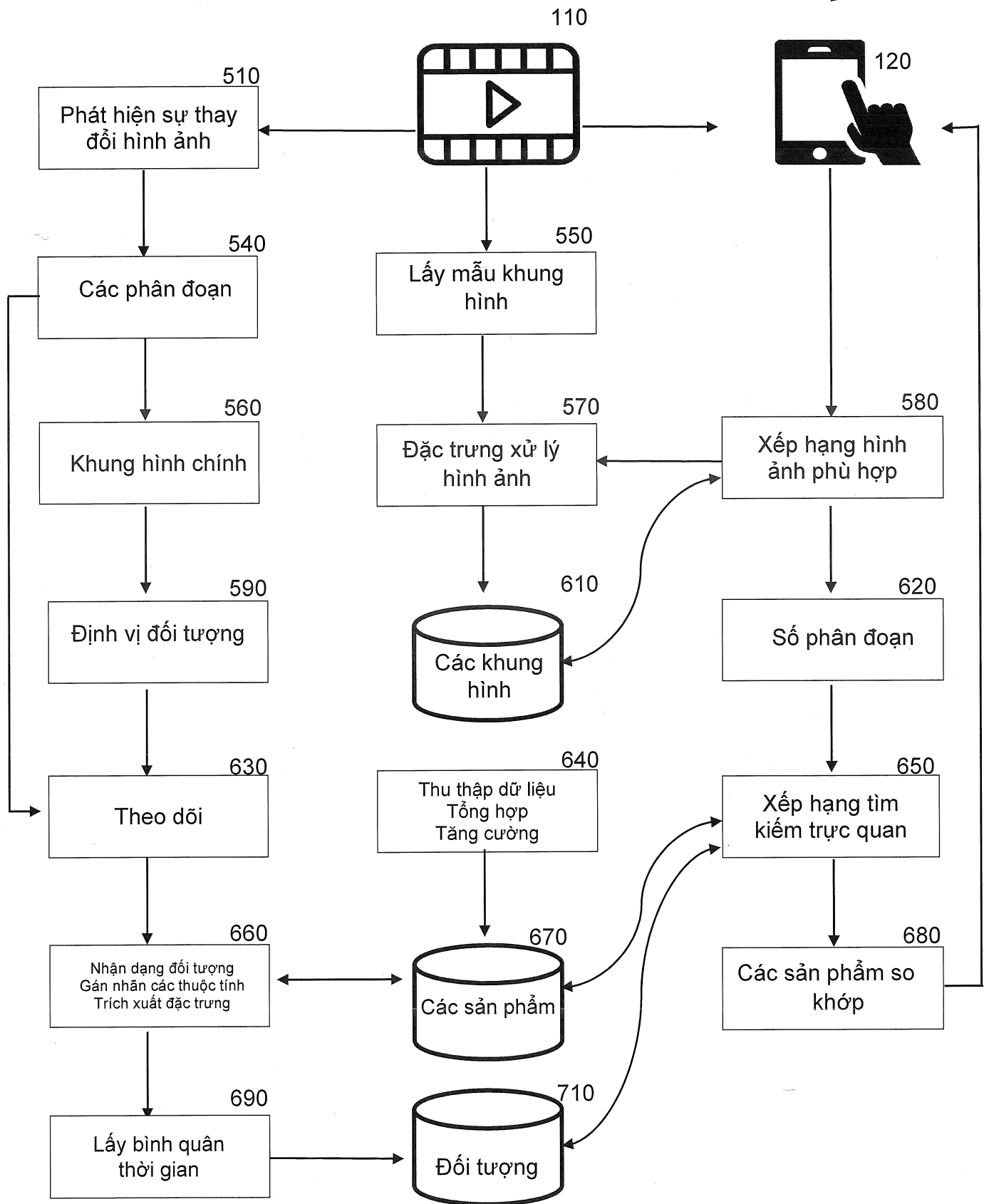


FIG. 5

6/6

