



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

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



1-0026459

(51)⁷ G06K 9/00

(13) B

(21) 1-2015-04712

(22) 09/12/2015

(45) 25/11/2020 392

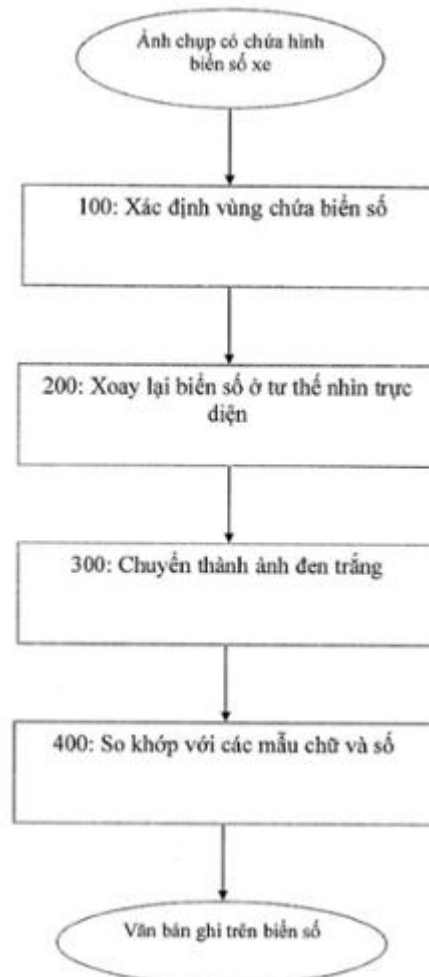
(43) 25/02/2016 335A

(73) Viện Nghiên cứu Công nghệ FPT - Trường đại học FPT (VN)
Số 8 Tôn Thất Thuyết, Mỹ Đình 2, Nam Từ Liêm, thành phố Hà Nội

(72) Phạm Bảo Thạch (VN); Hoàng Anh Tuấn (VN); Trần Nguyên Ngọc (VN); Đặng Minh Đức (VN); Từ Minh Phương (VN); Vũ Thanh Hải (VN).

(54) QUY TRÌNH XÁC ĐỊNH BIỂN SỐ XE

(57) Sáng chế đề xuất quy trình nhận đầu vào là ảnh chụp có chứa biển số xe và cho ra văn bản ghi trên biển số trong ảnh chụp đầu vào. Quy trình gồm bước xác định vùng chứa biển số trên ảnh, rồi biến đổi hình ảnh biển số về hình dạng chữ nhật, biến ảnh này thành ảnh đen trắng, và so khớp ảnh thu được với các mẫu hình ảnh của những ký tự có thể xuất hiện trên biển số, để xác định ký tự nào xuất hiện ở vị trí nào trên ảnh biển số. Quy trình tránh việc phải cắt hình ảnh biển số thành những mảnh nhỏ chứa các hình ảnh ký tự, do đó giúp làm tăng độ chính xác của việc xác định văn bản ghi trên biển số.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến quy trình nhận đầu vào là ảnh chụp có chứa biển số xe và cho ra văn bản ghi nội dung biển số trong ảnh chụp đầu vào.

Tình trạng kỹ thuật của sáng chế

Tự động xác định được nội dung văn bản ghi trên biển số là một nhu cầu phổ biến ở trong các hệ thống quản lý phương tiện giao thông trên đường. Chẳng hạn, ở các hệ thống trạm thu phí, hình ảnh camera lắp đặt tại trạm thu phí có thể chụp được ảnh biển số xe và từ ảnh này có thể tự động xác định văn bản ghi trên biển số đã được chụp.

Có nhiều sáng chế đã đề xuất những quy trình để xác định được văn bản ghi trên biển số xe, từ ảnh chụp các biển số này. Đa phần các cách tiếp cận để xác định được văn bản ghi trên biển số xe đã được đề xuất đều dựa trên một quy trình gồm hai bước quan trọng là:

thứ nhất, cắt hình chụp biển số thành từng mảnh, mỗi mảnh chứa hình chụp một ký tự;

với mỗi mảnh hình chụp một ký tự, sinh ra những đặc trưng ở mức độ trừu tượng hoá cao, và phân loại chúng thuộc về các loại ký tự nào.

Ví dụ bài báo Artificial neural networks-based vehicle license plate recognition của tác giả Kocer và Cevik in trong tạp chí Procedia Computer Science năm 2011, đề xuất sử dụng mạng nơ ron nhân tạo để phân loại các phần hình ảnh chụp ký tự của biển số về loại ký tự nào.

Tuy độ chính xác trong bước phân loại có thể cao, khó khăn lớn nhất của các cách tiếp cận nêu trên là việc cắt hình chụp biển số thành những mảnh chứa hình chụp các ký tự thường đạt độ chính xác không cao, dễ để lọt nét thừa không thuộc

ký tự vào mảnh chụp hình ký tự, hoặc để mất nét thuộc ký tự ra ngoài mảnh cắt. Điều này dẫn đến độ chính xác tổng thể của việc nhận dạng biển số bị giới hạn.

Bản chất kỹ thuật của sáng chế

Sáng chế đề xuất quy trình nhận đầu vào là ảnh chụp có chứa biển số xe và cho ra văn bản ghi nội dung trên biển số trong ảnh chụp đầu vào, trong đó bỏ qua bước cắt hình chụp thành từng mảnh chứa hình ảnh của các ký tự, do đó vượt qua được giới hạn về độ chính xác của việc nhận dạng văn bản trên biển số.

Cụ thể, quy trình nhận đầu vào là ảnh chụp có chứa biển số xe và cho ra văn bản ghi trên biển số trong ảnh chụp đầu vào, gồm các bước:

xác định vùng ảnh chứa biển số, gọi là vùng V, theo phương pháp như mô tả trong bài báo Rapid Object Detection using a Boosted Cascade of Simple Features của Paul Viola và Michael Jones in trong kỷ yếu hội nghị Computer Vision and Pattern Recognition năm 2001, sử dụng kỹ thuật học máy có giám sát, mô hình phân tầng với thuật toán huấn luyện tự động thích nghi dựa vào mô hình thống kê được xây dựng từ các đặc trưng tính toán bởi chênh lệch giá trị tổng cường độ sáng các vùng trên các ảnh chụp biển số đã biết vùng chứa biển số;

tạo mô hình biển số là tứ giác có bốn góc, và biến đổi mô hình này theo phương pháp chủ động mô tả hình dạng đối tượng (Active Shape Model) như mô tả trong bài báo Active shape models - their training and application của các tác giả Cootes, Taylor, Cooper và Graham in trong tạp chí Computer Vision and Image Understanding năm 1995, cho đến khi bốn góc của mô hình khớp với hình ảnh bốn góc biển số trong vùng V; rồi chuyển đổi vùng V, theo phương pháp chủ động mô tả hình dạng đối tượng, về hình ảnh mới, gọi là ảnh B, có bốn góc của biển số sau khi chuyển đổi trở thành bốn đỉnh của một hình chữ nhật;

ánh xạ giá trị màu đỏ, xanh dương và xanh lá cây ở mỗi điểm ảnh, trên ảnh B , đến một số thực, gọi là độ sáng điểm ảnh, theo một hàm ánh xạ nhất định, để thu được ảnh biến số xám hoặc đen trắng, gọi là ảnh X ;

với từng mẫu hình ảnh xám, hoặc đen trắng, của những ký tự có thể xuất hiện trên biến số, thực hiện tích chập giữa mẫu và ảnh X , và ghi nhận lại giá trị cực đại của tích chập tính được, cũng như vị trí trên ảnh X ứng với giá trị cực đại của tích chập; rồi chọn ra những mẫu có giá trị tích chập cao nhất và cao hơn một ngưỡng nhất định, và đặt các ký tự ứng với các mẫu đã chọn tại các vị trí trên ảnh X ứng với các giá trị cực đại của tích chập của các mẫu đã chọn, để thu được văn bản đầu ra.

Mô tả vắn tắt các hình vẽ

Hình 1 là sơ đồ khối của quy trình nhận đầu vào là ảnh chụp có chứa biến số xe và cho ra văn bản ghi trên biến số trong ảnh chụp đầu vào.

Mô tả chi tiết sáng chế

Hình 1 là sơ đồ khối của quy trình nhận đầu vào là ảnh chụp có chứa biến số xe và cho ra văn bản ghi trên biến số trong ảnh chụp đầu vào. Quy trình này gồm có các bước:

bước 100, xác định giới hạn của vùng trên ảnh chụp có chứa hình ảnh biến số xe;

bước 200, biến đổi hình ảnh vùng chứa biến số xe về hình ảnh mới thể hiện biến số xe được chụp chính diện;

bước 300, chuyển hình ảnh thu được ở bước trên về dạng ảnh xám hoặc đen trắng;

bước 400, lấy những hình ảnh mẫu của các ký tự để so sánh với từng cửa sổ trên hình ảnh thu được ở bước trên, và chọn ra những hình ảnh mẫu khớp nhất, ở những vị trí cửa sổ khớp nhất trên hình ảnh thu được ở bước trên, dựng lại thành văn bản ghi trên biến số.