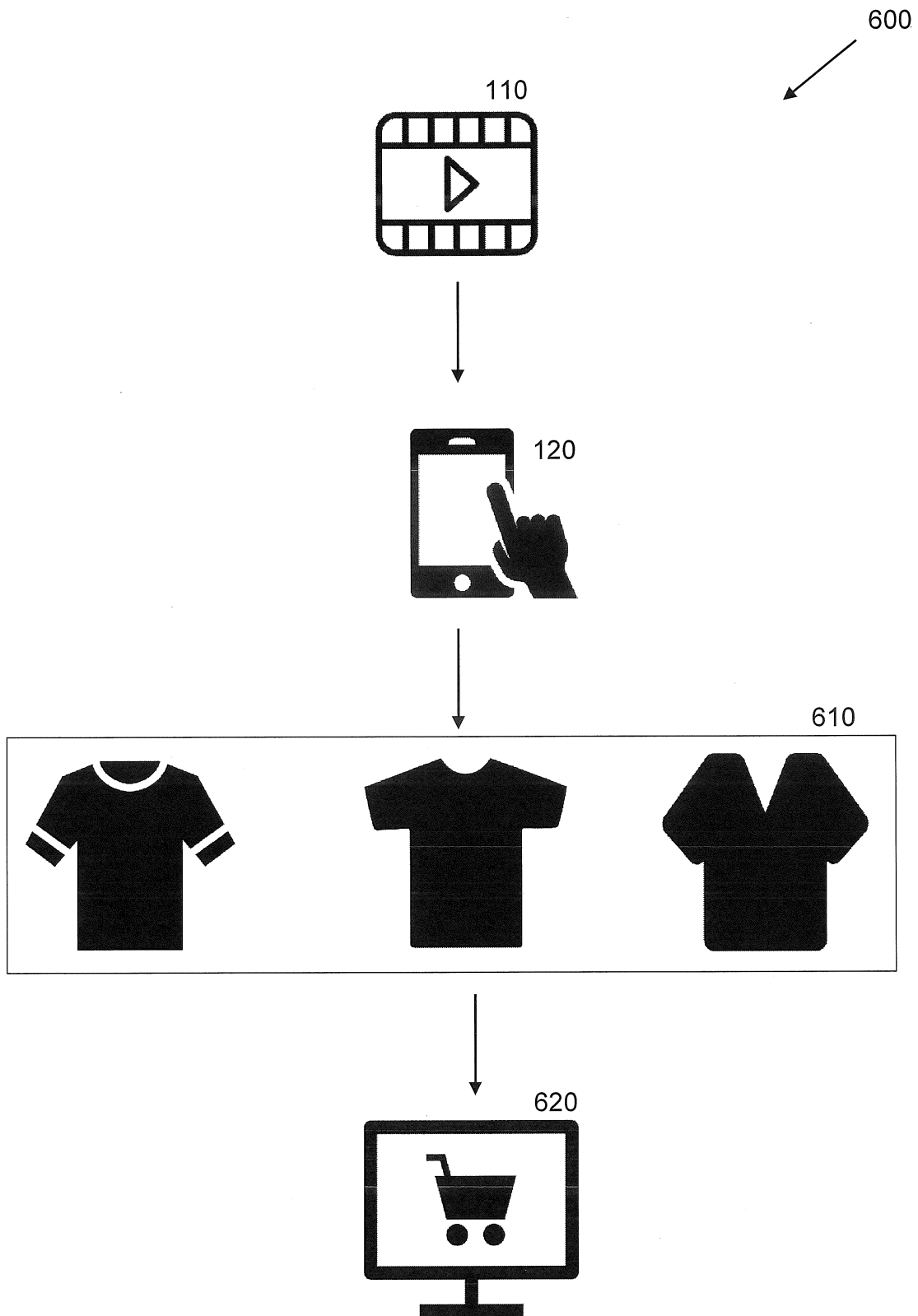


FIG. 6
-38-