Cụ thể việc thực hiện các bước nêu trên được trình bày ở dưới đây.

Ở bước 100, xác định vùng ảnh chứa biển số theo phương pháp như mô tả trong bài báo Rapid Object Detection using a Boosted Cascade of Simple Features của Paul Viola và Michael Jones in trong kỷ yếu hội nghị Computer Vision and Pattern Recognition năm 2001. Phương pháp này sử dụng kỹ thuật học máy có giám sát, mô hình phân tầng với thuật toán huấn luyện tự động thích nghi dựa vào mô hình thống kê được xây dựng từ các đặc trưng tính toán bởi chênh lệch giá trị tổng cường độ sáng các vùng trên các ảnh chụp biển số đã biết vùng chứa biển số.

Vùng ảnh chứa biển số, đầu ra của bước 100, được cho thành đầu vào của bước 200. Do ảnh chụp biển số có thể từ nhiều góc quay camera khác nhau, biển số thường không ở tư thế chính diện với camera. Do đó ảnh thu được của biển số thường bị méo mó, phần nằm gần camera sẽ to hơn phần nằm xa camera. Bước 200 có mục tiêu là biến đổi ảnh chụp biển số ở dạng méo mó như vậy về thành dạng giống như ảnh chụp chính diện của biển số.

Ở bước 200, trước tiên xác định những điểm nổi bật trên vùng ảnh chứa biển số. Ví dụ có những điểm nổi bật nằm ở bốn góc và bốn cạnh của biển số. Với một ảnh chụp biển số ở tư thế quay nghiêng, phần biển số nằm gần camera to hơn phần biển số nằm xa camera, bốn điểm ở bốn góc biển số tạo thành một tứ giác không phải là hình chữ nhật chuẩn. Những điểm này có thể được xác định bằng cách dựng một mô hình biển số gồm bốn điểm nằm ở bốn góc, rồi biến đổi mô hình này cho đến khi khớp với hình ảnh biển số, theo phương pháp chủ động mô tả hình dạng đối tượng như mô tả trong bài báo Active shape models - their training and application của các tác giả Cootes, Taylor, Cooper và Graham in trong tạp chí Computer Vision and Image Understanding năm 1995. Sau khi đã xác định xong những điểm nổi bật trên vùng ảnh chứa biển số, như trên, chuyển đổi ảnh về hình giống như chụp chính diện biển số, vẫn bằng phương pháp chủ động mô tả hình dạng đối tượng. Chẳng hạn, việc chuyển đổi thực hiện theo tiêu chí bốn điểm nổi bật ở bốn góc của biển số sau khi chuyển đổi trở thành bốn đỉnh của một hình chữ nhật.

Bước 300 được thực hiện bằng cách ánh xạ giá trị màu đỏ, xanh dương và xanh lá cây ở mỗi điểm ảnh, trên hình ảnh thu được sau bước 200, đến một số nguyên 0 hoặc 1, gọi là độ sáng điểm ảnh với 0 là đen và 1 là trắng, theo một hàm ánh xạ nhất định.

Ở bước 400, cần chuẩn bị sẵn những mẫu hình ảnh đen trắng của những ký tự có thể xuất hiện trên biển số, rồi với từng mẫu này thực hiện tích chập giữa mẫu và vùng hình ảnh thu được sau bước 300, và ghi nhận lại giá trị cực đại của tích chập tính được, cũng như vị trí trên vùng hình ảnh thu được sau bước 300 ứng với giá trị cực đại của tích chập. Sau khi đã thực hiện như trên cho mọi mẫu, chọn ra những mẫu có giá trị tích chập cao nhất và cao hơn một ngưỡng nhất định, và đặt các ký tự ứng với các mẫu đã chọn tại các vị trí trên vùng hình ảnh thu được sau bước 300 ứng với các giá trị cực đại của tích chập của các mẫu đã chọn, để thu được văn bản đầu ra. Chẳng hạn, trên vùng hình ảnh thu được sau bước 300 có chiều rộng 250 điểm ảnh và chiều cao 100 điểm ảnh, nếu xác định được mẫu ứng với ký tự số tám, số sáu và số chín có cực đại tích chập cao nhất và cao hơn ngưỡng đã đặt trước, tại các tọa độ $\{49,51\}$, $\{147,50\}$, $\{204,47\}$, thì văn bản thu được là “689”.

Yêu cầu bảo hộ

1. Quy trình xác định biển số xe, gồm các bước:

xác định vùng ảnh chứa biển số, gọi là vùng V bằng cách duyệt các vùng cửa sổ chữ nhật trên ảnh, với mỗi vùng xác định sự tồn tại các đặc trưng nhất định, mỗi đặc trưng là tổ hợp các vùng hình chữ nhật màu đen và trắng nằm kề nhau, màu của hình chữ nhật được xác định bằng cách tính toán chênh lệch giá trị tổng cường độ sáng của điểm ảnh thuộc hình chữ nhật, nếu chênh lệch lớn hơn một ngưỡng cho trước thì hình chữ nhật là màu trắng, nếu nhỏ hơn thì hình chữ nhật là màu đen; nếu tồn tại đủ các đặc trưng cần thiết thì cửa sổ chữ nhật được coi là vùng V ; việc xác định các đặc trưng nào là cần thiết cho vùng V được thực hiện bằng quá trình học, trong đó dữ liệu học là các vùng cửa sổ có chứa biển số và các vùng cửa sổ không chứa biển số, các đặc trưng chỉ xuất hiện trong vùng chứa biển mà không xuất hiện trong vùng không chứa biển được lưu lại và sử dụng làm đặc trưng cho quá trình xác định vùng V ;

tạo mô hình biển số là tứ giác có bốn góc, và biến đổi mô hình này theo phương pháp chủ động mô tả hình dạng đối tượng, theo đó việc tìm kiếm đường viền của biển số xe bằng cách lặp lại nhiều lần thao tác tìm một vị trí tốt hơn của một điểm neo bất kỳ trên đường viền, vị trí tốt hơn là vị trí mà tại đó nội dung của ảnh trùng khớp hơn với hình mẫu của điểm đó, hình mẫu của mỗi điểm neo được sinh ra từ việc huấn luyện: cho máy quan sát nhiều ảnh mà vị trí điểm neo được chỉ ra bằng người, sau đó máy học sẽ sinh ra hình mẫu của các điểm neo và đánh giá được độ khớp của một vùng ảnh bất kỳ với một mẫu điểm neo, độ trùng hợp càng cao thì nghĩa là vị trí điểm neo tìm được càng đúng, quá trình tìm điểm neo phải tuân theo một luật chi cho phép tìm một điểm neo cụ thể ở một số vị trí xác định quanh vị trí hiện tại của nó, luật này do máy học được trong quá trình huấn luyện ở trên, quá trình tìm kiếm đường viền dừng lại khi không thể tìm ra bất kỳ điểm neo tốt hơn nào nữa và đảm bảo khi bốn góc của mô hình khớp với hình ảnh bốn góc biển số trong

vùng V ; rồi chuyển đổi vùng V , theo phương pháp chủ động mô tả hình dạng đối tượng, về hình ảnh mới, gọi là ảnh B , có bốn góc của biên số sau khi chuyển đổi trở thành bốn đỉnh của một hình chữ nhật;

ánh xạ giá trị màu đỏ, xanh dương và xanh lá cây ở mỗi điểm ảnh, trên ảnh B , đến một số thực, gọi là độ sáng điểm ảnh, theo một hàm ánh xạ nhất định, để thu được ảnh biên số xám hoặc đen trắng, gọi là ảnh X ;

với từng mẫu hình ảnh xám, hoặc đen trắng, của những ký tự có thể xuất hiện trên biên số, thực hiện tính tích chập giữa mẫu và ảnh X , và ghi nhận lại giá trị cực đại của tích chập tính được, cũng như vị trí trên ảnh X ứng với giá trị cực đại của tích chập; rồi chọn ra những mẫu có giá trị tích chập cao nhất và cao hơn một ngưỡng nhất định, và đặt các ký tự ứng với các mẫu đã chọn tại các vị trí trên ảnh X ứng với các giá trị cực đại của tích chập của các mẫu đã chọn, để thu được văn bản đầu ra.

Hình 1